Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

Structure-activity relationship studies of the aromatic positions in cyclopentapeptide CXCR4 antagonists†

Jignesh Mungalpara,‡[a] Zack G. Zachariassen,‡[a] Stefanie Thiele,[b] Mette M. Rosenkilde,[b] and Jon Våbenø*[a]

Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX
DOI: 10.1039/b000000x

The cyclopentapeptide CXCR4 antagonist FC131 (cyclo(-Arg^1-Arg^2-2-Nal^3-Gly^4-D-Tyr^5-), **2**; 2-Nal = 3-(2-naphthyl)alanine) represents an excellent starting point for development of novel drug-like ligands with therapeutic potential in HIV, cancer, stem-cell mobilization, inflammation, and autoimmune diseases. While the structure-activity relationships for Arg^1, Arg^2, and Gly^4 are well established, less is understood about the roles of the aromatic residues 2-Nal^3 and D-Tyr^5. Here we report further structure-activity relationship studies of these two positions, which showed that (i) the distal aromatic ring of the 2-Nal^3 side chain is required in order to maintain high potency, and (ii) replacement of D-Tyr^5 with conformationally constrained analogues results in significantly reduced activity. However, a simplified analogue that contained Gly instead of D-Tyr^5 was only 13-fold less potent than **2**, which means that the D-Tyr^5 side chain is dispensable. These findings were rationalized based on molecular docking, and the collective structure-activity data for the cyclopentapeptides suggest that appropriately designed Arg^2-2-Nal^3 dipeptidomimetics have potential as CXCR4 antagonists.

Introduction

By now, the role of the G protein-coupled C-X-C chemokine receptor 4 (CXCR4) in HIV, cancer, stem-cell mobilization, inflammation, and autoimmune diseases is well established,[1] and several different antagonists for CXCR4 – both peptides and non-peptides – have been described in the literature over the last two decades.[2] The prototype non-peptide antagonist plerixafor (AMD3100), which is administered by subcutaneous injection, was approved for stem-cell mobilization in 2008 and is currently the only marketed CXCR4 antagonist. The molecular pharmacology of AMD3100[3, 4] and the structurally related non-peptide antagonists AMD3465[5] and AMD11070[6] has been extensively characterized.[7]

The majority of the reported peptide antagonists has been developed by Fujii and co-workers, starting from the 18-mer synthetic polyphemusin II analogue T22.[8] Extensive structure-activity relationship (SAR) and downsizing studies first led to the potent 14-mer antagonist T140 (**1**, Fig. 1),[9] and eventually to the discovery of the cyclopentapeptide FC131 (**2**, Fig. 1),[10] which we are currently using as lead compound for our ongoing efforts toward peptidomimetic CXCR4 antagonists. A 16-mer analogue of **1** that contains two additional C-terminal residues (CVX15, **3**, Fig. 1) was recently reported by Wu et al. as the ligand in an X-ray co-crystal structure of CXCR4.[11] Interestingly, the potent peptide antagonists **1**–**3** share an Arg^1-Arg^2-(n-Nal)^3-Xaa^4-L-/D-Tyr^5 pentapeptide motif (Fig. 1), a notable difference being that **1** and **2** both contain 3-(2-naphthyl)alanine (2-Nal) in position 3 while **3** contains the isomeric 3-(1-naphthyl)alanine (1-Nal).

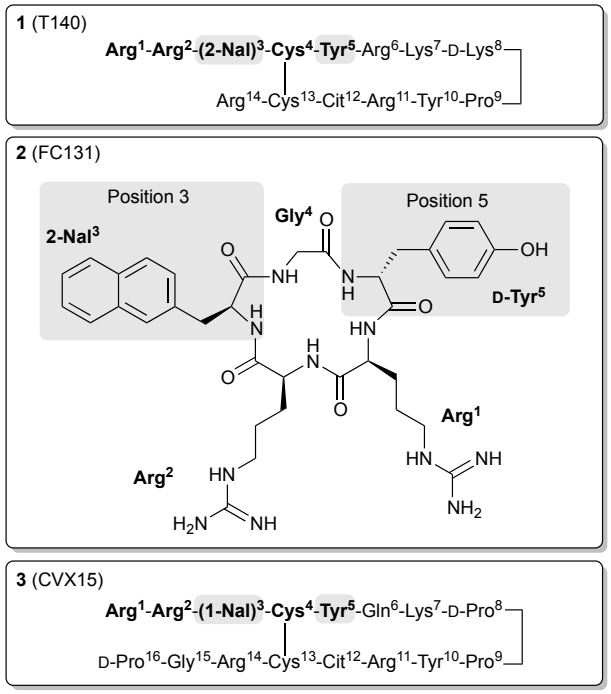

Figure 1 Structure of the lead cyclopentapeptide CXCR4 antagonist FC131 (**2**) and sequences of the larger peptide antagonists T140 (**1**) and CVX15 (**3**). The conserved Arg^1-Arg^2-(n-Nal)^3-Xaa^4-L-/D-Tyr^5 pentapeptide motif is shown in bold, and the aromatic positions 3 and 5 are highlighted with grey background. Cit = citrulline; 1-Nal = 3-(1-naphthyl)alanine; 2-Nal = 3-(2-naphthyl)alanine.

Aromatic rings and charged groups in ligand side chains have consistently been found to play a special role in binding and activation of peptidergic GPCRs.[12] Several SAR studies of the cyclopentapeptide **2** (Fig. 1) have been reported, and the relative importance of the two arginine residues is now well established: Arg^2 is crucial and serves as the anchor point for receptor binding, while Arg^1 plays a less important role.[13-15] The Gly residue in position 4 was originally introduced for synthetic reasons,[10] but a subsequent SAR study showed that the activity was reduced when Gly^4 was replaced with α-substituted nonpolar residues;[16] thus, the conformational and/or steric properties of Gly^4 are beneficial. In contrast, existing SAR data for the two aromatic residues 2-Nal^3 and D-Tyr^5 are less informative. Importantly, in a previously reported alanine scan of **2**, the Ala^3 and D-Ala^5 analogues were both classified as inactive (IC_{50} > 1 μM in both cases).[14] Consequently, the side chains of 2-Nal^3 and D-Tyr^5 were both considered as pharmacophoric elements for cyclopentapeptide CXCR4 antagonists. However, the roles and relative importance of the 2-Nal^3 and D-Tyr^5 side chains have been unclear, which has led to some ambiguity in pharmacophore definitions, i.e. whether one or both aromatic side chains are required for antagonistic activity.[17, 18]

On this background, we have performed further SAR studies of the two aromatic positions in the lead cyclopentapeptide **2**. Based on the nature of the already existing SAR data, two different approaches were used: for position 3 a classical SAR study, and for position 5 the introduction of conformational rigidity/flexibility. Here we report the findings of these SAR studies, an interpretation of the data based on molecular docking, and the implications for design of novel peptidomimetic CXCR4 antagonists.

Results and discussion

Design and SAR for position 3

In addition to the Ala^3 analogue that was classified as inactive,[14] only five analogues have previously been reported for position 3 (2-Nal). Inversion of stereochemistry (D-2-Nal^3) resulted in more than 25-fold reduction in affinity,[10] which shows that L-configuration is optimal in this position. Substitution with tryptophan (Trp^3) or a sulphur-containing Trp-analogue (3-(benzothiazol-2-yl)alanine; Bth^3) were in both cases shown to result in good activity,[14] which could be expected based on the structural similarity with 2-Nal. In contrast, *N*-methylation (*N*-Me-2-Nal^3) was shown to significantly reduce the antagonistic activity,[16] while introduction of a conformationally constrained Trp-derivative resulted in an inactive compound.[14]

However, analogues with aromatic side chains that are significantly different from 2-Nal have not previously been reported for this position. For the SAR study in position 3, we therefore replaced 2-Nal with aromatic and aliphatic residues of different size and shape, giving a compound series that contained small (**4**–**6**), medium (**7**–**10**), and large (**11**, **12**) side chains (Fig. 2). For reference purposes the known Ala^3 analogue **13**[14] was also included. Our synthetic strategy for preparation of cyclopentapeptide ligands and the biological assay for determination of antagonistic potency have recently been described.[15]

The analogues **4**–**13** were all significantly less potent than **2** (EC_{50} = 0.40 μM) (Fig. 2). The most potent compound was the isomeric 1-Nal^3 analogue **10** (EC_{50} = 5.6 μM, 14-fold reduction in potency), and the only compounds with activity (EC_{50} < 100 μM) contained a medium-sized substituent (**2** > **10** > **9**, **8**, **7**). Thus, the present SAR data show that position 3 of the cyclopentapeptides is very sensitive toward substitutions, and that 2-Nal remains the best residue in this position.

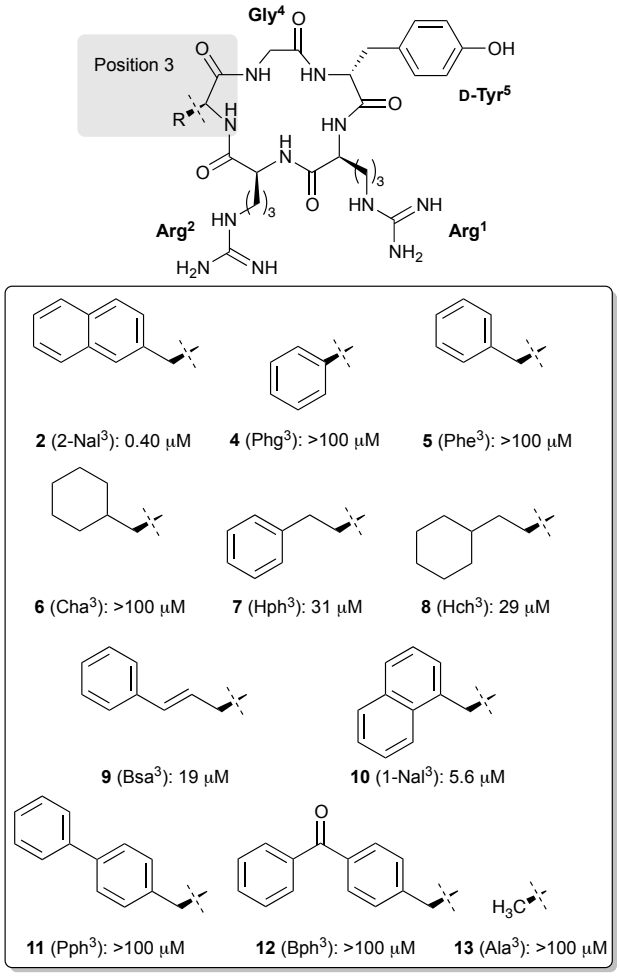

Figure 2 Structures and antagonistic potencies (EC_{50}) of the lead cyclopentapeptide **2** and the synthesized Xaa^3 analogues **4**–**13**.

Interestingly, Fujii et al. have previously reported a SAR study in the corresponding position 3 of the parent 14-mer peptide antagonist **1** (Fig. 1),[19] which showed that the Phe^3 analogue was only 6-fold less active than the 2-Nal^3 analogue (**1**), and 7-fold more active than the 1-Nal^3 analogue. This is in contrast to the present findings, where the Phe^3 analogue **5** was inactive, meaning that the main contribution comes from the distal aromatic ring of 2-Nal^3. Thus, the SAR trends obtained for the larger peptide scaffold **1** cannot automatically be transferred to the cyclopentapeptide scaffold of **2**. Moreover, the CXCR4 co-crystal peptide ligand **3** (Fig. 1) contains 1-Nal^3 instead of 2-Nal^3, while still being very potent (IC_{50} = 0.6 nM).[11] Thus, the present SAR data indicates that the 2-Nal^3 side chain of **2** approaches its subpocket in a different way than the corresponding side chain of the larger peptides **1** and **3**.

Design and SAR for position 5

For position 5 (D-Tyr), a larger number of cyclopentapeptide analogues have been reported in the literature.[14, 16, 20, 21] The L-Tyr^5 epimer was shown to be 35-fold less active than **2**,[10] meaning that the stereochemistry in this position is optimized. Regarding size, the smaller D-phenylglycine (D-Phg^5) analogue was shown to have low affinity,[21] while the larger D-naphthylalanine analogues had reduced affinity (D-1-Nal^5 >> D-2-Nal^5),[20] which shows that D-Tyr is the optimal size in position 5. Removal of the 4-hydroxyl group in D-Tyr^5 to give the D-Phe^5 analogue resulted in a 6-fold reduction in affinity,[14] while introduction of a halogen on the phenyl ring led to further affinity reduction.[21] Similarly, replacement of the 4-hydroxyl group with a 4-amino or a 4-methoxy group resulted in 13-fold and 64-fold reduction in affinity, respectively.[14] Overall, the reported phenyl-substituted analogues can be ranked by affinity as follows: 4-OH (**2**) > H > 2-F > 3-F > 4-NH_2 > 4-F > 4-OMe > 4-Cl > 4-Br.

Rational design of optimized ligands not only requires identification of the pharmacophoric groups, but also knowledge of the spatial orientation of these. For D-Tyr the global orientation of the side chain is described by the χ^1 torsion angle (Fig. 3), which can adopt three low-energy conformations: trans, gauche-(+), and gauche-(-). Our own pharmacophore modelling[17] and docking studies[15] of **2** and analogues suggest that χ^1 of D-Tyr^5 adopts the trans conformation ($\chi^1 \approx 180^\circ$) in the receptor-bound conformation, while docking studies by Demmer et al. suggest the gauche-(-) conformation ($\chi^1 \approx -60^\circ$).[22] However, SAR studies that address the rotameric state of the D-Tyr^5 side chain in the receptor-bound conformation of **2** have not been reported. We therefore designed a series of cyclopentapeptide analogues that contained conformational constraints in position 5. Since it was known that D-Tyr^5 could be replaced with D-Phe^5 without a drastic change in activity,[14] we used a series of constrained D-Phe mimetics (**15**–**18**, Fig. 3) instead of the corresponding hydroxylated D-Tyr mimetics, which were commercially unavailable. Thus, the known D-Phe analogue **14**[14] was used as the reference compound for this compound series; in our assay, **14** (EC_{50} = 0.85 μM) was approximately 2-fold less potent than **2** (EC_{50} = 0.40 μM), which reflects the contribution from the 4-hydroxyl group of D-Tyr^5.

In compound **15**, the phenyl ring was linked to C^α via a methylene group using the achiral 2-aminoindan-2-carboxylic acid (Aic^5) as building block. When considered as a D-Phe mimetic, χ^1 of Aic has approximately the same preference for the gauche-(+) and trans conformations, while the gauche-(-) conformation is unavailable.[23] Introduction of this amino acid (**15**) resulted in a 87-fold reduction of potency (EC_{50} = 74 μM) compared to **14**. In compound **16**, the phenyl ring was instead linked to N^α using 1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid (D-Tic^5). D-Tic is a chimera of D-Phe and D-pipecolic acid (D-Pic) where χ^1 can adopt the two gauche conformations, but not the trans conformation.[23] Compound **16** (EC_{50} = 25 μM) was 29-fold less potent than **14**; this is consistent with data for the corresponding hydroxylated D-Tyr mimetic D-Tic(7-OH)^5, which resulted in a 20-fold reduction in activity compared to **2**.[14] For the two final compounds in this series (**17** and **18**), we wanted to introduce 3-phenyl-D-Pro (D-Ppr), which exists as two diastereomers (3*S*- or 3*R*-configuration). D-Ppr is a chimera of D-Phe and D-Pro where χ^1 can access the trans conformation and one of the two gauche conformations (depending on which diastereomer that is used), but favours trans over gauche.[23] Since the enantiomerically pure (3*S*)-D-Ppr and (3*R*)-D-Ppr were commercially unavailable, racemic *trans*-Ppr (containing (3*S*)-D-Ppr and (3*R*)-L-Ppr) and *cis*-Ppr (containing (3*R*)-D-Ppr and (3*S*)-L-Ppr) were used as building blocks. The resulting compounds **17** and **18** (mixtures of two diastereomers, see ESI) were approximately equipotent, but still more than 40-fold less potent than **14**. Due to the relatively low activity of the diastereomeric mixtures, separation of the individual diastereomers was not attempted.

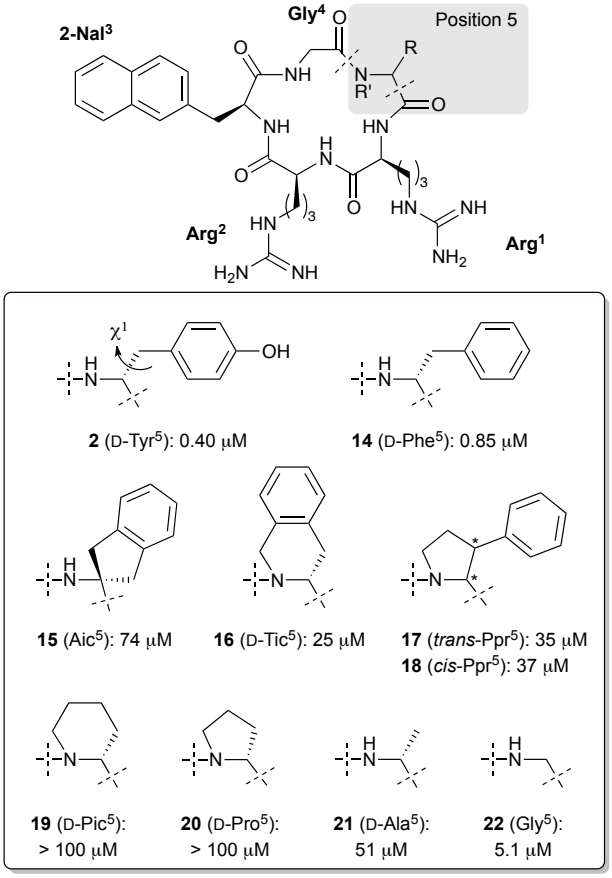

Figure 3 Structures and antagonistic potencies (EC_{50}) of the lead cyclopentapeptide **2** and the synthesized Xaa^5 analogues **14**–**22**, also showing the χ^1 torsion angle of D-Tyr^5 for **2**.

It was expected that one of the constrained analogues **15**–**18** would be significantly more potent than the others, reflecting the correct orientation of χ^1. However, all analogues showed moderate to low activity, which makes it difficult to draw any conclusions about the rotameric state for D-Tyr^5 in the receptor-bound conformation of **2**. Clearly, the beneficial effect of constraining χ^1 is outweighed by the simultaneous introduction of other structural elements that lead to unfavorable receptor interactions. First, introduction of 1-2 additional methylene groups leads to an increased steric demand, which the binding subpocket may not be able to accommodate. Second, Aic and D-Tic (**15** and **16**) not only constrain the χ^1 torsion angle, but also the χ^2 angle, which determines the plane of the phenyl ring.

Third, N^{α}-alkylation (**16-18**) removes the H-bond donor properties of the amide bond. Fourth, all the constrained D-Phe mimetics (**15**–**18**) will affect the backbone conformation: α,α-disubstituted amino acids (**15**) are known to stabilize/induce a helical backbone conformation,[24] while cyclic amino acids based on D-Pic and D-Pro (**16-18**) restrict the φ backbone torsion and promote trans/cis isomerization of the preceding amide bond.[25] In order to isolate the backbone effects imposed by D-Tic (**16**) and D-Ppr (**17** and **18**), we prepared the D-Pic[5] and D-Pro[5] analogues **19** and **20** as well as the known D-Ala[5] analogue **21**[14] (Fig. 3). Compounds **19** and **20** failed to produce any antagonistic activity (EC_{50} > 100 μM), while **21** was moderately active (EC_{50} = 51 μM), which confirms that the backbone effects are partly responsible for the relatively low potency of **16**–**18**. These findings are consistent with data for the *N*-methylated analogue (*N*-Me-D-Tyr[5]), which resulted in a 32-fold reduction in activity compared to **2**.[16] Realizing that introduction of conformational constraints in position 5 was counterproductive, we went in the opposite direction and prepared the Gly[5] analogue **22** (Fig. 3). Interestingly, this simplification resulted in an EC_{50} value of 5.1 μM, i.e., a 10-fold increase in potency compared to **21** and only a 6-fold reduction relative to **14**. Thus, the reduced size and/or the increased conformational flexibility of Gly[5] partly compensates for the side chain removal. Consequently, we do not longer consider the D-Tyr[5] side chain as an essential pharmacophoric element for cyclopentapeptide CXCR4 antagonists.

Rationalization of SAR for the two aromatic positions in terms of receptor binding

The sequence similarity between the 16-mer **3** and the cyclopentapeptide **2** (Fig. 1) suggests that the binding mode for **2** can be inferred from the co-crystal structure of **3** and CXCR4 (PDB code 3OE0).[11] However, inspection of this structure shows that the distance between Arg[1] and Tyr[5] in the receptor-bound conformation of **3** is not compatible with head-to-tail cyclization, which is required for **2**. This means that the position of some side chains in **2** must be shifted relative to the corresponding side chains in **3**. The crystal structure further shows that the aromatic side chains in positions 3 (1-Nal) and 5 (Tyr) of **3** are both located in hydrophobic regions around transmembrane helix (TMH) 5; 1-Nal[3] at the bottom of the pocket, and Tyr[5] higher up, close to the extracellular loop between TMHs 4 and 5.

In order to rationalize the present SAR data for the aromatic positions 3 and 5 in cyclopentapeptide ligands, compounds **2**, **14**, and **22** were docked to this X-ray structure using an optimized docking protocol as recently described.[15] Specifically, based on preliminary site-directed mutagenesis studies that identified Asp171 in TMH4 as important for the activity of **2**,[26] a H-bond constraint was placed on Asp171 to reduce the number of irrelevant poses. Moreover, Arg188 in extracellular loop 2 was temporarily mutated to alanine since the side chain partly restricted access to the sub-pocket containing Asp171.

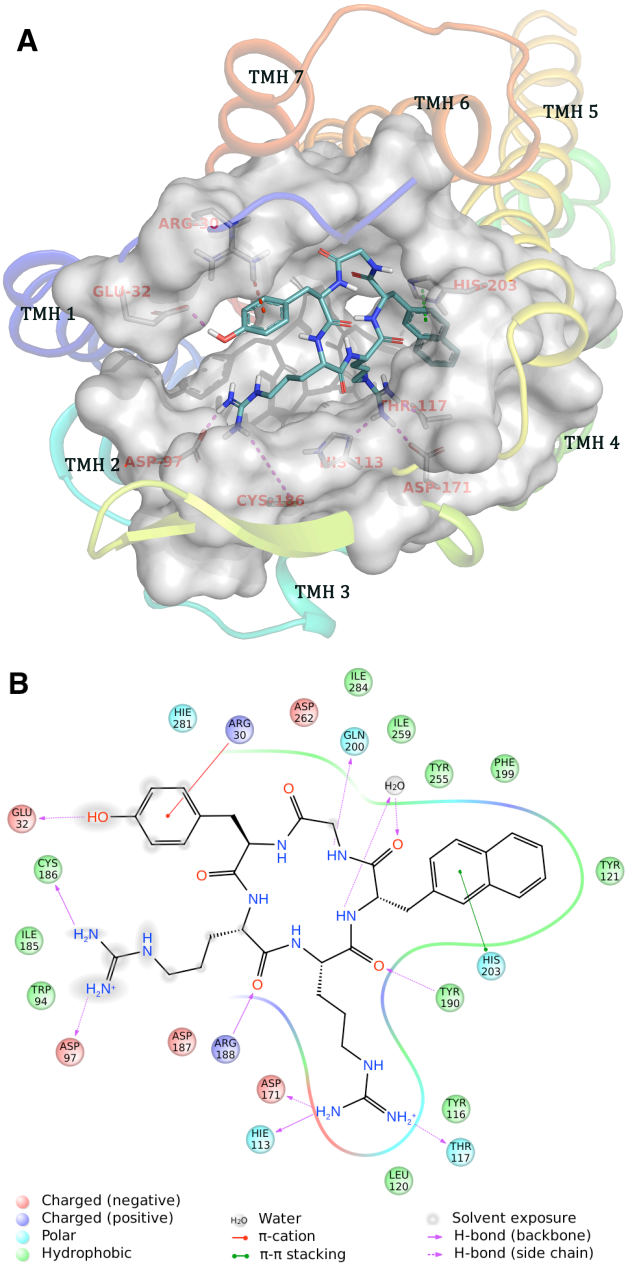

Figure 4 Proposed binding mode for the lead cyclopentapeptide **2**: (A) 3D representation showing the seven transmembrane helices (TMHs 1-7; coloured ribbons) and the binding pocket (grey surface; receptor atoms within 5 Å of the ligand) and selected key interactions; (B) 2D representation detailing the intermolecular interactions.

The docked ligands only differ in position 5 (**2**: D-Tyr[5]; **14**: D-Phe[5]; **22**: Gly[5]), and only minor differences were seen among their top scoring poses (data not shown); thus, they can be collectively represented by the binding mode for **2** (Fig. 4). The 2-Nal[3] side chain is accommodated in a well-defined hydrophobic subpocket mainly composed of residues in TMH 5 (Fig. 4A). The restrictions of this subpocket would explain the reduced potency of the Xaa[3] analogues **4**–**13**. In contrast, the side chain in position 5 (**2**: D-Tyr[5]; **14**: D-Phe[5]) is located at the opposite side of the transmembrane bundle, near the top of TMH 1, where it interacts with residues in the extracellular N-terminal fragment of CXCR4 (Fig. 4A). A favourable cation-π interaction is seen between the

guanidino group of Arg30 and the phenyl ring of D-Tyr^5 (**2**) or D-Phe^5 (**14**); the 4-hydroxyl group of **2** forms an additional H-bond with Glu32 (Fig. 4B), which would explain its higher activity. Still, the phenyl ring of **2** and **14** is located in a relatively open region and is partially solvent exposed (Fig. 4B), which represents an unfavourable contribution to binding. Overall, the favourable contributions outweigh the unfavourable, but the lack of a defined subpocket for Xaa^5 would explain why the D-Tyr^5 side chain can be removed without a dramatic loss of potency (**22**: Gly^5).

We have recently described the interactions between the two Arg residues and the receptor in detail;[15] briefly, the essential Arg^2 side chain sits deeply between TMHs 3 and 4, while the less important Arg^1 side chain is located higher up and points into a partly open region around TMH 2. When considering the entire cyclopentapeptide ligand, a general picture emerges where the Arg^2-2-Nal^3 fragment is buried in the major binding pocket (composed of TMHs 3, 4, 5, and 6), while the D-Tyr^5-Arg^1 fragment is located higher up in the minor binding pocket (composed of TMHs 1, 2, 3, and 7) and is partially solvent exposed. However, it should be noted that the Asp171-constraint used to generate these poses is based on preliminary site-directed mutagenesis data for **2**, and that alternative binding modes have been suggested.[27] Thus, further experimental studies are needed to fully establish the binding mode for cyclopentapeptide CXCR4 antagonists.

Experimental

General

All reagents and solvents were purchased and used as received. The individual amino acids were all N^{α}-Fmoc protected, and pentamethyl-2,3-dihydrobenzofuran-5-sulfonyl (Pbf) and *t*-Bu were used as protecting groups for the Arg and D-Tyr side chains, respectively. Preparative HPLC was performed with an XBridge™ C18 reversed phase column (250 mm × 19 mm, 10 µm particle size) on a Waters 600 Semi Prep System. Analytical HPLC was performed with an XBridge™ C18 reversed phase column (250 mm × 4.6 mm, 5 µm particle size) on a Waters 2695 system. Different gradients of CH_3CN-H_2O, containing 0.1% TFA, were used as eluting solvent for both preparative and analytical HPLC (flow rates of 15 mL/min and 1 mL/min, respectively), with photodiode array detection at 214 or 254 nm. HRMS spectra were obtained on an LTQ Orbitrap XL. ^{1}H and ^{13}C NMR spectra were recorded on a 400 MHz Varian spectrometer. Chemical shifts are expressed in ppm relative to methanol (^{1}H 3.31 ppm, ^{13}C 49.0 ppm). Coupling constants are given in hertz (Hz) and the values are given in δ scale.

Chemistry

The target compounds **4**–**22** were synthesized as recently described for **2**.[15] Briefly, the linear pentapeptide precursor was synthesized by standard Fmoc-based solid-phase peptide synthesis using a preloaded Fmoc-Gly Novasyn TGT resin. After cleavage from the resin, the linear side chain protected pentapeptide was cyclized head-to-tail in dilute solution to give the side chain protected cyclopentapeptide. Following deprotection, the crude peptide was purified by preparative HPLC and lyophilized to give the final products **4**–**22** as di(trifluoroacetate) salts. The identity of all final products was confirmed by HRMS and NMR, and all compounds were >95% pure as determined by analytical HPLC, see ESI.

Biology

The antagonistic potency of **4**–**22** was determined as recently reported for **2**; see ref. 12 for a detailed description. Briefly, the compounds were tested in a functional assay that measured inhibition of CXCL12-induced activation of human CXCR4, which was transiently expressed in COS-7 cells. Compounds **4**–**12** were tested in the range 10^{-10}–10^{-5} M, and the EC_{50} values were calculated by extrapolation of the curve under the assumption that the Hill coefficient was –1 and that the compounds were full antagonists. The other compounds were tested in the range 10^{-8}–10^{-4} M under the same assumptions.

Computational studies

The cyclopentapeptide ligands **2**, **14**, and **22** were docked to the CXCR4 structure using Schrödinger's induced fit docking workflow[28] as recently described.[15] Briefly, the X-ray co-crystal structure of human CXCR4 and the 16-mer peptide antagonist CVX15 (PDB code 3OE0)[11] was prepared for docking using the Protein Preparation Wizard workflow.[29] The three ligands were docked to this structure using our optimized protocol,[15] and the top 10 poses within an energy window of 30 kcal/mol were kept for each ligand. Visual inspection of the 30 generated poses resulted in the identification of the common binding mode for **2**, **14**, and **22** that is discussed above.

Conclusions

The SAR data presented here clearly show that the naphthyl group in position 3 is more important for activity than the phenol group in position 5 (**13** vs **21**), and that the distal aromatic ring of the 2-Nal^3 side chain is critical in order to maintain potency. The collective SAR data for the cyclopentapeptide CXCR4 antagonists, supported by molecular modeling, indicate that the Arg^2 and 2-Nal^3 side chains are buried in the receptor, while the side chains of D-Tyr^5 and Arg^1 are partly solvent exposed. In terms of peptidomimetic design, Arg^2-2-Nal^3 seems to serve as a minimal recognition motif, meaning that appropriately designed dipeptidomimetics have potential as CXCR4 antagonists. Tripeptidomimetic ligands based on the Arg^1-Arg^2-2-Nal^3 fragment represent an intermediary step along this path, and we are currently pursuing such compounds.

Acknowledgements

We thank Johann Eksteen, Jon Å. Aune, and Erik Thomassen for assisting in the synthesis of peptide ligands, and Associate Professor Dr. Bengt Erik Haug for valuable comments to the manuscript. We also thank Inger S. Simonsen and Randi Thøgersen for excellent technical assistance in the biological assay. Financial support for this project was obtained from the Research Council of Norway (grant 190728/V30) (J.M. and J.V.), from the University of Tromsø (Z.G.Z. and J.V.), and from the University of Copenhagen, the Danish Council for Independent Research | Medical Sciences, and the Aase and Einar Danielsen Foundation (S.T. and M.M.R.).

Notes and references

[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.*

† Electronic Supplementary Information (ESI) available: Antagonistic potencies for **4**–**22** in tabular format, and yields and characterization data (HRMS, ^{1}H and ^{13}C NMR, and purity) for **4**–**22**. See DOI: 10.1039/b000000x/

‡ These authors contributed equally to this work.

1. W. T. Choi, S. Duggineni, Y. Xu, Z. Huang and J. An, *J. Med. Chem.*, 2012, **55**, 977-994.
2. I. P. Singh and S. K. Chauthe, *Expert Opin. Ther. Pat.*, 2011, **21**, 227-269.
3. S. Hatse, K. Princen, L.-O. Gerlach, G. Bridger, G. Henson, E. De Clercq, T. W. Schwartz and D. Schols, *Mol. Pharmacol.*, 2001, **60**, 164-173.
4. M. M. Rosenkilde, L.-O. Gerlach, J. S. Jakobsen, R. T. Skerlj, G. J. Bridger and T. W. Schwartz, *J. Biol. Chem.*, 2004, **279**, 3033-3041.
5. M. M. Rosenkilde, L.-O. Gerlach, S. Hatse, R. T. Skerlj, D. Schols, G. J. Bridger and T. W. Schwartz, *J. Biol. Chem.*, 2007, **282**, 27354-27365.
6. R. M. Mosi, V. Anastassova, J. Cox, M. C. Darkes, S. R. Idzan, J. Labrecque, G. Lau, K. L. Nelson, K. Patel, Z. Santucci, R. S. Wong, R. T. Skerlj, G. J. Bridger, D. Huskens, D. Schols and S. P. Fricker, *Biochem. Pharmacol.*, 2012, **83**, 472-479.
7. A. Steen and M. M. Rosenkilde, in *Novel Developments in Stem Cell Mobilization: Focus on CXCR4*, eds. S. Fruehauf, W. J. Zeller and G. Calandra, Springer, New York, 2012, pp. 23-37.
8. M. Masuda, H. Nakashima, T. Ueda, H. Naba, R. Ikoma, A. Otaka, Y. Terakawa, H. Tamamura, T. Ibuka, T. Murakami, Y. Koyanagi, M. Waki, A. Matsumoto, N. Yamamoto, S. Funakoshi and N. Fujii, *Biochem. Biophys. Res. Commun.*, 1992, **189**, 845-850.
9. H. Tamamura, Y. Xu, T. Hattori, X. Zhang, R. Arakaki, K. Kanbara, A. Omagari, A. Otaka, T. Ibuka, N. Yamamoto, H. Nakashima and N. Fujii, *Biochem. Biophys. Res. Commun.*, 1998, **253**, 877-882.
10. N. Fujii, S. Oishi, K. Hiramatsu, T. Araki, S. Ueda, H. Tamamura, A. Otaka, S. Kusano, S. Terakubo, H. Nakashima, J. A. Broach, J. O. Trent, Z. X. Wang and S. C. Peiper, *Angew. Chem., Int. Ed.*, 2003, **42**, 3251-3253.
11. B. Wu, E. Y. T. Chien, C. D. Mol, G. Fenalti, W. Liu, V. Katritch, R. Abagyan, A. Brooun, P. Wells, F. C. Bi, D. J. Hamel, P. Kuhn, T. M. Handel, V. Cherezov and R. C. Stevens, *Science*, 2010, **330**, 1066-1071.
12. G. R. Marshall, *Biopolymers (Peptide Sci.)*, 2001, **60**, 246-277.
13. H. Tamamura, T. Araki, S. Ueda, Z. Wang, S. Oishi, A. Esaka, J. O. Trent, H. Nakashima, N. Yamamoto, S. C. Peiper, A. Otaka and N. Fujii, *J. Med. Chem.*, 2005, **48**, 3280-3289.
14. H. Tamamura, A. Esaka, T. Ogawa, T. Araki, S. Ueda, Z. Wang, J. O. Trent, H. Tsutsumi, H. Masuno, H. Nakashima, N. Yamamoto, S. C. Peiper, A. Otaka and N. Fujii, *Org. Biomol. Chem.*, 2005, **3**, 4392-4394.
15. J. Mungalpara, S. Thiele, Ø. Eriksen, J. Eksteen, M. M. Rosenkilde and J. Våbenø, *J. Med. Chem.*, 2012, **55**, 10287-10291.
16. S. Ueda, S. Oishi, Z. X. Wang, T. Araki, H. Tamamura, J. Cluzeau, H. Ohno, S. Kusano, H. Nakashima, J. O. Trent, S. C. Peiper and N. Fujii, *J. Med. Chem.*, 2007, **50**, 192-198.
17. J. Våbenø, G. V. Nikiforovich and G. R. Marshall, *Biopolymers (Peptide Sci.)*, 2006, **84**, 459-471.
18. S. Ueda, M. Kato, S. Inuki, H. Ohno, B. Evans, Z.-x. Wang, S. C. Peiper, K. Izumi, E. Kodama, M. Matsuoka, H. Nagasawa, S. Oishi and N. Fujii, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 4124-4129.
19. H. Tamamura, A. Omagari, K. Hiramatsu, S. Oishi, H. Habashita, T. Kanamoto, K. Gotoh, N. Yamamoto, H. Nakashima, A. Otaka and N. Fujii, *Bioorg. Med. Chem.*, 2002, **10**, 1417-1426.
20. T. Tanaka, H. Tsutsumi, W. Nomura, Y. Tanabe, N. Ohashi, A. Esaka, C. Ochiai, J. Sato, K. Itotani, T. Murakami, K. Ohba, N. Yamamoto, N. Fujii and H. Tamamura, *Org. Biomol. Chem.*, 2008, **6**, 4374-4377.
21. T. Tanaka, W. Nomura, T. Narumi, A. Esaka, S. Oishi, N. Ohashi, K. Itotani, B. J. Evans, Z. X. Wang, S. C. Peiper, N. Fujii and H. Tamamura, *Org. Biomol. Chem.*, 2009, **7**, 3805-3809.
22. O. Demmer, I. Dijkgraaf, U. Schumacher, L. Marinelli, S. Cosconati, E. Gourni, H. J. Wester and H. Kessler, *J. Med. Chem.*, 2011, **54**, 7648-7662.
23. S. E. Gibson, N. Guillo and M. J. Tozer, *Tetrahedron*, 1999, **55**, 585-615.
24. C. Toniolo, G. M. Bonora, A. Bavoso, E. Benedetti, B. di Blasio, V. Pavone and C. Pedone, *Biopolymers*, 1983, **22**, 205-215.
25. M. S. Weiss, A. Jabs and R. Hilgenfeld, *Nat. Struct. Biol.*, 1998, **5**, 676.
26. J. Mungalpara, S. Thiele, M. M. Rosenkilde and J. Våbenø, The 22nd American Peptide Symposium, San Diego, CA, June 25-30, 2011; poster YI-P2241.
27. O. Demmer, A. O. Frank, F. Hagn, M. Schottelius, L. Marinelli, S. Cosconati, R. Brack-Werner, S. Kremb, H.-J. Wester and H. Kessler, *Angew. Chem. Int. Ed. Engl.*, 2012, **51**, 8110-8113.
28. *Schrödinger Suite 2012 Induced Fit Docking protocol; Glide version 5.8, Schrödinger, LLC, New York, NY, 2012; Prime version 3.1, Schrödinger, LLC, New York, NY, 2012.*
29. *Schrödinger Suite 2011 Protein Preparation Wizard; Epik version 2.2, Schrödinger, LLC, New York, NY, 2011; Impact version 5.7, Schrödinger, LLC, New York, NY, 2011; Prime version 2.3, Schrödinger, LLC, New York, NY, 2011.*