(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 December 2011 (29.12.2011)

PCT

(10) International Publication Number
WO 2011/161266 A1

(51) International Patent Classification:
C07K 16/28 (2006.01) *A61P 37/00* (2006.01)
C07K 16/00 (2006.01)

(21) International Application Number:
PCT/EP2011/060738

(22) International Filing Date:
27 June 2011 (27.06.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/358,495 25 June 2010 (25.06.2010) US

(71) Applicant *(for all designated States except US)*: **ABLYNX NV** [BE/BE]; Technologiepark 21, B-9052 Ghent-Zwijnaarde (BE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **SCHOTTE, Peter** [BE/BE]; Hugo Verriestlaan 34, B-9840 De Pinte (BE). **BRIGE, Ann** [BE/BE]; Biezenkouter 15, B-9940 Ertvelde (BE). **DEWILDE, Maarten** [BE/BE]; Bergemeersenstraat 109, B-9300 Aalst (BE). **STALS, Hilde** [BE/BE]; Raas van Gaverestraat 114, B-9000 Gent (BE). **STRUBBE, Beatrijs** [BE/BE]; Marie-Thérèse de Courtebournelaan 8, B-9041 Oostakker (BE).

(74) Agents: **STEFFERL, Andreas** et al.; HOFFMANN · EITLE, Arabellastrasse 4, 81925 Munich (DE).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

WO 2011/161266 A1

(54) Title: IMPROVED IMMUNOGLOBULIN SINGLE VARIABLE DOMAINS AND CONSTRUCTS THEREOF DIRECTED AGAINST CXCR4

(57) Abstract: The present invention relates to amino acid sequences that are directed against CXCR4; as well as protein, constructs and compounds comprising the same; and also nucleic acids encoding the same.

IMPROVED IMMUNOGLOBULIN SINGLE VARIABLE DOMAINS AND CONSTRUCTS THEREOF DIRECTED AGAINST CXCR4

The present invention relates to amino acid sequences that are directed against CXCR4; as well as protein, constructs and compounds comprising the same; and also nucleic acids encoding the same.

Further aspects, embodiments, features, advantages, uses, applications and advantages from the present invention will become clear from the further description herein.

The international application WO 09/138519 by Ablynx N.V. entitled "*Amino acid sequences directed against CXCR4 and other GPCRs and compounds comprising the same*" describes amino acid sequences against G-protein coupled receptors (GPCRs) and in particular human CXCR4, Genbank accession number AF005058.

Unless explicitly mentioned otherwise herein, all terms mentioned herein have the meaning given in WO 09/138519 (or in the prior art cited in WO 09/138519). Also, where a method or technique is not specifically described herein, it can be performed as described in WO 09/138519 (or in the prior art cited in WO 09/138519). For example, the term "Nanobody" is as defined in WO 09/138519, and thus in a specific aspect generally denotes a VHH, a humanized VHH or a camelized VH (such as a camelized human VH) or generally a sequence optimized VHH (such as e.g. optimized for chemical stability and/or solubility, maximum overlap with known human framework regions and maximum expression).

WO 09/138519 describe a number of amino acid sequences and in particular VHHs and constructs thereof that are directed against human CXCR4 (see for example the amino acid sequences mentioned such as SEQ ID NO: 238 and SEQ ID NO: 239 in Table B-1.1 of WO 09/138519). WO 09/138519 also describes multivalent, multispecific and/or biparatopic constructs (as defined in WO 09/138519) that are directed against human CXCR4. Reference is for example made to the constructs referred to in Example 4 of WO 09/138519 such as SEQ ID NO: 264 in Table B-5 of WO 09/138519).

One particularly preferred example of an amino acid sequence against human CXCR4 from WO 09/138519 is the sequence called 238D2 (see SEQ ID NO: 238 in WO 09/138519):

EVQLVESGGGLVQTGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLEWVSGIKSSGDSTRYAGSVKGRFTISR DNAKNMLYLQMYSLKPEDTAVYYCAKSRVSRTGLYTYDNRGQGTQVTVSS (SEQ ID NO: 1)

One other particularly preferred example of an amino acid sequence against the human CXCR4 from WO 09/138519 is the sequence called 238D4 (see SEQ ID NO: 239 in WO 09/138519):

EVQLMESGGGLVQAGGSLRLSCAASGRTFNNYAMGWFRRAPGKEREFVAAITRSGVRSGVSAIYGDSVKDR FTISRDNAKNTLYLQMNSLKPEDTAVYTCAASAIGSGALRRFEYDYSGQGTQVTVSS (SEQ ID NO: 2)

WO 09/138519 further gives some non-limiting examples of multivalent, multispecific and/or biparatopic constructs that comprise 238D2 or 238D4 (see for example SEQ ID NO's: 261 to 266 in WO 09/138519 and in particular 238D2-20GS-238D4).

One other particularly preferred example of an amino acid sequence against the human CXCR4 from WO 09/138519 is the sequence called 238D2-20GS-238D4 (see SEQ ID NO: 264 in WO 09/138519):

EVQLVESGGGLVQTGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLEWVSGIKSSGDSTRYAGSVKGRFTISR DNAKNMLYLQMYSLKPEDTAVYYCAKSRVSRTGLYTYDNRGQGTQVTVSSGGGGSGGGGSGGGGSGGGG SEVQLMESGGGLVQAGGSLRLSCAASGRTFNNYAMGWFRRAPGKEREFVAAITRSGVRSGVSAIYGDSVKD RFTISRDNAKNTLYLQMNSLKPEDTAVYTCAASAIGSGALRRFEYDYSGQGTQVTVSS (SEQ ID NO: 3)

Generally, the anti-human CXCR4 amino acid sequences and constructs from WO 09/138519 show excellent biological activity and other desired properties. However, this does not mean that an anti-human CXCR4 amino acid sequence that would have (even further) improved properties would not be a valuable addition to the art.

The invention provides such improved anti-human CXCR4 amino acid sequences, and in particular (even further) improved variants of the sequences 238D2 (SEQ ID NO:1), 238D4 (SEQ ID NO:2) and 238D2-20GS-238D4 (SEQ ID NO:3).

In one aspect, the amino acid sequence provided by the invention is a variant of 238D2-20GS-238D4 that comprises, at position 5 of the 238D4 building block (numbering according to Figure 5 of this application), a valine instead of the original methionine residue. In this aspect, an amino acid sequence of the invention may be a variant of 238D2-20GS-238D4, also referred to herein as 4CXCR100 (SEQ ID NO: 4) that comprises, at position 5 of the 238D4 building block (numbering according to Figure 5 of this application), a valine residue:

EVQLVESGGGLVQTGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLEWVSGIKSSGDSTRYAGSVKGRFTISR DNAKNMLYLQMYSLKPEDTAVYYCAKSRVSRTGLYTYDNRGQGTQVTVSSGGGGSGGGGSGGGGSGGGG SEVQLVESGGGLVQAGGSLRLSCAASGRTFNNYAMGWFRRAPGKEREFVAAITRSGVRSGVSAIYGDSVKD RFTISRDNAKNTLYLQMNSLKPEDTAVYTCAASAIGSGALRRFEYDYSGQGTQVTVSS (SEQ ID NO: 4 or 4CXCR100)

In another aspect, the amino acid sequence provided by the invention is a variant of 238D2-20GS-238D4 that comprises i) at position 5 of the 238D4 building block (numbering according to Figure 1 of this application), a valine instead of the original methionine residue; and ii) at position 77 of the 238D2 building block (numbering according to Figure 1 of this application), a threonine instead of the original methionine residue. In this aspect, an amino acid sequence of the invention may be a variant of 238D2-20GS-238D4, also referred to herein as 4CXCR101 (SEQ ID NO: 5) that comprises i) at position 5 of the 238D4 building block (numbering according to Figure 5 of this application), a valine residue; and ii) at position 77 of the 238D2 building block (numbering according to Figure 5 of this application), a threonine:

EVQLVESGGGLVQTGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLEWVSGIKSSGDSTRYAGSVKGRFTISR DNAKNTLYLQMYSLKPEDTAVYYCAKSRVSRTGLYTYDNRGQGTQVTVSSGGGGSGGGGSGGGGSGGGGS EVQLVESGGGLVQAGGSLRLSCAASGRTFNNYAMGWFRRAPGKEREFVAAITRSGVRSGVSAIYGDSVKDR FTISRDNAKNTLYLQMNSLKPEDTAVYTCAASAIGSGALRRFEYDYSGQGTQVTVSS (SEQ ID NO: 5 or 4CXCR101)

Generally, an "optimized variant" of an amino acid sequence according to the invention is a variant that comprises one or more beneficial substitutions such as a substitutions increasing i) the degree of "humanization", ii) the chemical stability, and/or iii) the level of expression; while the potency (measured e.g. by the potency assay as described in the experimental part of WO 09/138519 or in this application) remains comparable (i.e. within a 10% deviation) to the wild type 238D2-20GS-238D4 (SEQ ID NO: 3) or comparable to the variant 4CXCR100 (SEQ ID NO: 4). Preferably, compared to the wild-type sequence of 238D2-20GS-238D4, an amino acid sequence of the invention contains at least one such substitution, and preferably at least two such substitutions, and preferably at least three humanizing substitutions and preferably at least 10 such humanizing substitutions. Also, again compared to the wild-type sequence 238D2-20GS-238D4, the amino acid sequences of the invention preferably comprise a maximum of 20 substitutions, and preferably a total of 15, 13, 11 or 10 substitutions (although the maximum number may in some cases not be critical, depending on the substitutions chosen). Some preferred, but non-limiting examples of such substitutions will become clear from the further description herein, and for example include, without limitation, for the 238D2 building block: T14P, M77T, Y82aN, K83R, and/or Q108L; and for the 238D4 building block: M5V, A14P, R39Q, K83R, T91Y, and/or Q108L (numbering according to Figure 5 of this application).

Also, as further described herein, the amino acid sequences of the invention may contain one or more other/further substitutions. Again, some preferred, but non-limiting examples of such

other/further substitutions will become clear from the further description herein, and for example may include (and preferably essentially consist of) one or more of the following substitutions (also referred to herein as *"substitutions (a) to (c)"*):

(a) one or more conservative amino acid substitutions; and/or

(b) one or more substitutions in which a "camelid" amino acid residue at a certain position is replaced by a different "camelid" amino acid residue that occurs at said position (for which reference is for example made to Figures 1-4). Some non-limiting examples of such substitutions are V5L, M43K (substitution to the residue that is most prevalent in this position in both human VH's as well as VHH's), S49A and/or A74S; and/or

(c) one or more substitutions that improve the (other) properties of the protein, such as substitutions that improve the long-term stability and/or properties under storage of the protein. These may for example and without limitation be substitutions that prevent or reduce oxidation events (for example, of methionine residues); that prevent or reduce pyroglutamate formation; and/or that prevent or reduce isomerisation or deamidation of aspartic acids or asparagines (for example, of DG, DS, NG or NS motifs). For such substitutions, reference is for example made to the International application WO 09/095235, which is generally directed to methods for stabilizing single immunoglobulin variable domains by means of such substitutions, and also gives some specific example of suitable substitutions (see for example pages 4 and 5 and pages 10 to 15). One example of such substitution may be to replace an NS motif at positions 82a and 82b with an NN motif;

(d) or any suitable combination of two or more of any of the foregoing substitutions (a) to (c).

It will be clear from the disclosure herein that the amino acid sequences of the invention contain at least one "amino acid difference" compared to each of the sequences of 238D2-20GS-238D4, respectively (in which the term *"amino acid difference"* is used herein in the same meaning as defined in WO 09/138519, namely as an insertion, deletion or substitution of a single amino acid residue on a position of the first sequence, compared to the second sequence; it being understood that two amino acid sequences can contain one, two, 5, 10, 11, 12 or more such amino acid differences. In the context of the present invention, any amino acid difference is preferably a substitution).

In particular, compared to the sequence of 238D2-20GS-238D4, the amino acid sequences of the invention contain at least one substitution (as defined herein), and may optionally contain one or

more further substitutions (such as any one of, or any suitable combination of any two or more of, the further substitutions (a) to (c) as mentioned herein).

In a preferred aspect of the invention, the amino acid sequences of the invention contain compared to the sequence 238D2-20GS-238D4 at least the substitutions: for the 238D2 building block: T14P, M77T, Y82aN, K83R, and Q108L and for the 238D4 building block: M5V (numbering according to Figure 5 of this application); also referred to herein as 4CXCR103 (SEQ ID NO: 6):

EVQLVESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLEWVSGIKSSGDSTRYAGSVKGRFTISR DNAKNTLYLQMNSLRPEDTAVYYCAKSRVSRTGLYTYDNRGQGTLVTVSSGGGGSGGGGSGGGGSGGGGS EVQLVESGGGLVQAGGSLRLSCAASGRTFNNYAMGWFRRAPGKEREFVAAITRSGVRSGVSAIYGDSVKDR FTISRDNAKNTLYLQMNSLKPEDTAVYTCAASAIGSGALRRFEYDYSGQGTQVTVSS (SEQ ID NO: 6) or 4CXCR103);

and may optionally contain one or more further humanizing substitutions (as described herein) and/or may optionally contain one or more further substitutions (such as any one of, or any suitable combination of any two or more of, the further substitutions (a) to (c) as mentioned herein); or a suitable combination of such humanizing substitutions and such other substitutions.

In a preferred aspect, the amino acid sequences of the invention contain a total of between 6 and 15, preferably between 9 and 13, such as 10, 11 or 12 amino acid substitutions compared to the wild-type sequence 238D2-20GS-238D4. As mentioned, these differences preferably at least comprise one and preferably both of the substitutions M5V in the 238D4 building block and/or M77T in the 238D2 building block, and at least one, preferably at least two, such as three, four or five or ten humanizing substitutions, and may optionally comprise one or more further substitutions (such as any one of, or any suitable combination of any two or more of, the further substitutions (a) to (c) as mentioned herein). Again, based on the disclosure herein and optionally after a limited degree of trial and error, the skilled person will be able to select (a suitable combination of) one or more such suitable humanizing and/or further substitutions.

In another specific aspect, the amino acid sequences of the invention contain a total of between 1 and 15, such as one, 5 or 10 amino acid differences compared to the sequence of 4CXCR100, in which at least one of these amino acid differences is the substitution M77T in the 238D2 building block and the other substitutions may for example be, and preferably are, either one or more further beneficial substitutions (better expression and better chemical stability) and/or one or more further substitutions (such as any one of, or any suitable combination of any two or more of, the further substitutions (a) to (c) as mentioned herein). Again, based on the disclosure herein and

optionally after a limited degree of trial and error, the skilled person will be able to select (a suitable combination of) one or more such suitable humanizing and/or further substitutions.

Also, most preferably, these amino acid differences compared to 238D2-20GS-238D4 and/or 4CXCR100 are most preferably located in the framework regions (defined according to Kabat, reference is again made to WO 09/138519), although it is not fully excluded that a very limited number of these amino acid differences (such as for example only one or two) may be present in the CDR's (as long as these do not detract (too much) from the desired affinity, on-rate or off-rate (for example, such amino acid differences in the CDR's may be introduced as a result of affinity maturation).

A preferred, but non-limiting aspect of an amino acid sequence of the invention is the amino acid sequences of the invention that contain compared to the sequence 238D2-20GS-238D4 at least the substitutions: for the 238D2 building block: T14P, M77T, Y82aN, K83R, and Q108L and for the 238D4 building block: M5V, A14P, R39Q, K83R, T91Y, and Q108L (numbering according to Figure 5 of this application); also referred to herein as 4CXCR104 (SEQ ID NO: 7):

EVQLVESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLEWVSGIKSSGDSTRYAGSVKGRFTISR DNAKNTLYLQMNSLRPEDTAVYYCAKSRVSRTGLYTYDNRGQGTLVTVSSGGGGSGGGGSGGGGSGGGGS EVQLVESGGGLVQPGGSLRLSCAASGRTFNNYAMGWFRQAPGKEREFVAAITRSGVRSGVSAIYGDSVKDR FTISRDNAKNTLYLQMNSLRPEDTAVYYCAASAIGSGALRRFEYDYSGQGTLVTVSS (SEQ ID NO: 7) or 4CXCR104);

and may optionally contain one or more further humanizing substitutions (as described herein) and/or may optionally contain one or more further substitutions (such as any one of, or any suitable combination of any two or more of, the further substitutions (a) to (c) as mentioned herein); or a suitable combination of such humanizing substitutions and such other substitutions.

Other amino acid sequences of the invention may for example contain a total of between one and five, such as one, two or three amino acid differences compared to the sequence of 4CXCR104 (while retaining all substitutions as defined above for 4CXCR104), in which such amino acid differences may for example be, and most preferably are, either one or more further humanizing substitutions and/or any one of, or any suitable combination of any two or more of, the further substitutions (a) to (c) as mentioned herein.

In a preferred aspect of the invention, it has been found that optimized variants of 238D2-20GS-238D4 such as 4CXCR103 and/or 4CXCR104 are more humanized, more stable, and give higher

expression or production yields and/or may have other advantages compared to other humanized variants of 238D2-20GS-238D4 that comprise e.g. a valine at position 5 for the 238D4 building block and a threonine at position 77 of the 238D2 building block (= 4CXCR101). Without being limited to any specific explanation or hypothesis, it is believed that this may be due to the fact that the preferred variants of 238D2-20GS-238D4 that comprise the substitutions as defined above for 4CXCR103 and/or 4CXCR104 allow the amino acid sequences of the invention to better able to fold into the desired immunoglobulin domain structure and/or, upon folding, to take on a more stable immunoglobulin domain structure.

Thus, in a specific, but non-limiting aspect, the invention relates to an amino acid sequence (i.e. an amino acid sequence of the invention) that is a variant of 238D2-20GS-238D4 (SEQ ID NO: 3) that comprises, compared to the amino acid sequence of 238D2-20GS-238D4, (i) the M5V mutation in the 238D4 building block; and (ii) the M77T mutation in the 238D2 building block; and optionally (iii) at least one to twenty, preferably at least one to eleven, and more preferably four or eleven substitutions (as defined herein i.e. one or more humanizing substitutions as well as one or more further suitable amino acid substitutions (preferably, any one of, or any suitable combination of any two or more of, the further substitutions (a) to (c) as mentioned herein).

As mentioned, such a variant of 238D2-20GS-238D4 preferably contains (i) a total of between 7 and 15, preferably between 9 and 13, such as 10, 11 or 12 amino acid differences compared to the wild-type sequence 238D2-20GS-238D4.

In another aspect, the invention relates to an amino acid sequence (i.e. an amino acid sequence of the invention) that is a variant of 4CXCR101 (SEQ ID NO: 5) that comprises, compared to the amino acid sequence of 4CXCR101, at least the substitutions: for the 238D2 building block: T14P, M77T, Y82aN, K83R, and Q108L and for the 238D4 building block: M5V (also referred herein as 4CXCR103 or SEQ ID NO: 6); and optionally (iii) as well one or more further suitable amino acid substitutions (preferably, any one of, or any suitable combination of any two or more of, the further substitutions (a) to (c) as mentioned herein). Such a variant of 4CXCR103 preferably contains (i) a total of between 1 and 10, preferably 5 amino acid substitutions compared to the sequence of 4CXCR103 (although the maximum number may in some cases not be critical, depending e.g. on the humanizing substitutions chosen).

In another aspect, the invention relates to an amino acid sequence (i.e. an amino acid sequence of the invention) that is a variant of 4CXCR103 that comprises, compared to the amino acid sequence of 4CXCR103, (i) for the 238D4 building block: A14P, R39Q, K83R, T91Y, and Q108L (also referred to

herein as 4CXCR104 or SEQ ID NO: 7); and optionally (ii) as well one or more further suitable amino acid substitutions, in which said amino acid differences are preferably substitutions and more preferably substitutions that are chosen from one or more further humanizing substitutions (compared to the humanizing substitutions already present in 4CXCR104) and/or from any one of, or any suitable combination of any two or more of, the further substitutions (a) to (c) as mentioned herein.

The invention also relates to proteins and polypeptides that comprise or essentially consist of an amino acid sequence of the invention.

An alignment of 238D2-20GS-238D4 (WO 09/138519) with 4CXCR238D2 and 4CXCR238D4 is given in Figure 5, which form the basis for 4CXCR100 (invention), 4CXCR101 (invention), 4CXCR103 (invention), and 4CXCR104 (invention).

With respect to any humanizing substitutions that may be present in the amino acid sequences of the invention (i.e. compared to 238D2-20GS-238D4 (WO 09/138519), 4CXCR100 (invention), 4CXCR101 (invention), 4CXCR103 (invention), and 4CXCR104 (invention)), it is remarked that as described in WO 09/138519, a humanizing substitution can generally be defined as a substitution whereby an amino acid residue that occurs in a framework regions of a camelid V_{HH} domain is replaced by a different amino acid that occurs at the same position in the framework region of a human V_H domain (and preferably, a human V_H3 domain). Thus, suitable humanizing substitutions will be clear to the skilled person based on the disclosure herein, the disclosure in WO 09/138519, and from a comparison of the amino acid sequence of a given V_{HH} sequence and one or more human V_H sequences.

Reference is for example made to the attached Figures 1-4 (which have been taken from Tables A-6 to A-9 of WO 09/138519), which list some of the amino acid residues that have been found to occur in the framework regions of camelid VHH domains, and the corresponding amino acid residue(s) that most often occur in the framework regions of a human V_H3 sequence (such as for example, the germline sequences DP-47, DP-51 or DP-29). The humanizing substitutions that can be taken from these Figures are also some of the preferred humanizing substitutions used in the invention; however, it may also be possible to use humanizing substitutions that have been obtained by comparison with other germline sequences (from the V_H3 class or sometimes also from other V_H classes). As generally known from WO 09/138519 (and from the patent applications from Ablynx N.V. and the further prior art mentioned in WO 09/138519), based on such sequence comparison, particularly suited and/or optimal humanizing substitutions (and combinations thereof) may

generally be determined by limited trial and error, i.e. by introducing one or more envisaged humanizing substitutions and testing the humanized variants thus obtained for one or more desired properties, such as melting temperature, affinity, potency, properties upon formatting, expression levels in a desired host organism, and/or other desired properties for VHH domains or Nanobodies or proteins/polypeptides comprising the same, for which again reference is made to WO 09/138519 and the further patent applications by Ablynx N.V. mentioned therein).

With regard to humanizing substitutions, it should be noted that for the purposes of the present application, any substitutions at any of the camelid "Hallmark residues" (see again WO09/068627, as well as Figures 1-4) should not be counted as a "humanizing substitution". Such substitutions at any of the Hallmark residues may or may not be present, and which when present may or may not be a substitution in which an amino acid residue in a VHH is replaced by an amino acid residue that occurs at the same position of a human V_H sequence. For example, such a substitution at a Hallmark residue may for example also be a substitution in which an amino acid residue that occurs at the Hallmark position is replaced by another amino acid residue that occurs at said position in camelid V_{HH} sequences (reference is again made to Figures 1-4).

It will be clear to the skilled person from the disclosure herein that the amino acid sequences of the invention are directed against human CXCR4 and are improved variants for 238D2-20GS-238D4 as described in WO 09/138519. Thus, the amino acid sequences of the invention can be used for the same purposes, uses and applications as described in WO 09/138519, for example to inhibit signaling that is mediated by human CXCR4 and/or its ligand(s); and/or in the prevention or treatment of diseases associated with an increased signalling of CXCR4, such as the various diseases in the group of cancer such as hematopoietic cancers like CLL, AML, ALL, MM, Non-Hodgkin lymphoma, solid tumors such as breast cancer, lung cancer, brain tumors, ovarian cancer, stromal chemoresistance of tumors, leukemia and other cancers, disrupting adhesive stromal interactions that confer tumor cell survival and drug resistance, mobilizing tumour cells from tissue sites and making them better accessible to conventional therapy, inhibiting of migration and dissemination of tumor cells (metastasis), inhibiting or paracrine growth and survival signals, inhibiting pro-angiogenesis effects of SDF-1, inflammation and inflammatory disorders such as bowel diseases (colitis, Crohn'disease, IBD), infectious diseases, psioriasis, autoimmune diseases (such as MS), sarcoidosis, transplant rejection, cystic fibrosis, asthma, chronic obstructive pulmonary disease, rheumatoid arthritis, viral infection, HIV, West Nile Virus encephalitis , common variable immunodeficiency. Furthermore, the amino acid sequences of the invention can be used for stem cell mobilization in various patients in need of stem cells after X-ray radiation such as e.g. cancer patients after radiation treatment to

replenish the stem cell pool after radiation in cancer patients, or in patients in need of more stem cells, e.g. in patients with ischemic diseases such as myocardial infarction (MI), stroke and/or diabetes (i.e. patients in need of tissue repair) wherein more stem cell would be re-transfused (after mobilization, screening, selection for lineage in need (e.g. cardiac, vascular lineages) and ex-vivo expansion of patient's own stem cells).

In particular, the amino acid sequences of the invention are very potent (i.e. EC50 values as measured e.g. in the experimental part in the pM range) antagonists of human CXCR4 and/or are inverse agonists in certain continuously active human CXCR4 mutants (see e.g. Example 5 of WO 09/138519). Reference is for example made to Example 5 and 6 on pages 222ff of WO 09/138519, as well as the further general disclosure of WO 09/138519. More in particular, the amino acid sequences of the invention may be used as an improved alternative to 238D2-20GS-238D4, and thus may in particular be used for the same purposes as described in WO 09/138519 for 238D2-20GS-238D4.

As already mentioned in WO 09/138519, one of these applications of 238D2-20GS-238D4 (and thus for the amino acid sequences of the invention) is as a building block in compounds or constructs that comprise, next to one or more amino acid sequences of the invention, and one or more other groups, residues, moieties, binding domains or binding units (as described in WO 09/138519). For example, as described in WO 09/138519, such one or more further binding domains or binding units may be other immunoglobulin single variable domains, VHH's, (single) domain antibodies, Nanobodies or dAb's, and these may for example be directed against a protein or binding unit that provides for increased half-life (for example, albumin or a binding unit or binding peptide that can bind to a serum protein such as albumin).

Such multispecific constructs are preferably proteins and polypeptides (i.e. encoded by a nucleotide sequence and/or capable of being expressed by a host or host cell), as also generally described in WO 09/138519.

As described herein "polypeptides of the invention" may be essentially as described for the "polypeptides of the invention" described in WO 09/138519 that comprise 238D2-20GS-238D4 and one or more substitutions (e.g. as disclosed herein) and are in particular and preferred 4CXCR103 and 4CXCR104, and most preferred are 4CXCR104.

Preferably, all amino acid sequences and Nanobodies of the invention and polypeptides of the invention are capable of undergoing essentially the same binding interactions as described in WO 09/138519 for 238D2-20GS-238D4, and polypeptides comprising 238D2-20GS-238D4.

Also, amino acid sequences of the invention preferably have a functional activity (as measured using the cAMP Hunter eXpress CXCR4 assay described in Example 2 in the experimental part) that is essentially the same than the functional activity of 238D2-20GS-238D4.

Moreover, amino acid sequences of the invention preferably have a inhibiting activity in the chemotaxis assay (see Example 4 in the experimental part) that is essentially the same than the activity of 238D2-20GS-238D4.

Polypeptides of the invention preferably have a functional activity as measured using the cAMP Hunter eXpress CXCR4 assay described in Example 3 in the experimental part that has an EC50 value that is lower than about 100nM such as between 100nM and 50nM or lower, more preferably that is 50nM or lower (see Example 4).

Polypeptides of the invention preferably have an inhibiting activity in the chemotaxis assay (see Example 4 in the experimental part) that has an IC50 value that is lower than 0.8 nM such as between 0.6 nM and 0.8 nM (see Example 4).

In another aspect, the invention relates to a nucleic acid that encodes an amino acid sequence of the invention, a Nanobody of the invention or a polypeptide of the invention (or a suitable fragment thereof). Such a nucleic acid will also be referred to herein as a "*nucleic acid of the invention*" and may for example be as essentially further described in WO 09/138519; and may in particular be in the form of a genetic construct, again as essentially further described in WO 09/138519 and may be the nucleic acid sequences as disclosed herein (SEQ ID NO: 11 to 14).

In another aspect, the invention relates to a host or host cell that expresses (or that under suitable circumstances is capable of expressing) an amino acid sequence of the invention, a Nanobody of the invention and/or a polypeptide of the invention; and/or that contains a nucleic acid of the invention. Such a host or host cell may again generally be as described in WO 09/138519 and may be *Pichia Pastoris* as disclosed in the experimental part of this application.

The invention also relates to methods for the production/expression of the amino acid sequences, Nanobodies and polypeptides of the invention. Such methods may generally comprise the steps of (i) the expression, in a suitable host cell or host organism or in another suitable expression system of a nucleic acid that encodes an amino acid sequence, a Nanobody or polypeptide of the invention, optionally followed by: (ii) isolating and/or purifying the amino acid sequence, Nanobody or polypeptide of the invention thus obtained. In particular, such a method may comprise the steps of (i) cultivating and/or maintaining a host of the invention under conditions that are such that said

host of the invention expresses and/or produces at least one amino acid sequence, Nanobody and/or polypeptide of the invention; optionally followed by (ii) isolating and/or purifying the amino acid sequence, Nanobody or polypeptide of the invention thus obtained. These methods again may essentially be performed as described in WO 09/138519.

In another aspect the invention relates to a method for producing an amino acid sequence, a Nanobody or protein or polypeptide according to the invention by the methods as described in WO2010/125187, which is incorporated herein by its entirety. In particular, the invention relates to a method for producing an amino acid sequence, a Nanobody or protein or polypeptide according to the invention, said method at least comprising the steps of culturing a host cell to produce said amino acid sequence, a Nanobody or protein or polypeptide according to the invention comprising:

a) cultivating said host cell in a culture medium under conditions that are such that said host cell will multiply;

b) maintaining said host cell under conditions that are such that said host cell expresses and/or produces said amino acid sequence, a Nanobody or protein or polypeptide according to the invention;

optionally followed by

c) secreting said amino acid sequence, Nanobody or protein or polypeptide according to the invention into the culture medium by said host cell; and

d) isolating and/or purifying the secreted amino acid sequence, Nanobody or protein or polypeptide according to the invention from the culture medium.

The invention further relates to a method as described herein, further applying conditions that promote the formation of disulfide bridges in and/or after step a), step b), step c), and/or step d). The conditions that promote the formation of disulfide brides are selected from one or more of the following:

a) addition of oxidizing agents, preferably oxidizing metal ions, preferably one or more selected from Cu2+, Fe2+, Fe3+ and Zn2+;

b) enhancing expression of a thiol isomerase;

c) adapting the culturing conditions by one or more selected from the following: lowering culturing temperature and/or optimizing the culturing medium, including but not limited to reduction of methanol feed for hosts requiring a methanol feed, lowering conductivity of the culture medium, addition of yeast extract and/or peptone, or any combination thereof;

d) refolding the amino acid sequence, Nanobody or protein or polypeptide according to the invention in the presence of redox-buffer, preferably in the additional presence of denaturant;

e) treating the amino acid sequence, Nanobody or protein or polypeptide according to the invention by oxygenation, increasing temperature, increasing pH, or high pressure, or any combination thereof; and

f) combinations of any of a) through e).

The oxidizing agent, preferably oxidizing metal ions, more preferably one or more selected from Cu2+, Fe2+, Fe3+ and Zn2+, are added to at least one production step of the an amino acid sequence, a Nanobody or protein or polypeptide according to the invention, preferably selected from: culturing the host to produce the amino acid sequence, Nanobody or protein or polypeptide according to the invention, the culture supernatant comprising the amino acid sequence, Nanobody or protein or polypeptide according to the invention after removal of the host, any step of purifying the amino acid sequence, Nanobody or protein or polypeptide according to the invention, or the purified amino acid sequence, Nanobody or protein or polypeptide according to the invention.

In particular, the present invention relates to a method as described herein, wherein said addition of oxidizing agents, preferably oxidizing metal ions, preferably Cu2+, Fe2+, Fe3+ and Zn2+, more preferably 1-10 mM Cu2+ is performed alone, or in combination with one or more of the conditions according to claim 2 b) to e) and/or in combination with one or more of the conditions as described herein.

In particular, the present invention relates to a method as described herein, wherein said thiol isomerase is selected from PDI, calsequestrin and other PDI-related proteins comprising, but not limited to ERp72, ERp57, ERp60, ERp44, ERp5, ERp27 and PDIR, preferably PDI, or coexpressing genes (such as HAC1P, Bip/Kar2p) that increase the basal expression level of foldases and chaperones.

In particular, the present invention relates to a method as described herein, wherein said culturing temperature is lowered by 5°C as compared to the standard culturing temperature for the host organism.

In particular, the present invention relates to a method as described herein, wherein said methanol feed is lowered by 30-80% as compared to the standard methanol feed for the respective host.

In particular, the present invention relates to a method as described herein, wherein said conductivity of the culture medium is lowered by 30% to 80% as compared to the standard medium for the respective host.

In particular, the present invention relates to a method as described herein, wherein yeast extract and/or peptone are added to the culture medium at a concentration in the feed of 0 to 20%.

In particular, the present invention relates to a method as described herein, wherein refolding the amino acid sequence, Nanobody or protein or polypeptide according to the invention in the presence of denaturant and redox-buffer is performed using 2M guanidinium hydrochloride and 1:5 mM/mM cystamine/cysteamine.

In particular, the present invention relates to a method as described herein, wherein the amino acid sequence, Nanobody or protein or polypeptide according to the invention is treated by increasing the temperature to 40-60°C, preferably 55°C, increasing pH to pH 8-9, and/or high pressure to 250-5000 bar, preferably about 1000-2000 bar, optionally combined with oxygenation by purging with oxygen.

In particular, the present invention relates to a method as described herein, wherein the amino acid sequence, Nanobody or protein or polypeptide according to the invention is attached to a stationary phase of a chromatographic column.

In particular, the present invention relates to a method as described herein, wherein said eukaryotic host is selected from insect cells, mammalian cells, and lower eukaryotic hosts comprising yeasts such as Pichia, Hansenula, Saccharomyces, Kluyveromyces, Candida, Torulopsis, Torulaspora, Schizosaccharomyces, Citeromyces, Pachysolen, Debaromyces, Metschunikowia, Rhodosporidium, Leucosporidium, Botryoascus, Sporidiobolus, Endomycopsis, preferably Pichia pastoris.

In another aspect, the invention relates to methods for the production/expression of the amino acid sequences, Nanobodies and polypeptides of the invention wherein the production/expression is improved (i.e. 2 to 3.5 higher expression of the selected optimized amino acid sequences, Nanobodies and polypeptides of the invention) compared to the reference compound 4CXCR100 while having at the same time essentially the same activity than 4CXCR100, i.e. has an IC50 value in the chemotaxis assay of about 0.8 nM or lower (see Example 4) and/or an EC50 value in the cAMP assay of about 55pM or lower (see Example 3). Such methods may generally comprise the steps of (i) the expression, in a suitable host cell or host organism or in another suitable expression system of a nucleic acid that encodes an amino acid sequence, Nanobody or polypeptide of the invention (and in

particular 4CXCR104 or 4CXCR103, or most preferred 4CXCR104), optionally followed by: (ii) isolating and/or purifying the amino acid sequence, Nanobody or polypeptide of the invention thus obtained. In particular, such a method may comprise the steps of (i) cultivating and/or maintaining a host of the invention under conditions that are such that said host of the invention expresses and/or produces at least one amino acid sequence and/or polypeptide of the invention; optionally followed by (ii) isolating and/or purifying the amino acid sequence, Nanobody or polypeptide of the invention thus obtained.

One specific method for the production/expression of the amino acid sequences, Nanobodies and polypeptides of the invention is described in the International application of Ablynx N.V. entitled "Method for the production of domain antibodies", which has an international filing date of April 30, 2010 and application no.: PCT/EP2010/055916.

The invention further relates to a product or composition containing or comprising at least one amino acid sequence of the invention, at least one polypeptide of the invention (or a suitable fragment thereof) and/or at least one nucleic acid of the invention, and optionally one or more further components of such compositions known per se, i.e. depending on the intended use of the composition. Such a product or composition may for example be a pharmaceutical composition (as described herein), a veterinary composition or a product or composition for diagnostic use (as also described herein).

The invention also relates to the use of an amino acid sequence, Nanobody or polypeptide of the invention, or of a composition comprising the same, in (methods or compositions for) modulating and in particular inhibiting (as defined herein and/or in WO 09/138519) CXCR4 and/or CXCR4 -mediated signalling, either in vitro (e.g. in an in vitro or cellular assay) or in vivo (e.g. in an a single cell or in a multicellular organism, and in particular in a mammal, and more in particular in a human being, such as in a human being that is at risk of or suffers from a disease or disorder associated with CXCR4 and/or its ligands).

The invention also relates to methods for modulating and in particular inhibiting (as defined herein and/or in WO 09/138519) CXCR4 and/or CXCR4 -mediated signalling, either in vitro (e.g. in an in vitro or cellular assay) or in vivo (e.g. in an a single cell or multicellular organism, and in particular in a mammal, and more in particular in a human being, such as in a human being that is at risk of or suffers from a disease or disorder associated with CXCR4 and/or its ligands), which method comprises at least the step of contacting CXCR4 with at least one amino acid sequence, Nanobody or

polypeptide of the invention, or with a composition comprising the same, in a manner and in an amount suitable to modulate and in particular to block CXCR4 and/or CXCR4 -mediated signalling.

The invention also relates to the use of an one amino acid sequence, Nanobody or polypeptide of the invention in the preparation of a composition (such as, without limitation, a pharmaceutical composition or preparation as further described herein) for modulating and in particular inhibiting (as defined herein and/or in WO 09/138519) CXCR4 and/or CXCR4 -mediated signalling, either in vitro (e.g. in an in vitro or cellular assay) or in vivo (e.g. in an a single cell or in a multicellular organism, and in particular in a mammal, and more in particular in a human being, such as in a human being that is at risk of or suffers from a disease or disorder associated with CXCR4 and/or its ligands).

The invention will now be further described by reference to the following non-limiting Experimental Part and the non-limiting Figures, in which:

FIGURES

Figures 1-4: Tables giving sequence comparisons of the framework regions of human V_H3 domains and V_{HH} sequences (data taken from WO 09/138519).

Figure 5: Alignment of clones of interest with their humanized counterparts vs. VH3-23/JH5 with germline counterparts and residues substituted during optimization underlined

Figure 6: Expression level analysis of 3 clones of respectively 4CXCR101, 4CXCR103 and 4CXCR104 was compared to one clone of the wild type construct 4CXCR100.

Figure 7: The functional activity of the Nanobody construct variants 4CXCR101, 4CXCR103 and 4CXCR104 was compared to the parental Nanobody construct 4CXCR100 in a cAMP assay

Figure 8: The functional activity of the Nanobody construct variants 4CXCR101, 4CXCR103 and 4CXCR104 was compared to the parental Nanobody construct 4CXCR100 in a chemotaxis assay

Figure 9: RP-HPLC chromatograms of 4CXCR100 (panel A), 4CXCR103 (panel B), 4CXCR104 (panel C), and 4CXCR101 (panel D) before and after treatment with 20 mM H_2O_2 for 3 hours at room temperature. The insets show a zoom on the main peak and the prepeak, which represents the oxidised protein (wavelength = 280 nm).

Figure 10: The functional activity of the Nanobody construct variant 4CXCR104 was compared to the benchmark antibody in a pERK assay.

EXPERIMENTAL PART

Example 1: Sequence optimization of the two building blocks of 238D2-20GS-238D4

The protein sequence of parent 4CXCR238D2 (SEQ ID NO: 1) and 4CXCR238D4 (SEQ ID NO: 2) are aligned to the human VH3-23 (DP-47) and JH5 germlines (SEQ ID NO: 8) in Figure 5. Amino acid differences relative to the human germline sequence are represented by letters, identical amino acids by dots. Amino acid differences in framework regions that are underlined in Figure 5 were selected for conversion in the variants whereas the others were left untouched. The following 2 building blocks were obtained (Table B-1):

Table B-1:

4CXCR0003 (SEQ ID NO: 9): EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLEWVSGIKSSGDSTRYAGSVKGRFTISRDN AKNTLYLQMNSLRPEDTAVYYCAKSRVSRTGLYTYDNRGQGTLVTVSS
4CXCR0024 (SEQ ID NO: 10): EVQLLESGGGLVQPGGSLRLSCAASGRTFNNYAMGWFRQAPGKEREFVAAITRSGVRSGVSAIYGDSVKDRFTI SRDNAKNTLYLQMNSLRPEDTAVYYCAASAIGSGALRRFEYDYSGQGTLVTVSS

Purified, monovalent material was produced from nucleic acid encoding for 4CXCR238D2, 4CXCR0003, 4CXCR238D4 and 4CXCR0024. These materials (also referred to here as Nanobodies) were then characterized in a number of assays, like binding ELISA and ligand displacement and others (see Table B-2). In the binding ELISA assay, CXCR4 lipoparticles were immobilized per well on 96-well Maxisorp plates by overnight coating at 4°C. Following inhibiting with 4% Marvel in PBS, the Nanobodies were added and bound Nanobody (harbouring a c-Myc tag) were detected via sequential mouse anti-Myc and rabbit anti-mouse-HRP detection. In the heat treatment setup, samples were heated (Tm + 10 °C) prior to the ELISA. The displacement assay was carried out as described in WO 09/138519. In brief, 40 pM of [^{125}I] SDF-1 ligand (in-house labeled) was allowed to bind 2 µg of hCXCR4/HEK293 membrane extracts in presence or absence of the Nanobodies. After incubation for 1 hour at 4°C, membrane extracts were washed and the total amount of bound ligand radioactive counts per minute (cpm) are determined. Aspecific binding of the radio-labeled ligand to the membrane extracts (non-CXCR4 related) was determined by addition of excess unlabelled SDF-1 (100 nM) to compete all radio-ligand from the CXCR4 receptor. The aspecific binding value for each plate was subtracted from the total binding (cpm in absence of Nanobody) and the cpm values obtained for each Nanobody, and % residual [^{125}I] SDF-1 binding in presence of Nanobody was

calculated. In addition, the variants were analyzed in the thermal shift assay (TSA) and by differential scanning calorimetry (DSC) as well. Briefly, in the thermal shift assay Nanobodies (0.2 mg/ml) at different pH (range 3.5-9.0) were subjected to a heating cycle (37 → 90 °C, 0.05 °C/min) in presence of Sypro Orange. At higher temperatures samples started to unfold and Sypro Orange got unquenched by binding to the hydrophobic patches of the Nanobody which became surface exposed. The fluorescence signal of Sypro Orange was monitored during the heating cycle. In order to obtain the melting temperature, the first derivative of the signal was calculated. The maximum peak corresponded to the melting temperature of the Nanobody in a particular buffer. For the DSC assays, samples (0.5 mg/ml for Tm determination, 0.3 mg/ml for reversibility studies) were analyzed with the MicroCal Automated VP-capillary Differential Scanning Calorimeter. Samples were heated (1 °C/min) till 95°C for Tm determination. For reversibility measurements, samples were heated (1 °C/min) till 5°C above Tm and cooled down (1 °C/min) again. The percentage reversibility was calculated by looking at the ratio of Cp at the melting temperature and a baseline point. Finally, all variants were tested by analytical size exclusion chromatography (SEC) to check on unwanted multimerization (Phenomenex column (BioSep-Sec2000, flow 200 µl/min in PBS)).

Table B-2: Overview characterization data parenteral compound and sequence optimized variants.

	Binding ELISA			Ligand displace-ment					
Nanobody ID	EC50 untreated	EC50 treated	ratio (treated/ untreated)	Ki	Tm TSA (°C)	Tm DSC (°C)	reversibility (%) (DSC)	Analytical SEC	% human vs VH3-23/JH5
4CXCR238D2	399pM	457pM	1,1	174pM	61,4	61,3	90,8	Ok	89.7
4CXCR0003	479pM	416pM	0,9	195pM	61,5	61,5	89,5	Ok	96.6
4CXCR238D4	236pM	448pM	1,9	98pM	71,0	70,8	92,7	Ok	80.5
4CXCR0024	150pM	462pM	3,1	107pM	76,5	76,4	96,6	Ok	87.4

Overall, the variants showed comparable binding and ligand inhibiting characteristics compared to their parental counterparts. In addition, the Tm of 4CXCR0024 increased by around 10% upon sequence optimization. These sequence optimized building blocks were used then to make the biparatopic constructs 4CXCR101 (SEQ ID NO: 5), 4CXCR103 (SEQ ID NO: 6) and 4CXCR104 (SEQ ID NO: 7), the only difference being position five of each building block which was valine in the final construct for which there is no expectation that this mutation (L5V) has an influence on the behavior of the final construct.

Example 2: Production at small scale and medium scale of 4CXCR0101, 4CXCR103 and 4CXCR104

a) Cloning, copy number determination and expression analysis.

In this example we describe cloning of the parental 4CXCR100 (238D2-20GS-238D4 with M5V, i.e. also referred to as SEQ ID NO: 4) and the 3 sequence optimized anti-CXCR4 Nanobody constructs; respectively 4CXCR101 (SEQ ID NO: 5), 4CXCR103 (SEQ ID NO: 6) and 4CXCR104 (SEQ ID NO: 7), their copy number determination and expression levels after production *Pichia pastoris* X33 in shake flasks. The subunits in these bivalent Nanobodies are fused head-to-tail with a 20G/S linker.

The 4 different Nanobody constructs were produced after recombinant expression in *Pichia pastoris* expression system, based on *Pichia pastoris* optimized nucleic acid sequences and the commercially available system from Invitrogen/RTC using X-33 as a host strain. The genes coding for the different variants 4CXCR101, 4CXCR103 and 4CXCR104 were designed in correspondence to their amino acid sequence. The genes were synthesized at Geneart and codon use favorable for Pichia expression was introduced (see Table B-3).

Table B-3: Nucleic acid sequences encoding protein products 4CXCR100, 4CXCR101, 4CXCR103 and 4CXCR104

4CXCR100 – SEQ ID NO: 11 GAGGTGCAATTGGTGGAGTCTGGGGGAGGCCTGGTGCAAACTGGGGGGTCTCTGAGACTCTCCTGTGCAGC CTCTGGATTCACCTTTAGTAGCTATGCCATGAGTTGGGTCCGCCAGGCTCCAGGAAAGGGGCTCGAGTGGGTC TCAGGTATTAAAAGTAGTGGTGATAGTACAAGATATGCGGGCTCCGTGAAGGGCCGATTCACCATCTCCAGA GACAACGCCAAGAATATGCTGTATCTGCAAATGTACAGCCTGAAACCTGAGGACACGGCCGTGTATTACTGTG CAAAGTCACGGGTTAGCCGAACCGGATTATATACGTACGACAACAGGGGCCAGGGGACCCAGGTCACCGTCT CCTCAGGTGGCGGTGGCAGTGGAGGTGGCGGATCCGGCGGGGGAGGTAGTGGGGGTGGGGGCTCAGAGG TGCAATTGGTGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGGGCTCTCTGAGACTCTCCTGTGCAGCCTCTG GACGCACCTTCAATAACTATGCCATGGGCTGGTTCCGCCGGGCTCCAGGGAAGGAGCGTGAATTTGTTGCAG CTATTACCCGGAGTGGAGTCCGGAGTGGTGTTAGCGCAATCTATGGAGACTCCGTGAAGGACCGATTCACCA TCTCCAGAGACAACGCCAAGAACACGCTGTATCTGCAAATGAACAGCCTGAAACCTGAGGACACGGCCGTTTA CACCTGTGCGGCCTCGGCAATAGGATCTGGAGCGCTACGACGTTTTGAGTATGACTACTCGGGCCAGGGGAC CCAGGTCACCGTCTCCTCATGATAA
4CXCR101– SEQ ID NO: 12 GAGGTGCAATTGGTTGAATCTGGTGGTGGATTGGTTCAAACAGGTGGAAGCTTGAGACTTTCCTGTGCAGCTT CTGGTTTCACCTTTAGCTCCTACGCAATGAGTTGGGTCAGACAAGCTCCAGGTAAAGGTCTGGAATGGGTATC CGGTATTAAGTCTTCTGGTGATTCGACTAGATACGCTGGTAGTGTTAAGGGCAGATTCACTATTTCGCGAGAT AACGCTAAGAACACTTTGTACTTGCAAATGTACTCCCTTAAGCCTGAAGATACCGCTGTCTACTACTGCGCAAA GAGCAGAGTTTCCAGAACTGGACTATACACCTACGACAACAGAGGACAAGGCACACAGGTCACCGTGTCTTC AGGTGGCGGTGGCAGTGGAGGTGGCGGATCCGGCGGGGGAGGTAGTGGGGGTGGGGGCTCAGAAGTTCA ATTGGTCGAAAGTGGTGGAGGTTTGGTACAAGCAGGTGGATCTCTGAGACTGTCTTGTGCTGCTTCTGGTAG GACGTTTAACAACTACGCTATGGGTTGGTTTCGAAGAGCCCCAGGAAAAGAAAGGGAATTTGTTGCAGCTAT CACACGTTCAGGTGTTAGATCAGGTGTCTCTGCTATATACGGAGATAGCGTCAAGGATAGGTTCACGATCAGT

CGTGATAACGCCAAGAACACCCTTTACCTGCAAATGAACTCGTTGAAACCTGAGGATACTGCTGTTTACACTTG TGCAGCTAGCGCTATTGGTAGTGGCGCTTTGCGTAGATTCGAGTACGACTACTCTGGTCAAGGAACACAGGTC ACCGTCTCCTCA
4CXCR103 – SEQ ID NO: 13: GAGGTGCAATTGGTTGAATCTGGTGGAGGTTTGGTACAACCAGGTGGTAGCCTAAGACTTTCCTGTGCTGCTA GTGGATTCACCTTCTCTTCCTACGCAATGAGTTGGGTAAGACAAGCTCCAGGAAAGGGATTGGAATGGGTTA GCGGAATAAAGTCTTCGGGTGATAGCACACGTTACGCTGGTTCAGTCAAAGGTAGGTTCACGATTAGTCGAG ATAACGCAAAGAACACTCTATACCTGCAGATGAACTCCTTAAGACCTGAGGACACAGCTGTGTACTACTGTGC AAAGTCTCGTGTTAGCAGAACTGGTCTTTACACCTACGACAACAGAGGACAAGGAACGTTGGTCACCGTTTCC TCAGGTGGCGGTGGCAGTGGAGGTGGCGGATCCGGCGGGGGAGGTAGTGGGGGTGGGGGCTCAGAAGTTC AATTGGTCGAAAGTGGTGGAGGTTTGGTACAAGCAGGTGGATCTCTGAGACTGTCTTGTGCTGCTTCTGGTA GGACGTTTAACAACTACGCTATGGGTTGGTTTCGAAGAGCCCCAGGAAAAGAAAGGGAATTTGTTGCAGCTA TCACACGTTCAGGTGTTAGATCAGGTGTCTCTGCTATATACGGAGATAGCGTCAAGGATAGGTTCACGATCAG TCGTGATAACGCCAAGAACACCCTTTACCTGCAAATGAACTCGTTGAAACCTGAGGATACTGCTGTTTACACTT GTGCAGCTAGCGCTATTGGTAGTGGCGCTTTGCGTAGATTCGAGTACGACTACTCTGGTCAAGGAACACAGGT CACCGTCTCCTCA
4CXCR104 – SEQ ID NO: 14: GAGGTGCAATTGGTTGAATCTGGTGGAGGTTTGGTACAACCAGGTGGTAGCCTAAGACTTTCCTGTGCTGCTA GTGGATTCACCTTCTCTTCCTACGCAATGAGTTGGGTAAGACAAGCTCCAGGAAAGGGATTGGAATGGGTTA GCGGAATAAAGTCTTCGGGTGATAGCACACGTTACGCTGGTTCAGTCAAAGGTAGGTTCACGATTAGTCGAG ATAACGCAAAGAACACTCTATACCTGCAGATGAACTCCTTAAGACCTGAGGACACAGCTGTGTACTACTGTGC AAAGTCTCGTGTTAGCAGAACTGGTCTTTACACCTACGACAACAGAGGACAAGGAACGTTGGTCACCGTTTCC TCAGGTGGCGGTGGCAGTGGAGGTGGCGGATCCGGCGGGGGAGGTAGTGGGGGTGGGGGCTCAGAAGTTC AATTGGTCGAATCTGGTGGAGGTTTGGTTCAACCAGGTGGTTCCTTGAGACTGTCTTGTGCTGCTTCTGGTAG GACGTTTAACAACTACGCTATGGGTTGGTTTCGTCAAGCTCCAGGTAAAGAGCGAGAATTTGTTGCTGCAATT ACAAGATCTGGCGTTAGATCTGGAGTATCCGCAATATACGGTGATTCCGTTAAGGACAGATTCACAATCTCAC GTGATAACGCCAAGAACACTTTGTACCTGCAAATGAACTCCCTAAGACCAGAAGACACTGCAGTCTACTATTG TGCTGCTTCAGCTATTGGTTCTGGTGCTTTGAGACGATTTGAGTACGATTACTCTGGTCAAGGAACCTTGGTCA CCGTCTCCTCA

Transformation of the X-33 strain was done with the obtained expression vectors, and clones were selected on zeocin containing plates. Clones were picked *ad random* and were streaked on a new zeocin plate. A qPCR was performed to rank the clones according to their copy numbers. For each Nanobody construct, 4 different clones with a high copy number (>3) were selected based on a qPCR copy number screening assay. Next, the respective clones of each construct were tested for their expression level in shake flask.

Figure 6 shows the relative expression levels versus the parental construct after SDS-Page analysis of a medium sample. All clones from one construct showed equal expression levels. The expression level of the different constructs could be ranked according to the level of sequence optimization, which was the highest for 4CXCR104, followed by 4CXCR103, and the lowest expression level for 4CXCR101 and the parental Nanobody 4CXCR100. Hence, it appears that the M77T mutation in the 4CXCR238D2 (SEQ ID NO: 1) building block, introduced for sequence optimization, resulted in a

reduced expression level. Completely unexpectedly, this reduced expression was not only rescued, but even significantly improved by further mutations in the 4CXCR238D2 (SEQ ID NO: 1) building block (cf. expression level of 4CXCR103 with 4CXCR101 and 4CXCR102). In fact, the expression level was further increased by mutating the 4CXCR238D4 (SEQ ID NO: 2) building block in addition (cf. expression level of 4CXCR104 with 4CXCR103).

b) Production via fermentation at medium scale (2L).

The different constructs 4CXCR100, 4CXCR101, 4CXCR103 and 4CXCR104 were further evaluated for their expression level at 2L fermentor scale. Baseline high cell density *Pichia pastoris* (X33) fermentation conditions were used with the following parameter settings: temperature set constant to 30°C, pH 5 during the biomass production and the following induction phase, dissolved oxygen set constant to 30%, and antifoam A204 (Sigma) for foam control. Cell biomass was accumulated during the first batch and glycerol fed batch phase, followed by the MeOH induction phase, starting at low MeOH feed rate (4 mL/L.h with adaptation phase), during which the Nanobody was secreted into the fermentation medium. The estimated expression titres of the 3 different constructs in the Ably/Hyp-A medium are shown in B-4, confirm that also at fermentation scale production, 4CXCR104 had the highest expression levels, followed by 4CXCR103 and the parental 4CXCR100. The lowest expression was again observed for 4CXCR101, which was 35% lower compared to the parental construct 4CXCR100. Compared to the parental Nanobody, the production yields of the sequence optimized constructs 4CXCR103 and 4CXCR104 were 2 to 3.5 times higher. These increased production yields resulting from the amino acid changes is wholly unexpected, but have a clear economical advantage.

Table B-4: Overview of the estimated expression yields for the different anti-CXCR4 Nanobody constructs after SDS-Page analysis

Construct	Wet Cell Weight at end of fermentation (g/L)	Estimated yield (mg/L clarified fermentation medium)	Estimated yield (mg/L fermentation broth)
4CXCR100	412 g/L	1200 mg/L	706 mg/L
4CXCR101	374 g/L	800 mg/L	495 mg/L
4CXCR103	360 g/L	2500 mg/L	1593 mg/L
4CXCR104	348 g/L	5000 mg/L	2504 mg/L

Quality controls of the clarified fermentation broth via RP-HPLC analysis after a clean-up step demonstrated that 4CXCR104 contained a fraction of unpaired cysteine variant (10%), which could be re-oxidized by copper treatment for 4hrs at the end of induction phase, using 1mM $CuSO_4$.

c) Purification of 4CXCR100: 4CXCR101, 4CXCR103 and 4CXCR104

The parental 4CXCR100 Nanobody, expressed in *Pichia pastoris* was purified using a downstream process consisting of five DSP steps; centrifuged culture broth was clarified by a microfiltration-TFF step (Hydrosart, 0.2µm cassette), followed by an ultrafiltration-TFF step (Hydrosart 10kDa, Sartocon, Sartorius). The product was captured on an SP Sepharose FF equilibrated in 1/10 PBS, pH 6.5 and eluted with PBS pH7.3, 1M NaCl, followed by a polish step using Poros 50HS, equilibrated in 25mM citric acid pH4.0, and eluted using PBS, 1M NaCl pH 7.3 to remove product related impurities such as degradation products. Prior to the size exclusion chromatography (SEC) step in PBS, using Superdex75pg, to remove possible HMW variants, the Nanobody® was concentrated via NFF using Vivaspin (5kDa MWCO, 20mL Sartorius). DNA and endotoxins were removed via an anion exchange chromatography (AEX) step in flow-through mode, using Source 30Q in PBS. Finally, the product is sterile filtered through a 0.22µm filter prior to freezing at −70°C.

4CXCR101, 4CXCR103 and 4CXCR104 were purified via a 2-step purification procedure to provide rapidly a limited amount of product, with low HCP and endotoxin content. After harvest of the fermentor, part of the spent medium was clarified via a microfiltration-TFF step (Hydrosart, 0.2µm cassette) followed by a capture step using Toyopearl GigaCapS-650M equilibrated in 25mM NaPi, pH6.8 and eluted with 250mM NaCl. LPS-removal was done by OGP (N-octyl--D-glucopyranoside) treatment followed by SEC in PBS. Prior to this final SEC step; all samples were incubated ON at 4°C at pH 6.8 in the presence of 100µM $CuSO_4$, to re-oxidize the unpaired cysteine variant. Purity of the batches was analyzed via SDS-page, RPC, SEC and LC/MS (data not shown), and was comparable for all three.

Example 3: Functional activity of 4CXCR0101, 4CXCR103 and 4CXCR104 in cAMP assay

The functional activity of the Nanobody construct variants 4CXCR101, 4CXCR103 and 4CXCR104 was compared to the parental Nanobody construct 4CXCR100 (238D2-20GS-238D4 with M5V, SEQ ID NO: 4) in a cAMP assay (cAMP Hunter™ eXpress CXCR4 Assay, DiscoveRx). For this, cells were seeded in a 96-well plate at a density of 30,000 cells/well in OCC2 medium and incubated overnight in a 37°C,

5% CO_2, humidified incubator. The medium was aspirated the next day and 45 µL/well HBSS/10 mM Hepes/Antibody Reagent mix was added. This mix was composed of 1/3 Antibody Reagent and 2/3 HBSS/10 mM Hepes. A ½ serial dilution of Nanobody, together with 25 nM SDF-1 (R&D Systems) and 20 µM forskolin (provided with kit) was added to the plate and incubated for 30 minutes at 37°C. Afterwards, 60 µL cAMP Detection Reagent/cAMP Solution D mixture, composed of 1 part Substrate Reagent 2, 5 parts Substrate Reagent 1 and 19 parts cAMP lysis buffer, was transferred to the wells and incubated for 1 hour at room temperature protected from light. Finally, 60 µL cAMP solution A was added to each well and incubated for 3 hours at room temperature protected from light. The plates were read with Tecan Infinite F200 using the luminescence program. The results of this assay (ran in triplicate) are shown in Figure 7 and the IC50 values are summarized in Table B-5. In conclusion, no large shift in the functional activity of all the CXCR100 variants was observed (see also Table B-5 below).

Table B-5: functional activity of the Nanobody construct variants 4CXCR101, 4CXCR103 and 4CXCR104 was compared to the parental Nanobody construct 4CXCR100 in a cAMP assay:

[illegible]	29,1 nM	25,0 nM - 33,8 nM
[illegible]	35,6 nM	29,6 nM - 42,7 nM
[illegible]	36,1 nM	30,1 nM - 43,3 nM
[illegible]	42,0 nM	32,7 nM - 53,9 nM

Example 4: Functional activity of 4CXCR0101, 4CXCR103 and 4CXCR104 in chemotaxis assay

The functional activity of the three Nanobody variants 4CXCR101, 4CXCR103 and 4CXCR104 was compared to the parental Nanobody 4CXCR100 in a chemotaxis assay using the Jurkat cell line (Jurkat E6-1; ATCC). For this, cells were seeded 1 day before the experiment at a cell concentration of $0.5x10^6$ cell/ml in complete medium (RPMI1640 + 10% FBS). The following day, SDF-1 (200pM final concentration, R&D Systems) and serially diluted Nanobodies were added to the bottom of a small chemotaxis plate (Neuprobe 106-5) in a total volume of 29µl. A chemotaxis filter membrane (ChemoTx® Disposibla, pore size 5µm) was placed on top of the wells, ensuring that the membrane was in contact with the solution in the wells below. Nanobody dilution (10µl at 5X the serially diluted final concentration as below the membrane in each well) was added on top of the membrane, followed by 40µl of Jurkat cell suspension ($6.25x10^6$ cell/ml). The plates were incubated for 3 hours at 37°C in a humidified incubator (5% CO_2). After incubation, the filters were carefully removed and the cells in the well below were resuspended in the existing solution. The complete

cell suspension was transferred to the corresponding wells of white polystyrene Costar plates. After this, 30µl of Cell Titer Glo reagent (Promega G7571) was added to each well, followed by a 10 minute incubation, with shaking in the dark. Luminescence was measured (1 sec/well) using Envision 2103 Multilabel Reader with emission filter 700 (Perkin Elmer). The results of this assay (duplicate run) are shown in Figure 8 and the IC50 values are summarized in Table B-6. As also seen for the cAMP assay, no significant difference in potency was observed for the sequence optimized variants of 4CXCR100.

Table B-6: The functional activity of the three Nanobody variants 4CXCR101, 4CXCR103 and 4CXCR104 was compared to the parental Nanobody 4CXCR100 in a chemotaxis assay using the Jurkat cell line

Variant	IC50 (M)
4CXCR100	6.31×10^{-10}
4CXCR101	6.26×10^{-10}
4CXCR103	6.19×10^{-10}
4CXCR104	7.88×10^{-10}

Example 5: Chemical stability of 4CXCR0101, 4CXCR103 and 4CXCR104 in oxidation assay

Forced oxidation experiments in the presence of H_2O_2 have shown that the 238D2 building block of the biparatopic 4CXCR100 is susceptible to oxidation. The oxidation site was identified as M77 by peptide mapping and LC/MSMS (results not shown). In order to reduce the vulnerability to oxidation and thus to increase the chemical stability of the Nanobody, a M77T mutation was included during sequence optimisation of 4CXCR100 into 4CXCR101, 4CXCR103 and 4CXCR104.

To verify that 4CXCR101, 4CXCR103 and 4CXCR104 are less susceptible to oxidation compared to 4CXCR100, the molecules were subjected to a forced oxidation reaction. Briefly, the four Nanobodies were diluted to 1 mg/mL in D-PBS and incubated for 3 hours at room temperature in the presence of 20 mM H_2O_2. After incubation, the samples were desalted and analyzed on RP-HPLC. The resulting chromatograms are shown in Figure 9 and the integration data are summarized in B-7. These data clearly demonstrated that 4CXCR101, 4CXCR103 and 4CXCR104 are less susceptible to oxidation: forced oxidation of 4CXCR101, 4CXCR103 and 4CXCR104 resulted in only a minor effect on the relative pre peak area corresponding to the oxidised variant (increase to 4-9%), whereas the pre peak of the 4CXCR100 molecule dramatically increased to 55.6%. These data confirm that 4CXCR101, 4CXCR103 and 4CXCR104, containing the M77T mutation, are chemically more stable than 4CXCR100.

Table B-7: Integration results from RP-HPLC analysis of 4CXCR100, 4CXCR101, 4CXCR103 and 4CXCR104 before and after treatment with 20 mM H_2O_2.

Nanobody	Area % Prepeak (oxidised)	Area % Main peak
4CXCR100	5.8	91.8
4CXCR100 + 20 mM H_2O_2	55.6	40.7
4CXCR101	5.2	91.4
4CXCR101 + 20 mM H_2O_2	7.6	88.8
4CXCR103	5.0	91.5
4CXCR103 + 20 mM H_2O_2	9.2	86.2
4CXCR104	1.1	95.2
4CXCR104 + 20 mM H_2O_2	4.7	88.1

Methods

RP-HPLC experiments were carried out on an Agilent 1200 series instrument from Agilent Technologies (Palo Alto, USA). The RP-HPLC conditions were as follows:

Column : Zorbax 300SB-C3, 4.6 x 150 mm, 5 µm (Agilent, Part.No. 883995-909)

Solvent A : 0.1% TFA in H_2O

Solvent B : 0.1% TFA in 99.9% ACN

Column temperature : 70°C

Flow : 1 mL/min

Amount injected : 10 µg

Gradient:

Time (min)	Mobile phase B (%)
0	10
2.5	10
3	28
27	36
27.5	95
30	95
30.1	10
34	10

Example 6: Potency of 4CXCR104 in a Cell based potency assay: Cellular Erk

In order to examine the potency of the constructs of the invention, a cellular Erk phosphorylation assay was used comparing 4CXCR104 with a benchmark antibody (an anti-CXCR4 monoclonal antibody).

Upon ligand binding, CXCR4 is stabilized in a conformation that activates heterotrimeric G-protein, of which Gi is a major component. However, other G-proteins and non-G-proteins mediated pathways are also used. A wide variety of downstream effector pathways are activated such as adenylate cyclase and phospholipase C that influence intracellular concentrations of second messengers (cyclic AMP, diacylglycerol, inositol 1,4,5 trisphosphate and Ca2+) but also mediate Extracellular Signal Regulated kinase (ERK1/2) phosphorylation. Activation of Gαi/o, Gαs, Gαq/11 or Gα12/13 modulates ERK1/2 activation via numerous mechanisms. In addition, both α and βγ-subunits of G proteins can stimulate ERK1/2 phosphorylation through transactivation of receptor tyrosine kinases (RTKs). GPCRs have also been shown to mediate ERK1/2 activation in a G protein-independent but β-arrestin dependent manner.

The Cellul'erk - 50000 tests of Cisbio (catalog nr 64ERKPEI) was used, essentially according to the manufacturer's instructions. In short, 10^4 stable CXCR4 transfected CHO-K1 cells were plated in 96 well Tissue Culture Treated Plates (white bottom) (Corning // Cat#: 3917 // 18210022) according to plate layout, and incubated for 24h in a humid chamber (wet tissues) at 37°C in CO2 incubator. 24h after plating of the cells, the medium was replaced by assay medium, comprising SDF-1α or SDF-1α and the test compounds. The cells were incubated for 10 min at RT, after which a final lysis/Erk blocking reagent was added. The cell debris was further assayed. The results are presented in Table B-8, and graphically represented in Figure 10.

Table B-8 comparison of 4CXCR104 with benchmark Ab) in Erk potency assay

	4CXCR104	Benchmark Ab
Best-fit values		
Bottom	8767	7690
Top	45164	43250
LogIC50	-7,718	-6,599
HillSlope	-1,329	-1,329
IC50	1,92E-08	2,52E-07
Span	36397	35560

Std. Error		
Bottom	744	1183
Top	594,3	464,5
LogIC50	0,04128	0,05061
HillSlope	0,08474	0,08474
Span	952,8	1261
95% Confidence Intervals		
Bottom	7282 to 10251	5329 to 10050
Top	43978 to 46349	42323 to 44177
LogIC50	-7.800 to -7.635	-6.700 to -6.498
HillSlope	-1.498 to -1.160	-1.498 to -1.160
IC50	1.585e-008 to 2.316e-008	1.994e-007 to 3.175e-007
Span	34496 to 38298	33045 to 38076

The results demonstrate that the constructs of the invention, and in particular 4CXCR104, inhibit the SDF1α-induced phosphorylation of ERK more potent than the benchmark Ab.

Example 7: Potencies of CXCR4 constructs in comparison with AMD3100

In order to examine the potency of the constructs of the invention, 4CXCR104 was compared to AMD3100 (plerixafor, mozobil) in an ERK phosphorylation assay, a cAMP assay and in a migration assay.

AMD3100 is a bicyclam compound that binds CXCR4 (see *e.g.,* Antimicrob Agents Chemother 2000, 44: 1667-11673; AMD3100 is also known as plerixafor or mozobil).

The cAMP was performed essentially as described in Example 3, but with 111.5 nM SDF1α. The chemotaxis (or migration) assay was performed essentially as described in Example 4, but with 1 nM SDF1α. The Erk assay was performed essentially as described in Example 6, but with 10 nM SDF1α.

The results are depicted in Table B-9

B-9 Overview IC50 results of 4CXCR104 compared to plerixafor in ERK phosphorylation, cAMP and migration assays

Assay type	IC_{50} 4CXCR104	IC_{50} plerixafor
ERK phosphorylation assay (10 nM SDF1)	19.2 ± 7.35 nM (n = 81)	1741 nM ± 158 nM (n = 2)
cAMP assay (111.5 nM SDF1)	16.6 nM ± 5.87 nM (n = 11)	8740 nM ± 5176 nM (n = 2)

Migration assay (1 nM SDF1)	0.59 nM ± 0.13 nM (n = 4)	494 nM ± 238 nM (n = 6)

The results suggest that the Nanobodies® have significantly better potency in the cAMP assay, the migration assay and the ERK phosphorylation assay at physiologically relevant ligand concentrations, as compared with the SME benchmark molecule.

5

The terms and expressions which have been employed are used as terms of description and not of limitation, and there is no intention in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, it being recognized that various modifications are possible within the scope of the invention.

10 All references disclosed herein are incorporated by reference, in particular for the teaching that is referenced hereinabove.

What is claimed is:

CLAIMS

1. Amino acid sequence that is a variant of 238D2-20GS-238D4 (SEQ ID NO:3) that comprises, compared to the amino acid sequence of 238D2-20GS-238D4, (i) at least the mutation M77T in the 238D2 building block; and (ii) optionally at least one, preferably at least two, and more preferably three, four of five humanizing substitutions; and (iii) optionally one or more further suitable amino acid substitutions in any of the framework regions.

2. Amino acid sequence selected from the group of 4CXCR104 (SEQ ID NO: 7), 4CXCR103 (SEQ ID NO: 6) and 4CXCR101 (SEQ ID NO: 5).

3. Amino acid sequence with amino acid sequence as shown in SEQ ID NO: 7.

4. Amino acid sequence with amino acid sequence as shown in SEQ ID NO: 6.

5. Amino acid sequence with amino acid sequence as shown in SEQ ID NO: 5.

6. Protein or polypeptide that essentially consists of an amino acid sequence according to any of claims 1 to 5.

7. Protein or polypeptide that comprises an amino acid sequence according to any of claims 1 to 5.

8. Protein or polypeptide that comprises an amino acid sequence according to any of claims 1 to 5 and one or more other groups, residues, moieties, binding domains or binding units.

9. Protein or polypeptide that comprises an amino acid sequence according to any of claims 1 to 5 and one or more other immunoglobulin single variable domains, VHH's, (single) domain antibodies, Nanobodies or dAb's.

10. Protein or polypeptide that comprises an amino acid sequence according to any of claims 1 to 5 and one or more other immunoglobulin single variable domains, VHH's, (single) domain antibodies, Nanobodies or dAb's against human CXCR4.

11. Protein or polypeptide that comprises an amino acid sequence according to any of claims 1 to 5 and one or more other immunoglobulin single variable domains, VHH's, (single) domain antibodies, Nanobodies or dAb's against human CXCR4.

12. Protein or polypeptide that comprises an amino acid sequence according to any of claims 1 to 5 and another Nanobody that is a Nanobody against human CXCR4.

13. Protein or polypeptide according to any of claims 6 to 12 that has been provided with increased half-life, for example through suitable modification such as through pegylation, by fusion to albumin, by including a immunoglobulin single variable domain that can bind to serum albumin, or by attachment of a serum albumin binding peptide.

14. Nucleic acid sequence encoding the amino acid sequence of any of claims 1 to 5, or the protein or polypeptide of any of claims 6 to 13.

15. Nucleic acid sequence according to claim 14, wherein the nucleic acid sequence is selected from the group of nucleic acid sequences with SEQ ID NO's: 11 to 14.

16. Pharmaceutical composition comprising the amino acid sequence of any of claims 1 to 5, or the protein or polypeptide of any of claims 6 to 13 and optionally a pharmaceutically acceptable excipient.

17. The amino acid sequence of any of claims 1 to 5, or the protein or polypeptide of any of claims 6 to 13 for use as a medicament.

18. The amino acid sequence of any of claims 1 to 5, or the protein or polypeptide of any of claims 6 to 13 for use as a medicament to inhibit signaling that is mediated by human CXCR4 and/or its ligand(s); and/or in the prevention or treatment of diseases associated with an increased signalling of CXCR4, such as the various diseases in the group of cancer such as hematopoietic cancers like CLL, AML, ALL, MM, Non-Hodgkin lymphoma, solid tumors such as breast cancer, lung cancer, brain tumors, ovarian cancer, stromal chemoresistance of tumors, leukemia and other cancers, disrupting adhesive stromal interactions that confer tumor cell survival and drug resistance, mobilizing tumor cells form tissue sites and making them better accessible to

conventional therapy, inhibiting of migration and dissemination of tumor cells (metastasis), inhibiting or paracrine growth and survival signals, inhibiting pro-angiogenesis effects of SDF-1, inflammation and inflammatory disorders such as bowel diseases (colitis, Crohn'disease, IBD), infectious diseases, psioriasis, autoimmune diseases (such as MS), sarcoidosis, transplant rejection, cystic fibrosis, asthma, chronic obstructive pulmonary disease, rheumatoid arthritis, viral infection, HIV, West Nile Virus encephalitis, common variable immunodeficiency. Furthermore, the amino acid sequences of the invention can be used for stem cell mobilization in various patients in need of stem cells after X-ray radiation such as e.g. cancer patients after radiation treatment to replenish the stem cell pool after radiation in cancer patients, or in patients in need of more stem cells, e.g. in patients with ischemic diseases such as myocardial infarction (MI), stroke and/or diabetes (i.e. patients in need of tissue repair) wherein more stem cell would be re-transfused (after mobilization, screening, selection for lineage in need (e.g. cardiac, vascular lineages) and ex-vivo expansion of patient's own stem cells).

19. A method of inhibiting signaling that is mediated by human CXCR4 in a human suffering from a diseases selected from cancers like CLL, AML, ALL, MM, Non-Hodgkin lymphoma, solid tumors such as breast cancer, lung cancer, brain tumors, ovarian cancer, stromal chemoresistance of tumors, leukemia and other cancers, disrupting adhesive stromal interactions that confer tumor cell survival and drug resistance, mobilizing tumor cells form tissue sites and making them better accessible to conventional therapy, inhibiting of migration and dissemination of tumor cells (metastasis), inhibiting or paracrine growth and survival signals, inhibiting pro-angiogenesis effects of SDF-1, inflammation and inflammatory disorders such as bowel diseases (colitis, Crohn'disease, IBD), infectious diseases, psioriasis, autoimmune diseases (such as MS), sarcoidosis, transplant rejection, cystic fibrosis, asthma, chronic obstructive pulmonary disease, rheumatoid arthritis, viral infection, HIV, West Nile Virus encephalitis , common variable immunodeficiency. Furthermore, the amino acid sequences of the invention can be used for stem cell mobilization in various patients in need of stem cells after X-ray radiation such as e.g. cancer patients after radiation treatment to replenish the stem cell pool after radiation in cancer patients, or in patients in need of more stem cells, e.g. in patients with ischemic diseases such as myocardial infarction (MI), stroke and/or diabetes (i.e. patients in need of tissue repair) wherein more stem cell would be re-transfused (after mobilization, screening, selection for lineage in need (e.g. cardiac, vascular lineages) and ex-vivo expansion of patient's own stem cells).

20. Host cell comprising the nucleic acid of any of claims 14 and 15.

21. Method for the production of an amino acid sequence according to claim 1, wherein the substitutions in the framework region of SEQ ID NO: 3 are such that the expression level is improved by 2 fold while the IC50 value in the chemotaxis assay is 0.8 nM or lower.

22. Method for producing an amino acid sequence of any of claims 1 to 5, or the protein or polypeptide of any of claims 6 to 13, at least comprising the steps of culturing a host cell to produce said amino acid sequence or said protein or polypeptide comprising:

a) cultivating said host cell in a culture medium under conditions that are such that said host cell will multiply;

b) maintaining said host cell under conditions that are such that said host cell expresses and/or produces said amino acid sequence or said protein or polypeptide;

optionally followed by

c) secreting said amino acid sequence or said protein or polypeptide into the culture medium by said host cell; and

d) isolating and/or purifying the secreted amino acid sequence or protein or polypeptide from the culture medium.

23. The method according to claim 22, further applying conditions that promote the formation of disulfide bridges in and/or after step a), step b), step c), and/or step d).

Figure 1: Comparison of human V_H3 and Camelid V_{HH}'s - Framework 1:

Pos.	Amino acid residue(s):		V_{HH} Ent.	V_{HH} Var.
	Human V_H3	*Camelid V_{HH}'s*		
1	**E**, Q	**Q**, A, E	-	-
2	**V**	**V**	0.2	1
3	**Q**	**Q**, K	0.3	2
4	**L**	**L**	0.1	1
5	**V**, L	**Q**, E, L, V	0.8	3
6	**E**	**E**, D, Q, A	0.8	4
7	**S**, T	**S**, F	0.3	2
8	**G**, R	**G**	0.1	1
9	**G**	**G**	0	1
10	**G**, V	**G**, D, R	0.3	2
11	**Hallmark residue:** L, M, S, V,W; preferably L		0.8	2
12	**V**, I	**V**, A	0.2	2
13	**Q**, K, R	**Q**, E, K, P, R	0.4	4
14	**P**	**A**, Q, A, G, P, S, T, V	1	5
15	**G**	**G**	0	1
16	**G**, R	**G**, A, E, D	0.4	3
17	**S**	**S**, F	0.5	2
18	**L**	**L**, V	0.1	1
19	**R**, K	**R**, K, L, N, S, T	0.6	4

Figure 1: Comparison of human V_H3 and Camelid V_{HH}'s - Framework 1 (continued):

20	**L**	**L**, E, I, V	0.5	4
21	**S**	**S**, A, F, T	0.2	3
22	**C**	**C**	0	1
23	**A**, T	**A**, D, E, P, S, T, V	1.3	5
24	**A**	**A**, I, L, S, T, V	1	6
25	**S**	**S, A, F, P, T**	0.5	5
26	**G**	**G, A, D, E, R, S, T, V**	0.7	7
27	**F**	**S, F, R, L, P, G, N,**	2.3	13
28	**T**	**N, T, E, D, S, I, R, A, G, R, F, Y**	1.7	11
29	**F, V**	**F,L, D, S, I, G, V, A**	1.9	11
30	**S, D, G**	**N, S, E, G, A, D, M, T**	1.8	11

Figure 2: Comparison of human V_H3 and Camelid V_{HH}'s - Framework 2:

Pos.	Amino acid residue(s):		V_{HH} Ent.	V_{HH} Var.
	Human V_H3	*Camelid V_{HH}'s*		
36	**W**	**W**	0.1	1
37	**Hallmark residue:** $F^{(1)}$, H, I, L, Y or V, preferably $F^{(1)}$ or Y		1.1	6
38	**R**	**R**	0.2	1
39	**Q**	**Q**, H, P, R	0.3	2
40	**A**	**A**, F, G, L, P, T, V	0.9	7
41	**P**, S, T	**P**, A, L, S	0.4	3
42	**G**	**G**, E	0.2	2
43	**K**	**K**, D, E, N, Q, R, T, V	0.7	6
44	**Hallmark residue:** $G^{(2)}$, $E^{(3)}$, A, D, Q, R, S, L; preferably $G^{(2)}$, $E^{(3)}$ or Q; most preferably $G^{(2)}$ or $E^{(3)}$.		1.3	5
45	**Hallmark residue**: $L^{(2)}$, $R^{(3)}$, C, I, L, P, Q, V; preferably $L^{(2)}$ or $R^{(3)}$		0.6	4
46	**E**, V	**E**, D, K, Q, V	0.4	2
47	**Hallmark residue:** $W^{(2)}$, $L^{(1)}$ or $F^{(1)}$, A, G, I, M, R, S, V or Y; preferably $W^{(2)}$, $L^{(1)}$, $F^{(1)}$ or R		1.9	9
48	**V**	**V**, I, L	0.4	3
49	**S**, A, G	**A**, S, G, T, V	0.8	3

Figure 3: Comparison of human V_H3 and Camelid V_{HH}'s - Framework 3:

Pos.	**Amino acid residue(s):**		V_{HH} Ent.	V_{HH} Var.
	Human V_H3	*Camelid V_{HH}'s*		
66	**R**	**R**	0.1	1
67	**F**	**F**, L, V	0.1	1
68	**T**	**T**, A, N, S	0.5	4
69	**I**	**I**, L, M, V	0.4	4
70	**S**	**S**, A, F, T	0.3	4
71	**R**	**R**, G, H, I, L, K, Q, S, T, W	1.2	8
72	**D**, E	**D**, E, G, N, V	0.5	4
73	**N**, D, G	**N**, A, D, F, I, K, L, R, S, T, V, Y	1.2	9
74	**A**, S	**A**, D, G, N, P, S, T, V	1	7
75	**K**	**K**, A, E, K, L, N, Q, R	0.9	6
76	**N**, S	**N**, D, K, R, S, T, Y	0.9	6
77	S, T, I	**T**, A, E, I, M, P, S	0.8	5
78	**L**, A	**V**, L,A, F, G, I, M	1.2	5
79	**Y**, H	**Y**, A, D, F, H, N, S, T	1	7
80	**L**	**L**, F, V	0.1	1
81	**Q**	**Q**, E, I, L, R, T	0.6	5
82	**M**	**M**, I, L, V	0.2	2
82a	**N**, G	**N**, D, G, H, S, T	0.8	4
82b	**S**	**S**, N, D, G, R, T	1	6
82c	**L**	**L**, P, V	0.1	2

Figure 3: Comparison of human V_H3 and Camelid V_{HH}'s - Framework 3: (continued)

83	**Hallmark residue:** R, K[(5)], N, E[(5)], G, I, M, Q or T; preferably K or R; most preferably K		0.9	7
84	**Hallmark residue:** P[(5)], A, D, L, R, S, T, V; preferably P		0.7	6
85	**E**, G	**E**, D, G, Q	0.5	3
86	**D**	**D**	0	1
87	**T**, M	**T**, A, S	0.2	3
88	**A**	**A, G, S**	0.3	2
89	**V, L**	**V, A, D, I, L, M, N, R, T**	1.4	6
90	**Y**	**Y, F**	0	1
91	**Y, H**	**Y, D, F, H, L, S, T, V**	0.6	4
92	**C**	**C**	0	1
93	**A, K, T**	**A, N, G, H, K, N, R, S, T, V, Y**	1.4	10
94	**K, R, T**	**A, V, C, F, G, I, K, L, R, S or T**	1.6	9

Figure 4: Comparison of human V_H3 and Camelid V_{HH}'s - Framework 4:

Pos.	**Amino acid residue(s):**		V_{HH} Ent.	V_{HH} Var.
	Human V_H3	*Camelid V_{HH}'s*		
103	**Hallmark residue:** $W^{(4)}$, $P^{(6)}$, $R^{(6)}$, S; preferably W		0.4	2
104	**Hallmark residue:** G or D; preferably G		0.1	1
105	**Q**, R	**Q**, E, K, P, R	0.6	4
106	**G**	**G**	0.1	1
107	**T**	**T**, A, I	0.3	2
108	**Hallmark residue:** Q, $L^{(7)}$ or R; preferably Q or $L^{(7)}$		0.4	3
109	**V**	**V**	0.1	1
110	**T**	**T**, I, A	0.2	1
111	**V**	**V**, A, I	0.3	2
112	**S**	**S**, F	0.3	1
113	**S**	**S**, A, L, P, T	0.4	3

Figure 5

```
                          10        20        30        40        50            60

Kabat#          : .........|.........|.........|.........|.........|..abcde.......|

VH3-23/JH5      : EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLEWVSAISG----SGGSTYYA
4CXCR238D2      : ....V........T.................................G.KS----..D..R..
4CXCR238D4      : ....M........A............R..NN...G.F.R....ER.F.AAITRSGVR..V.AI.G
                            70        80           90        100           110

Kabat#          : .........|.........|..abc.......|.........|abcdefg.........|...

VH3-23/JH5      : DSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAK---------------WGQGTLVTVSS
4CXCR238D2      : G............A..M.....Y..KP..........SRVSRTGLYTY--DNR....Q.....
4CXCR238D4      : ....D........A............KP......T..ASAIGSGALRRFEYDYS....Q.....
```

Figure 6

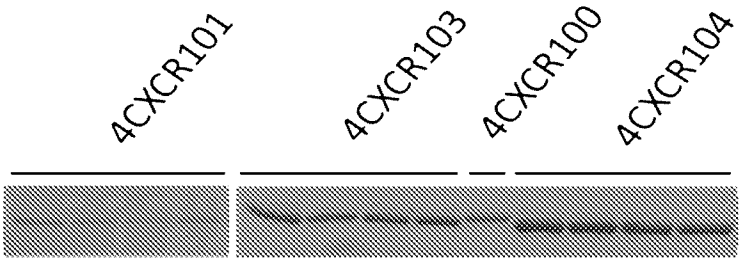


Figure 7

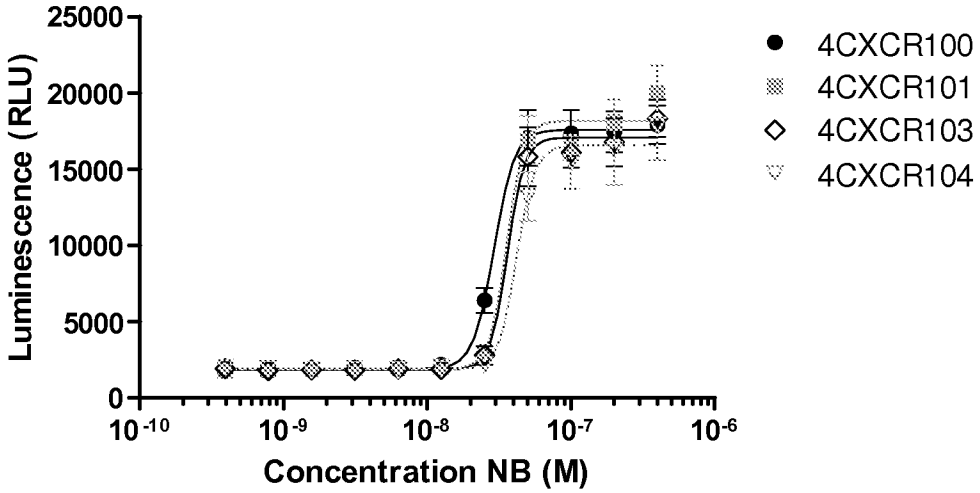


Figure 8:

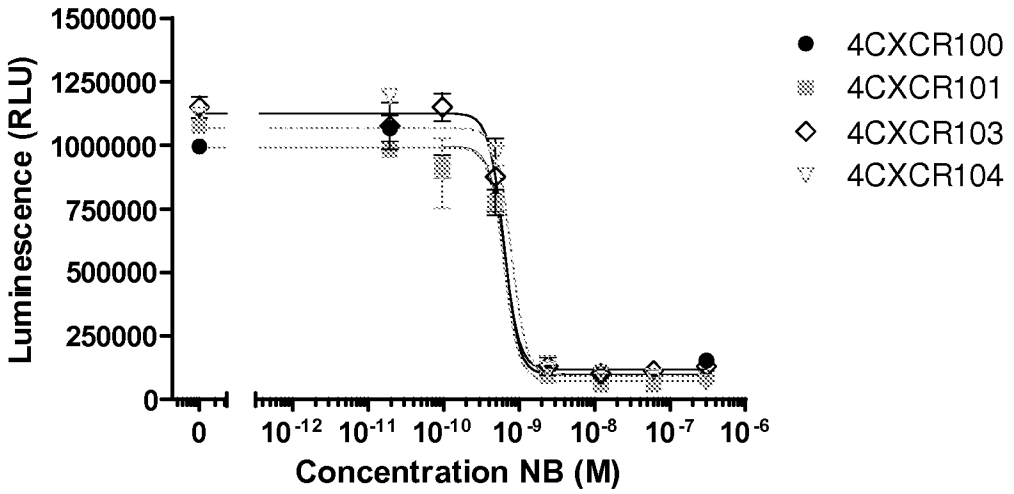
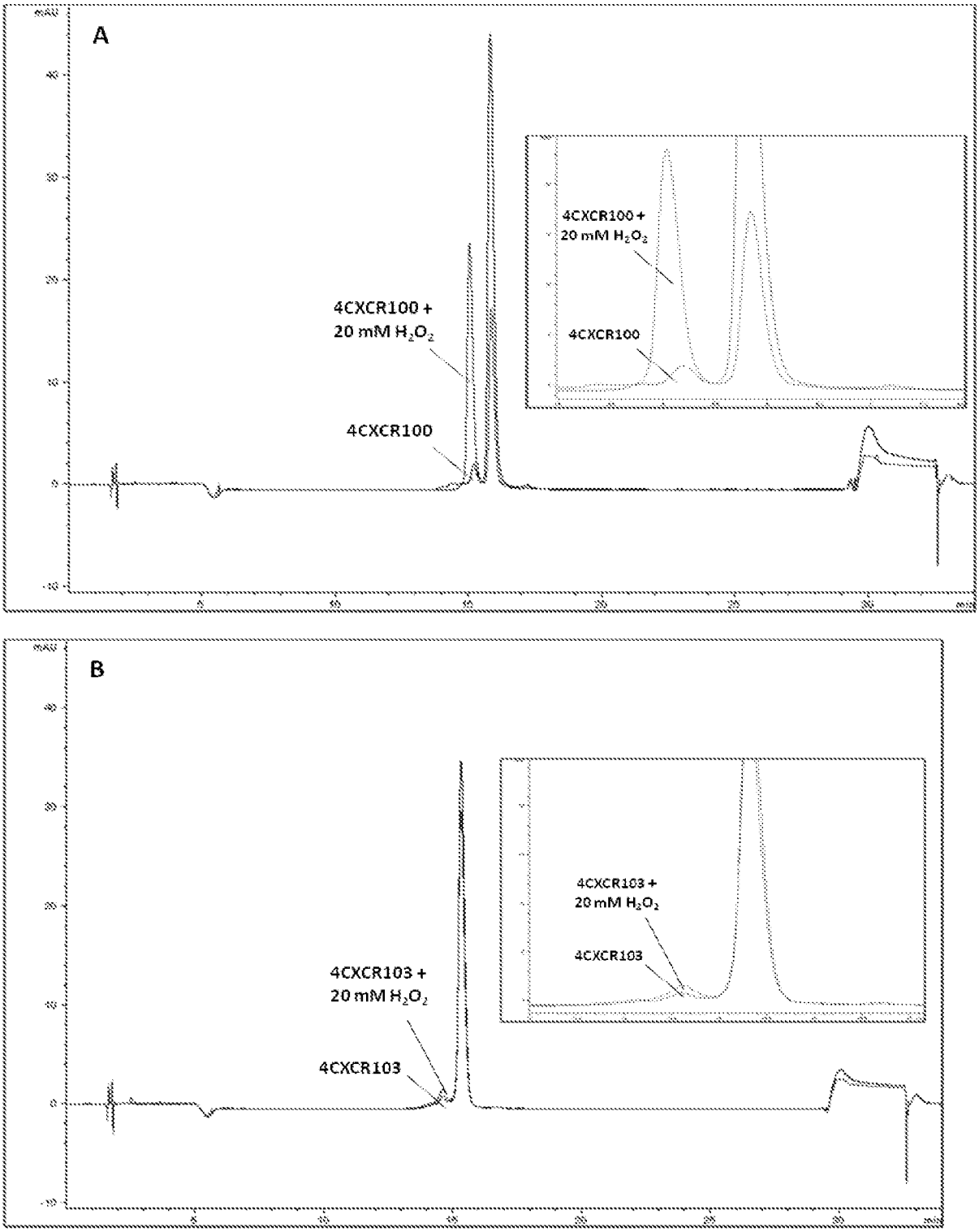
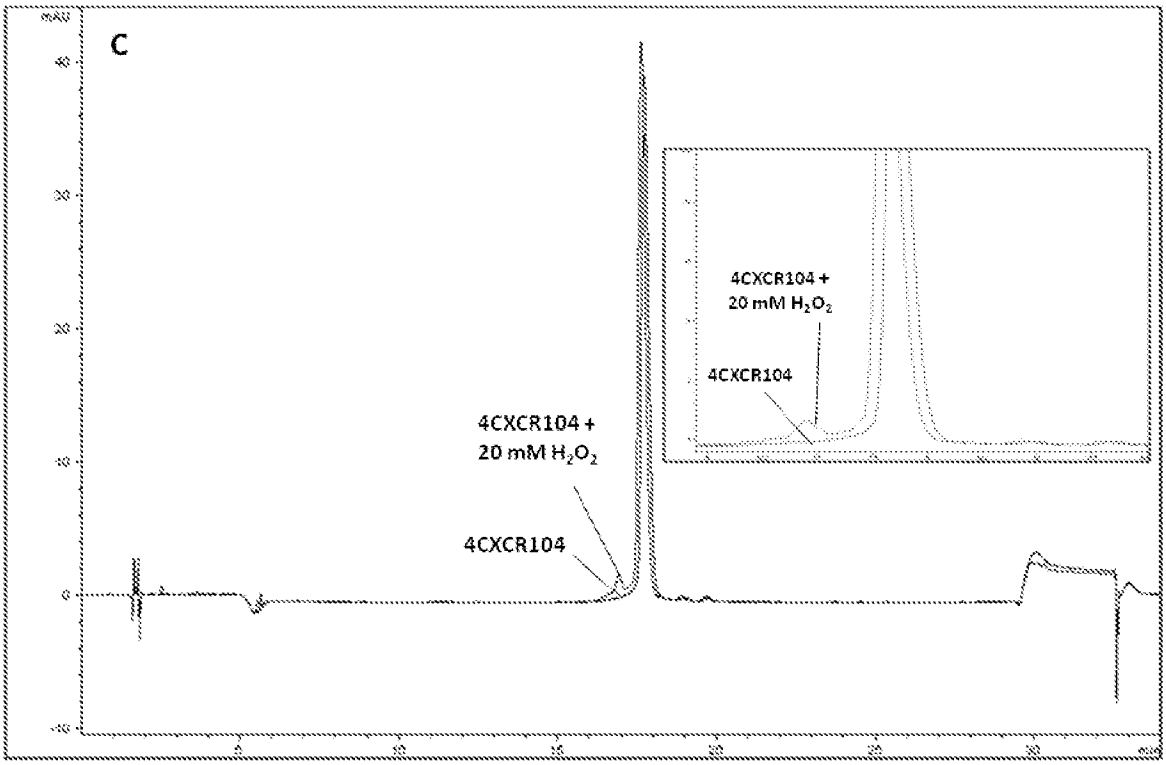

1500000
1250000
1000000
750000
500000
250000
0
Luminescence (RLU)
0
10-12
10-11
10-10
10-9
10-8
10-7
10-6
Concentration NB (M)
4CXCR100
4CXCR101
4CXCR103
4CXCR104

Figure 9:


A
4CXCR100 +
20 mM H_2O_2
4CXCR100
4CXCR100 +
20 mM H_2O_2
4CXCR100
B
4CXCR103 +
20 mM H_2O_2
4CXCR103
4CXCR103 +
20 mM H_2O_2
4CXCR103


C
4CXCR104 +
20 mM H_2O_2
4CXCR104
4CXCR104 +
20 mM H_2O_2
4CXCR104

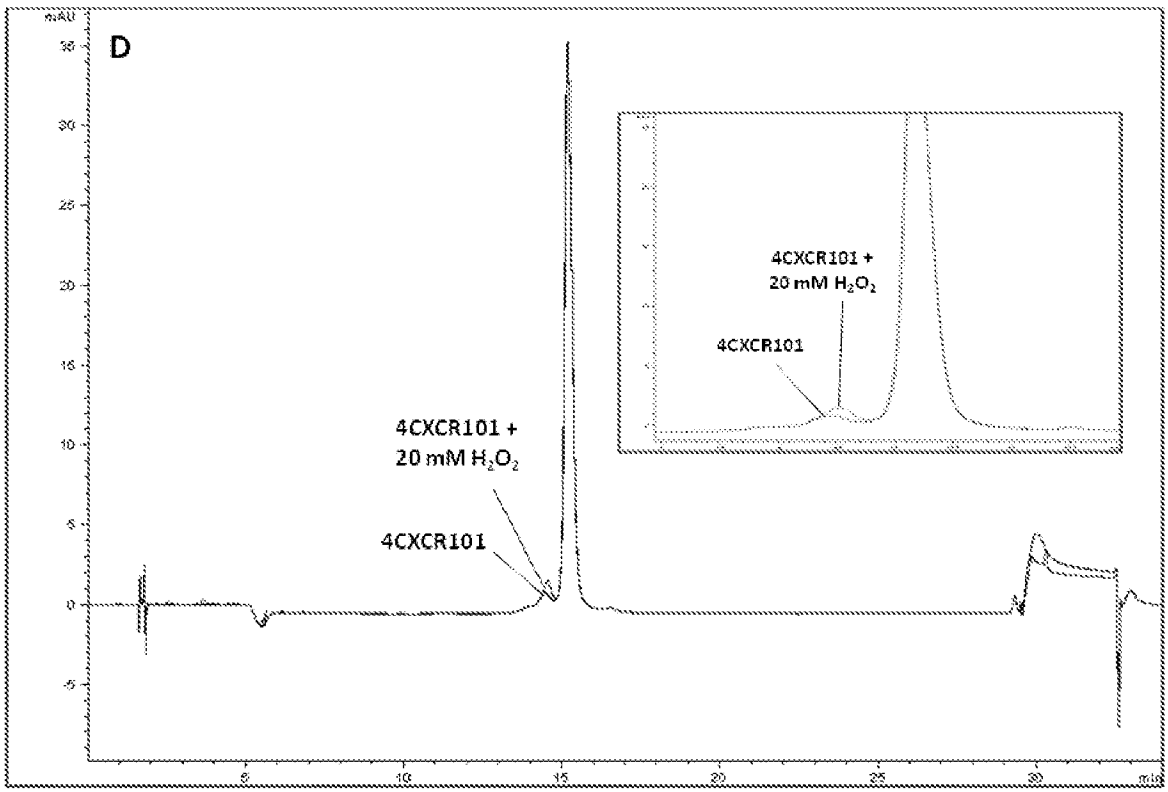
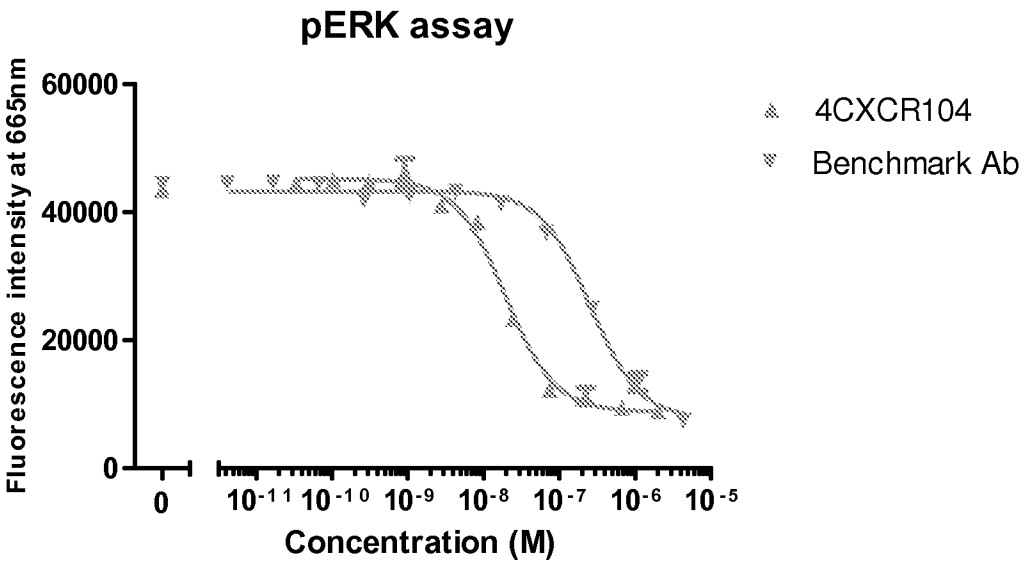

D
4CXCR101 +
20 mM H_2O_2
4CXCR101
4CXCR101 +
20 mM H_2O_2
4CXCR101

FIGURE 10


pERK assay
Fluorescence intensity at 665nm
60000
40000
20000
0
0
10-11 10-10 10-9 10-8 10-7 10-6 10-5
Concentration (M)
4CXCR104
Benchmark Ab

5

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2011/060738

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K16/28 C07K16/00 A61P37/00
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2010/043650 A2 (ABLYNX NV [BE]; BLANCHETOT CRISTOPH [NL]; SMIT MARTINE [NL]; LEURS REG) 22 April 2010 (2010-04-22) the whole document -----	1,7-23
X	WO 2009/138519 A1 (ABLYNX NV [BE]; BLANCHETOT CHRISTOPH [NL]; SMIT MARTINE [NL]; LEURS RE) 19 November 2009 (2009-11-19) the whole document -----	1,7-23
	-/--	

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
4 October 2011	14/10/2011

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Lechner, Oskar

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	XU CHEN ET AL: "Human anti-CXCR4 antibodies undergo VH replacement, exhibit functional V-region sulfation, and define CXCR4 antigenic heterogeneity.", JOURNAL OF IMMUNOLOGY (BALTIMORE, MD. : 1950) 15 AUG 2007 LNKD- PUBMED:17675502, vol. 179, no. 4, 15 August 2007 (2007-08-15), pages 2408-2418, XP002610999, ISSN: 0022-1767 the whole document page 2416, right-hand column, paragraph 2-3	1-16
X	CARNEC XAVIER ET AL: "Anti-CXCR4 monoclonal antibodies recognizing overlapping epitopes differ significantly in their ability to inhibit entry of human immunodeficiency virus type 1.", JOURNAL OF VIROLOGY FEB 2005 LNKD- PUBMED:15650218, vol. 79, no. 3, February 2005 (2005-02), pages 1930-1933, XP2234668, ISSN: 0022-538X abstract; figure 4; table 1	1-16
X	BARIBAUD F ET AL: "Antigenically distinct conformations of CXCR4.", JOURNAL OF VIROLOGY OCT 2001 LNKD- PUBMED:11533159, vol. 75, no. 19, October 2001 (2001-10), pages 8957-8967, XP002518542, ISSN: 0022-538X the whole document figure 6; table 2	1-16
A	WO 2008/060367 A2 (MEDAREX INC [US]; KUHNE MICHELLE [US]; BRAMS PETER [US]; TANAMACHI DAW) 22 May 2008 (2008-05-22) abstract	1-16
A	BRELOT A ET AL: "Identification of residues of CXCR4 critical for human immunodeficiency virus coreceptor and chemokine receptor activities.", THE JOURNAL OF BIOLOGICAL CHEMISTRY 4 AUG 2000 LNKD- PUBMED:10825158, vol. 275, no. 31, 4 August 2000 (2000-08-04), pages 23736-23744, XP002611000, ISSN: 0021-9258 the whole document	1-16

-/--

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WONG REBECCA S Y ET AL: "Comparison of the potential multiple binding modes of bicyclam, monocylam, and noncyclam small-molecule CXC chemokine receptor 4 inhibitors.", MOLECULAR PHARMACOLOGY DEC 2008 LNKD-PUBMED:18768385, vol. 74, no. 6, December 2008 (2008-12), pages 1485-1495, XP002611001, ISSN: 1521-0111 the whole document	1-16

T	"Data sheet of monoclonal antibody MAB173-anti CXCR4", INTERNET CITATION , 13 November 2007 (2007-11-13), XP002458711, Retrieved from the Internet: URL:http://www.rndsystems.com/pdf/mab173.pdf [retrieved on 2007-11-13] the whole document	

A	VADAY GAYLE G ET AL: "CXCR4 and CXCL12 (SDF-1) in prostate cancer: inhibitory effects of human single chain Fv antibodies", CLINICAL CANCER RESEARCH, THE AMERICAN ASSOCIATION FOR CANCER RESEARCH, US, vol. 10, no. 16, 15 August 2004 (2004-08-15), pages 5630-5639, XP002397863, ISSN: 1078-0432, DOI: DOI:10.1158/1078-0432.CCR-03-0633 the whole document	1-16

A	KHAN ABID ET AL: "Small molecule CXCR4 chemokine receptor antagonists: developing drug candidates.", CURRENT MEDICINAL CHEMISTRY 2007 LNKD-PUBMED:17896975, vol. 14, no. 21, 2007, pages 2257-2277, XP002611002, ISSN: 0929-8673 the whole document	1-16

	-/--	

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	HARMSEN M M ET AL: "Properties, production, and applications of camelid single-domain antibody fragments.", APPLIED MICROBIOLOGY AND BIOTECHNOLOGY NOV 2007 LNKD- PUBMED:17704915, vol. 77, no. 1, November 2007 (2007-11), pages 13-22, XP002611003, ISSN: 0175-7598 the whole document	1-16
A	ROOVERS R C ET AL: "Nanobodies in therapeutic applications", CURRENT OPINION IN MOLECULAR THERAPEUTICS, CURRENT DRUGS, LONDON, GB, vol. 9, no. 4, 1 January 2007 (2007-01-01) , pages 327-335, XP009093747, ISSN: 1464-8431 the whole document	1-16
T	JÄHNICHEN S: "CXCR4 nanobodies (VHH-based single variable domains) potently inhibit chemotaxis and HIV-1 replication and mobilize stem cells.", PROC NATL ACAD SCI U S A, vol. 107, no. 47, November 2010 (2010-11), pages 20565-20570, XP002611004, the whole document	1-16
T	BURGER JAN A ET AL: "Potential of CXCR4 antagonists for the treatment of metastatic lung cancer", EXPERT REVIEW OF ANTICANCER THERAPY, FUTURE DRUGS LTD, UK, vol. 11, no. 4, 1 April 2011 (2011-04-01), pages 621-630, XP009152669, ISSN: 1744-8328	1-16

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2011/060738

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2010043650 A2	22-04-2010	EP 2352764 A2	10-08-2011
WO 2009138519 A1	19-11-2009	AU 2009248049 A1	19-11-2009
		CA 2724208 A1	19-11-2009
		CN 102099378 A	15-06-2011
		EP 2285833 A1	23-02-2011
		JP 2011523550 A	18-08-2011
		KR 20110020825 A	03-03-2011
		US 2011206660 A1	25-08-2011
WO 2008060367 A2	22-05-2008	AR 063086 A1	23-12-2008
		AU 2007320024 A1	22-05-2008
		CA 2665239 A1	22-05-2008
		CL 28352007 A1	30-05-2008
		CN 101528259 A	09-09-2009
		EA 200900424 A1	28-08-2009
		EP 2066351 A2	10-06-2009
		JP 2010505830 A	25-02-2010
		KR 20090064589 A	19-06-2009
		US 2010104508 A1	29-04-2010