A Novel Transmembrane CXCR4 Variant That Expands the WHIM Genotype-Phenotype Paradigm: Supplementary Materials

Supplementary Results

Molecular dynamics simulations

D84 represents $D^{2\times50}$, the most conserved amino acid in helix II of class A G protein–coupled receptors (GPCRs), such as CXCR4. D84 is a critical, conserved residue of the Na+ ion binding pocket, which is essential for physiologic signaling of class A GPCRs.[1,2] Variants in the $D^{2\times50}$ position have been implicated in diseases such as the melanocortin-4 receptor $MC4R^{D90N}$ in severe early-onset obesity[3] and type 2 vasopressin receptor $V2R^{D85N}$ in nephrogenic diabetes insipidus.[4] (Please note that the numbers 2×50, 3×39, 3×35, etc, refer to generic numbering of amino acid positions in GPCRs according to Ballesteros and Weinstein.[5,6])

At the initiation of molecular dynamics (MD) simulations, the $CXCR4^{WT}$ and $CXCR4^{D84H}$ receptor contained Na^+ in the allosteric site (see Supplementary Methods). The ion remained bound at the allosteric site throughout the simulation course for 900 ns in wild-type CXCR4 ($CXCR4^{WT}$) and was coordinated by $D84^{2\times50}$, $S123^{3\times39}$, and $N119^{3\times35}$ with additional stabilization of this coordination sphere mediated by hydrogen bonding of the amide N-H of N119 to the carboxylate moiety in D84 (Supplementary Figure 3A and Supplementary Video 1). Located just above this Na^+ coordination sphere was a hydrophobic gate comprising amino acid residues $Y255^{6\times51}$, $Y116^{3\times32}$, $W252^{6\times48}$, and $F87^{2\times53}$, which impedes water entry and solvolysis of this coordinated Na^+, further stabilizing CXCR4 in its inactivated state. In contrast, the H84 histidine moiety of $CXCR4^{D84H}$ less efficiently coordinates Na^+ in the allosteric pocket and is devoid of the N119-D84 H-bonding interaction seen in $CXCR4^{WT}$, resulting in destabilization and opening of the hydrophobic gate, water-mediated solvolysis, and ejection of Na^+ toward the extracellular region of $CXCR4^{D84H}$ at 25 ns after simulation initiation (Supplementary Figure 3A and Supplementary Video 2).

Impaired access of Na^+ to the allosteric site has been linked to destabilization of the inactive receptor form and constitutive activity of $CXCR4^{N119A}$ and $CXCR4^{N119S}$ variants.[7,8] Thus, the D84H conformational change toward the active receptor state was investigated by analyzing rearrangements in residue contacts that constitute a conserved signal initiation phase in class A GPCRs, mainly characterized by collapse of the allosteric binding pocket.[1,2] Collapse of the Na^+

pocket leads to a denser repacking of the 4 residue pairs (2×46:2×50, 2×50:3×39, 2×50:7×45, and 7×45:7×49), initiating movement of transmembrane (TM) 7 toward TM3.[2,8] The residue-residue contact scores (RRCSs) between these residue pairs in $CXCR4^{WT}$ and $CXCR4^{D84H}$ during MD simulations were calculated and compared with RRCS in available CXCR4 crystal structures generated with antagonists.[9,10] The D84H substitution resulted in higher RRCSs and hence denser packing of all 4 residue pairs when compared with the $CXCR4^{WT}$ or x-ray structure of CXCR4 cocrystallized with antagonist IT1t (Supplementary Figure 3B). The Cα distances for the most conserved Na^+ coordinating residues $D^{2\times50}$ and $S123^{3\times39}$ were also determined. The measurement distributions for $CXCR4^{WT}$ and $CXCR4^{D84H}$ showed shortening of this distance in the presence of D84H in comparison with WT (Supplementary Figure 3C). Such increased proximity of $D^{2\times50}$ and $S^{3\times39}$ was observed in constitutively active variants of CXCR4 and in experimental structures of class A GPCRs in their active state.[7]

Supplementary Methods

Patients

The studies were approved by IRB of the respective institutions, including Institutional Review of Duke University (IRB Number: Pro00066839), Belarusian Research Center for Pediatric Oncology (IRB Number: 200809-RD), and University Medical Center Freiburg (IRB Number: 354/19)

Sample collection and processing

EDTA or heparin blood samples from P1, P2, and P4 were collected from consented participants at 3 academic centers. Ficoll-Hypaque density gradient centrifugation isolated peripheral blood mononuclear cells (PBMCs), followed by separation of plasma, red blood cell lysis, and buffy coat isolation according to standard methodologies.

Genetic testing

Genetic testing was performed using specific panels utilizing next-generation sequencing (Supplementary Table 1A).

Histopathologic studies

Peripheral blood smears for P1 were created in the Duke Central Automated Laboratory using routine clinical laboratory procedures. Briefly, Wright-Giemsa–stained peripheral blood smears were created using an SP-10 automated slide-making and slide-staining machine (Sysmex).

Images were obtained on an Olympus BX41 microscope equipped with Olympus UPlanFL N optic at the 100× oil magnification and 1.30 numerical aperture and were digitalized using Olympus DP23 camera and Olympus cellSens Entry software.

Plasmid constructs of CXCR4 wild-type and variants

Plasmid constructs pcDNA3.1_CXCR4^{WT} and pcDNA3.1_CXCR4^{D84H} were generated by cloning CXCR4 wild-type (CXCR4^{WT}) and CXCR4^{D84H} variant (CXCR4^{D84H}) complementary DNA (cDNA) into pcDNA3.1(Hygro+). The cDNA of human *CXCR4* and its variant p.D84H carrying the c.250G>C variation and harboring the consensus Kozak sequence at the 5′ end (GCCGCCACCatg) were synthesized by GENEWIZ (Leipzig, Germany). The wild-type (*CXCR4*^{WT}) and variant (*CXCR4*^{D84H}) cDNA were cloned into pcDNA3.1(Hygro+) via NheI/EcoRV restriction sites to obtain pcDNA3.1_CXCR4^{WT} and pcDNA3.1_CXCR4^{D84H}, respectively. The *CXCR4* variants (p.R334X and p.E343K) were generated by site-directed mutagenesis, using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies), the primers listed in Supplementary Table 1B, and pcDNA3.1_CXCR4 as template. Endotoxin-free plasmid DNA was purified (EndoFree Plasmid MaxiPrep Kit; Qiagen), and sequence of the construct was verified using CMV forward and BGH reverse primers (Microsynth AG).

Reagents

Mavorixafor was synthesized by ChemPartner and dissolved as a 10 mM stock solution in dimethyl sulfoxide (DMSO). Human CXCL12 (also known as stromal cell–derived factor 1 alpha, SDF1α) was purchased from PeproTech, Inc.

Cell culture, cell lines, and transient transfections

CXCR4-negative K562 cells (CCL-243; ATCC) were cultured in Iscove's Modified Dulbecco's Medium (IMDM; Gibco; Thermo Fisher Scientific Inc) supplemented with 10% fetal bovine serum (FBS; Sigma-Aldrich) and antibiotic-antimycotic (Gibco; Thermo Fisher Scientific Inc) agents. Cells were starved overnight in IMDM, 0.5% bovine serum albumin (BSA; Sigma-Aldrich), and antibiotic-antimycotic agents. Stable K562 cell lines expressing CXCR4 were established by electroporation (Gene Pulser Xcell Electroporation System; Bio-Rad Laboratories, Inc) with linearized (Bgl II restriction enzyme) pcDNA3.1-CXCR4 plasmids in Ingenio Electroporation Solution (Mirus Bio LLC). Stable transfected cells were selected for 14 days with 400 µg/mL hygromycin B (Invitrogen; Thermo Fisher Scientific Inc), followed by sorting of CXCR4-positive cells (BD Pharmingen 12G5-APC monoclonal antibody [mAb]; BD

Biosciences). Clonal cell lines with comparable CXCR4 expression were selected for further experiments.

Transiently expressing K562 cells (2.5×10^6 cells in 250 µL/cuvette) were established by transfecting 25 µg pcDNA3.1_CXCR4 plasmids per cuvette in Ingenio Solution (Mirus Bio LLC) by electroporation using a Gene Pulser Xcell device (BioRad Laboratories, Inc) and exponential decay protocol (250 V, 1000 µF) in 0.4-cm cuvettes and were immediately transferred to 6-well plates containing 2.25 mL preequilibrated culture media. The cells were used for further experiments 24 hours after transfection.

PBMCs were cultured in Roswell Park Memorial Institute (RPMI) 1640 medium (Gibco) supplemented with 20% FBS (Sigma-Aldrich), 100 U/mL penicillin, 100 µg/mL streptomycin (Gibco), and 2 mM L-glutamine (Gibco). The cells were serum-starved for 1 hour at 37 °C with 5% CO_2 in RPMI 1640 medium containing 1% FBS, 100 U/mL penicillin, streptomycin, and 2 mM L-glutamine at a final concentration of 5×10^5 cells/mL before the assays were performed.

Functional characterization assays

Internalization, chemotaxis, calcium mobilization, and cyclic adenosine monophosphate (cAMP) enzyme-linked immunosorbent assay (ELISA) were performed using PBMCs from healthy donors (HDs) (age- and sex-matched shipping control [HD-SC] and in-house control [HD-IH]), 3 patients (P1, P2, and P4) harboring a heterozygous $CXCR4^{D84H}$, and transfected K562 cells.[11] CXCR4 expression was measured via flow cytometry (CytoFLEX Flow Cytometer) and analyzed in FCS Express software, as previously published.[11]

Internalization assay

Transiently transfected K562 cells were seeded 1×10^5 cells/well in 96-well plates 24 hours after transfection and serum-starved overnight. Stable K562 clones expressing CXCR4 (1×10^5 cells/well) were seeded in 96-well plates and serum-starved for 24 hours. Cells were then resuspended in warm incubation buffer (Hanks' Balanced Salt Solution [HBSS] with Ca^{2+} and Mg^{2+} + 0.5% BSA + 20 mM HEPES buffer pH 7.4) and stimulated with CXCL12 for 45 minutes or 4 hours at 37 °C with 5% CO_2. After incubation, the cells were washed twice with cold incubation buffer and then stained with anti-CXCR4 12G5-APC mAb (BD Biosciences; 1:20 dilution in incubation buffer) for 20 minutes at 4 °C. After being washed and resuspended in flow buffer (HBSS with Ca^{2+} and Mg^{2+} + 0.1% BSA + 20 mM HEPES buffer pH 7.4), the samples were measured via flow cytometry (CytoFLEX; Beckman Coulter Diagnostics) and analyzed

using FCS Express flow cytometry software (De Novo Software). Cells were gated based on the forward and side scatter and isotype control, and the mean fluorescence intensity (MFI) of the CXCR4-positive population was used in subsequent analysis.

Freshly isolated PBMCs from the heparinized whole blood of HDs and patients were suspended in a preequilibrated culture medium for 1 hour at 37 °C with 5% CO_2, after which 5 × 10^5 cells/mL were seeded in 96-well plates and serum-starved for 1 hour at 37 °C with 5% CO_2. Internalization was then carried out as described above. Lymphocyte subtyping was performed using fluorescent mAbs specific for T-cell antigens CD3, CD8, and CD4; B-cell antigen CD19; and natural killer–cell antigen CD56. Background fluorescence was evaluated using the corresponding immunoglobulin-isotype control mAbs. Cells were then washed and resuspended in flow buffer, and the samples were measured and analyzed via flow cytometry as described.

Chemotaxis assay

Transiently transfected K562 cells were serum-starved in IMDM supplemented with 0.5% BSA. After another 24 hours, chemotaxis was performed using 6.5-mm Transwell plates with 8.0-μm pore size (Corning Incorporated). Prior to chemotaxis, cells were stained with 500 nM Calcein AM Viability Dye (Thermo Fisher Scientific Inc) in IMDM supplemented with 0.5% BSA for 15 minutes at room temperature (RT) in the dark. Aliquots of 2.0 × 10^5 cells in 100 μL IMDM supplemented with 0.5% BSA and 20 μg/mL fibronectin (Sigma-Aldrich) were added to the plate inserts, and 600 μL buffer with or without indicated CXCL12 concentrations was added to the bottom wells. Cells were allowed to migrate in response to CXCL12 at 37 °C with 5% CO_2 for 4 hours. After removal of the plate inserts, the migrated cells were centrifuged and resuspended in Dulbecco's PBS containing flow cytometry counting beads (Precision Count Beads; BioLegend, Inc).

Freshly isolated and serum-starved PBMCs were resuspended in chemotaxis buffer (RPMI 1640 media containing 20 mM HEPES, L-glutamine, and 0.5% BSA) at 2.0 × 10^6 cells/mL. Cells were pretreated with the indicated concentrations of mavorixafor for 30 minutes at 37 °C with 5% CO_2. Chemotaxis was assayed immediately after treatment with the drug by placing 1.0 × 10^5 cells in 100 μL in the upper chamber of a Transwell 24-well plate separated by a 6.5-mm insert with 3-μm pores (Corning Life Sciences) from the lower chamber containing 600 μL of buffer with the indicated concentration of CXCL12. After incubation for 2.5 hours at 37 °C in a 5% CO_2 incubator, the cells that migrated into the lower chamber were collected by centrifugation and resuspended in the assay buffer. Cells were blocked using Human TruStain

FcX (BioLegend, Inc) for 10 minutes at RT, and subsequently lymphocyte subtyping was performed. A fixed number of flow cytometry counting beads (Precision Count Beads; BioLegend) were added to each sample.

Both migrated cells and counting beads were counted by flow cytometry (CytoFLEX). Data were analyzed using FCS Express software, and the total number of migrated cells was calculated according to the counted and total number of beads present in the sample. The highest chemotactic response in PBMCs from HDs was achieved upon 10 nM CXCL12 stimulation, and in K562 cells transfected with *CXCR4*^{WT}, upon 2 nM CXCL12 stimulation. Therefore, chemotactic index was calculated as the number of cells migrating in each condition relative to the number of cells migrating in response to 10 nM CXCL12 in the corresponding HD for PBMCs or relative to the number of cells migrating in response to 2 nM CXCL12 transfected with *CXCR4*^{WT} for K562 cells.

Calcium mobilization assay

Stable K562 clones expressing CXCR4 (1.0×10^5 cells/well) were seeded in black 96-well plates with a transparent bottom coated with poly-L-lysine (BioCoat; Corning) and serum-starved for 24 hours. Medium was removed, and cells were loaded with 100 µL of Fluo-4 AM (3 µM; Invitrogen) dye solution for 45 minutes at 37 °C. Dye solution was prepared by diluting Fluo-4 powder in DMSO and 10% Pluronic F-127 (Thermo Fisher Scientific Inc) to 3 mM and then preparing the final dilution in assay buffer (HBSS with Ca^{2+} and Mg^{2+}, 20 mM HEPES, 0.375 g/L $NaHCO_3$, 0.1% BSA + 0.77 g/L probenecid [Thermo Fisher Scientific Inc]). Subsequently, 100 µL of assay buffer alone or assay buffer with compound dilutions was added, and the plates were equilibrated in the plate reader for an additional 20 minutes at 37 °C. PBMCs (3×10^5 cells/well) were incubated with the Screen Quest Calbryte-520 Probenecid-Free and Wash-Free Calcium Assay Kit (AAT Bioquest, Inc) according to the manufacturer's instructions. CXCL12 was injected with a simultaneous measurement of fluorescent signal (FlexStation 3 Multi-Mode Microplate Reader; Molecular Devices, LLC). Raw traces were analyzed in SoftMax Pro 7 Software (Molecular Devices, LLC). The relative fluorescence units were calculated as the difference between maximal and minimal signal after treatment injection, normalized to the baseline signal before injection. Maximal effect and half-maximal effective concentration were calculated in Prism (GraphPad Software).

cAMP enzyme-linked immunosorbent assay

Stable K562 clones expressing CXCR4 were used in ELISA. cAMP was detected in stable K562 clones expressing CXCR4 by using cAMP-Screen Cyclic AMP Immunoassay System (Applied Biosystems, Thermo Fisher Scientific Inc). The clones (2×10^5 cells) were starved overnight in 200 µL of starvation medium on a 96-well plate. The medium was removed, and cells were preincubated with 3-isobutyl-1-methylxanthine (0.5 mM; Sigma-Aldrich) diluted in assay buffer (HBSS with Ca^{2+} and Mg^{2+}, 20 mM HEPES, 0.375 g/L $NaHCO_3$, 0.1% BSA) for 15 minutes at 37 °C, after which 10 µM forskolin (Sigma-Aldrich) was added either alone or together with 100 nM CXCL12 for 30 minutes at 37 °C. The samples were then processed according to the kit manufacturer's instructions.

Phospho flow cytometry

Stable K562 clones expressing CXCR4 were seeded (5×10^5 cells/well) in starvation medium in a 96-well plate overnight. PBMCs were starved for 1 hour. After stimulation, the cells were fixed with fixation solution (eBioscience; Thermo Fisher Scientific Inc) for 10 minutes at RT and then permeabilized for 5 minutes with ice-cold methanol on ice, followed by a washing step with wash/permeabilization solution (eBioscience; Thermo Fisher Scientific Inc). Cells were then stained with Alexa Fluor 647 Mouse Anti-ERK1/2 (pT202/pY204, 1/10 dilution; BD Biosciences) and Alexa Fluor 488 Mouse anti-Akt (pS473, 1/10 dilution; BD Biosciences) for 1 hour at RT in darkness. Cells were then washed 2 times in wash/permeabilization solution and resuspended in flow buffer (HBSS with Ca^{2+} and Mg^{2+} + 0.1% BSA + 20 mM HEPES pH 7.4). Samples were measured via flow cytometry (CytoFLEX) and analyzed in FCS Express software. After gating cells based on the forward/side scatter, MFI of the respective staining was exported for further analysis.

Ligand binding inhibition assay

Ligand binding inhibition by mavorixafor was measured as previously published.[12] Stable K562 CXCR4-expressing cells were washed once with assay buffer (HBSS + 20 mM HEPES buffer + 0.2% BSA, pH 7.4) and then incubated for 15 minutes at RT with test compound diluted in assay buffer at dose-dependent concentrations. Subsequently, human CXCL12 Alexa Fluor 647 (26 ng/mL; Almac) was added to the compound-preincubated cells. The cells were incubated for 30 minutes at RT, after which the cells were washed twice in assay buffer, fixed with 1% paraformaldehyde in PBS, and analyzed by flow cytometry (CytoFLEX). MFI of CXCL12-AF647 was determined (FCS Express software). The percentage of inhibition was calculated according to the formula $(1 - [(MFI - MFI_{NC})/(MFI_{PC} - MFI_{NC})]) * 100$, where *MFI* is the MFI of CXCR4-expressing cells in the presence of an inhibitor, MFI_{NC} is the MFI of mock-transfected cells in the

presence of the ligand, and MFI_{PC} is the MFI of CXCR4-expressing cells in the presence of the ligand alone. Half maximal inhibitory concentration (IC_{50}) was calculated in Prism (GraphPad Software).

Model building

The starting model was built based on the crystal structure of human CXCR4 (RCSB Protein Data Bank 3ODU), with the small-molecule antagonist in the crystal structure removed. The N-terminal and the C-terminal were capped with an acetyl(ACE) group and N-methylamide (NME) groups after truncating a disorder random loop (GLY306–GLN328), respectively. The T4 lysozyme fusion inserted between TM5 and TM6 of CXCR4 was also removed, and the 2 residues, SER229 and LYS230 at the T4 lysozyme junction site, were reconnected and minimized. In addition, thermostabilizing L125W variant was reversed. Molecular operating environment was used to add hydrogens to the initial coordinates of the protein, set protonation states, and correct bond orders.

To build the membrane simulation box, PACKMOL-Memgen was used to embed the processed protein into the lipid bilayer.[13] The orientation of CXCR4 in the palmitoyl-oleoyl-phosphatidylcholine (POPC) lipid bilayer was carefully adjusted so that the regions exposed to the solvent and buried in the lipids matched the annotation of CXCR4 topology from UniProt (P61073). The sodium ion at the allosteric site was placed by overlaying a crystal structure of β1-adrenergic receptor (RCSB Protein Data Bank 4BVN) with both the wild-type and variant *CXCR4*.

Molecular dynamic (MD) simulations and RRCS calculations

All MD simulations were performed with GROMACS 2021.[14] The Amber force field ff19SB[15] and the TIP3P water model[16] were used. The Amber LIPID17 force field and the Joung-Cheatham force field parameters[17] were also used for the lipids and the ions, respectively. The CXCR4 protein in the POPC lipid bilayer was solvated with 15 Å buffer of water in a rectangular box and neutralized with counterions. In addition, 0.15 M NaCl was added to mimic the salt concentration under physiologic conditions. The particle mesh Ewald method[18] was employed to calculate electrostatic energies. The cutoff of nonbonded interactions was set to 10 Å, and the long-range dispersion correction was used for van der Waals interactions.

The whole system was first minimized using a steepest descent algorithm, with the maximal number of cycles set as 50,000. Subsequently, the system was heated from 297 K to

310 K over a period of 500 ps with position restraints of 200 kcal/mol, and this was followed by an isothermal–isobaric (NPT) equilibration for 10 ns without any position restraint.

The production phase was simulated in an NPT ensemble with the Parrinello-Rahman barostat for 300 ns. Temperature was controlled with a Nosé-Hoover thermostat at 310 K, and the collision frequency was 2.0-1 ps. Pressure was controlled by semi-isotropic position scaling, and the pressure relaxation time was set to 5.0 ps. Hydrogen atoms were constrained using the LINear Constraint Solver algorithm.[19] A time step of 2 fs was used for the production runs. The production phases for the WT and D84H systems were all 900 ns.

RRCS calculations were performed as previously published.[2]

Prevalence estimations

Prevalence of D84H was calculated using average allele frequencies reported in Genome Aggregation Database v2.1,[20,21] Trans-Omics for Precision Medicine Freeze 8 via BRAVO,[22] and the National Heart, Lung, and Blood Institute Exome Sequencing Project.[23]

Statistical analysis

Statistical analysis was performed using data generated from K562 cells and PBMCs for the following end points: CXCR4 receptor internalization, cAMP production, chemotactic index, and phosphorylated protein kinase B (pAKT) and extracellular signal-regulated kinase (pERK). A 2-sample Mann-Whitney nonparametric test was used in most cases comparing cells expressing the indicated CXCR4 variant with WT cells. Chemotactic index at 2 nM CXCL12 was evaluated using a 1-sample Wilcoxon signed-rank test to determine whether the median of normalized value was significantly different from each variant. Since statistical testing was intended for hypothesis generation, no adjustments were made for multiple testing. Statistical analyses were conducted using GraphPad Prism software, version 9.0.1. *P* values <.05 were considered statistically significant and set as **P*<.05; ***P*<.01; and ****P*<.001. The number of independent experiments is stated in each figure legend. Calculation of maximum effect, half maximal effective concentration, and half maximal inhibitory concentration (IC_{50}) parameters of CXCL12 or mavorixafor was performed in GraphPad Prism version 9.0.1 by log (agonist/antagonist) vs response-variable slope (4 parameters) function.

Supplementary Tables

Supplementary Table 1. (A) Genetic testing panels used to screen patients. (B) Primers used to generate plasmids for WT and variant genes. The *CXCR4* variants were generated by site-directed mutagenesis, using the QuikChange II XL Site-Directed Mutagenesis Kit, the primers listed, and pcDNA3.1_CXCR4 as template.

A

Patient	Tests	Number of genes	Genes tested
P1	Bone Marrow Failure Syndromes Panel (Cincinnati Children's Hospital, Cincinnati, Ohio)	118	*AP3BI, BRCA2, BRIP1 (FANCJ), CSF3R, CXCR4, DKC1, ELAHE (previously ELA2), ERCC4, FANCA, FANCB, FANCC, FANCD2, FANCF, FANCG, FANCI, FANCL, FANCM, G6PC3, GATA1, GATA2, GFI1, HAX1, LAMTOR2, LYST, MPL, NHP2 , NOP10, PALB2, RAB27A, RAC2, RAD51C , RBM8A, RMRP, RPL5, RPL11, RPL15, RPL26, RPL35A, RPS7, RPS10, RPS17, RPS19, RPS24, RPS26, RTEL1, SBDS, SLC37A4, SLX4 (FANCP), SRP72, TAZ, TERC, TERT, TINF2, USB1, VPS13B, VPS45, WAS, WRAP53*
P2	Primary Immunodeficiency Panel (Invitae, San Francisco, California)	≤474	*ACD, ACP5, ACTB,ADA, ADA2, ADAM17, ADAR, AICDA, AIRE, AK2, AP3B1, ATM, B2M, BCL10, BLNK, BLOC1S6, BTK, CARD11, CARD14, CARD9, CASP10, CASP8, CD247, CD27, CD3D, CD3E, CD3G, CD40LG, CD79A, CD79B, CD8A, CEBPE, CHD7, CIITA, CLPB, COPA, CORO1A, CR2, CSF2RA, CSF3R, CTC1, CTLA4, CTPS1, CTSC, CXCR4, CYBA, CYBB, DCLRE1C, DKC1, DNMT3B, DOCK2, DOCK8, ELANE, EPG5,*

			FADD, FAS, FASLG, FERMT3, FOXN1, FOXP3, FPR1, G6PC3, GATA2, GFI1, HAX1, ICOS, IFIH1, IFNGR1, IFNGR2, IGLL1, IKBKB, IL10, IL10RA, IL10RB, IL12B, IL12RB1, IL17F, IL17RA, IL17RC, IL1RN, IL21, IL21R, IL2RA, IL2RG, IL36RN, IL7R, IRAK4, IRF7, IRF8, ISG15, ITCH, ITGB2, ITK, JAGN1, JAK3, LAMTOR2, LCK, LIG4, LPIN2, LRBA, LYST, MAGT1, MALT1, MAP3K14, MEFV, MVK, MYD88, NBN, NCF2, NCF4, NFAT5, NFKB2, NFKBIA, NHEJ1, NHP2, NLRC4, NLRP12, NLRP3, NOD2, NOP10, ORAI1, PARN, PGM3, PIK3CD, PIK3R1, PLCG2, PMM2, PNP, POLE, PRF1, PRKCD, PRKDC, PSMB8, PSTPIP1, PTPRC, RAB27A, RAC2, RAG1, RAG2, RBCK1, RFX5, RFXANK, RFXAP, RHOH, RMRP, RNASEH2A, RNASEH2B, RNASEH2C, RORC, RTEL1, SAMHD1,SEMA3E, SH2D1A, SH3BP2,SLC29A3, SLC35C1, SLC37A4, SLC7A7, SMARCAL1, SP110, SPINK5, STAT1, STAT2, STAT3, STAT5B, STIM1, STK4, STX11, STXBP2, TAP1, TAP2, TAPBP, TAZ, TBX1, TCN2, TERC, TERT, TICAM1, TINF2, TLR3, TMC6, TMC8, TMEM173, TNFRSF13B, TNFRSF13C, TNFRSF1A, TNFRSF4, TNFSF12, TPP2, TRAF3, TRAF3IP2, TREX1, TRNT1, TTC7A, TYK2, UNC13D, UNC93B1, UNG, VPS13B, VPS45, WAS, WIPF1, XIAP, ZAP70, ZBTB24
P3	Primary Immunodeficiency Panel (Laboratory of Molecular Genetic Research of the Belarusian	290	*ACP5, ACTB, ADA, ADAR, AGA, AICDA, AIRE, AK2, ALG13, AP3B1, APOL1, ATM, B2M, BCL10, BLM, BLNK, BLOC1S6, BTK, CA2, CARD11, CARD14, CARD9, CASP10, CASP8, CCBE1, CD19, CD247, CD27, CD3D, CD3E, CD3G, CD4, CD40, CD40LG, CD79A, CD79B, CD8A, CEBPE,*

Research Center for Pediatric Oncology, Hematology, and Immunology, Minsk, Belarus)	*CECR1, CFHR1, CFHR3, CHD7, CIITA, CLCN7, CNBP, COLEC11, COPA, CORO1A, CREBBP, CSF2RA, CSF3R, CTLA4, CTPS1, CTSC, CXCR4, CYBA, CYBB, DCLRE1B, DCLRE1C, DKC1, DNMT3B, DOCK2, DOCK8, ELF4, EPG5, FADD, FAS, FASLG, FCGR1A, FCGR2A, FCGR2B, FCGR3A, FCGR3B, FCGRT, FERMT3, FOXN1, FOXP3, G6PC, G6PC3, G6PD, GATA1, GATA2, GFI1, GJC2, GTF2H5, HAX1, ICOS, IFIH1, IFNGR1, IFNGR2, IGLL1, IKBKB, IKBKG, IKZF1, IL10, IL10RA, IL10RB, IL12B, IL12RB1, IL12RB2, IL17F, IL17RA, IL17RC, IL18, IL1RN, IL21, IL21R, IL2RA, IL2RG, IL36RN, IL7R, INSR, IRAK4, IRF7, IRF8, ISG15, ITCH, ITGB2, ITK, JAGN1, JAK2, JAK3, KRAS, LAMTOR2, LCK, LIG1, LIG4, LPIN2, LRBA, LRRC8A, LYST, MAGT1, MAL, MALT1, MAN2B1, MANBA, MAP3K14, MBD5, MBL2, MC2R, MCM4, MEFV, MKL1, MLPH, MOGS, MPO, MS4A1, MSH6, MTHFD1, MVK, MYD88, MYO5A, NBN, NCF1, NCF2, NCF4, NCSTN, NDNL2, NFAT5, NFKB1, NFKB2, NFKBIA, NHEJ1, NHP2, NKX2-5, NLRC4, NLRP12, NLRP3, NOD2, NOP10, NRAS, ORAI1, OSTM1, PARN, PCCA, PCCB, PEPD, PIGA, PIK3CD, PIK3R1, PLCG2, PMM2, PMS2, PNP, POLE, PRF1, PRKDC, PRPS1, PSENEN, PSMB8, PSTPIP1, PTPN11, PTPRC, PTRF, RAB27A, RAC2, RAG1, RAG2, RASGRP2, RBCK1, RECQL4, RFX5, RFXANK, RFXAP, RHOH, RMRP, RNASEH2A, RNASEH2B, RNASEH2C, RNF168, RNF31, RORC, RPSA, RTEL1, SAMD9L, SAMHD1, SBDS, SEMA3E, SERAC1, SH2D1A, SH3BP2, SKIV2L, SLC11A1, SLC29A3, SLC35C1, SLC37A4, SLC39A4,*

			SLC46A1, SMARCAL1, SP110, SPINK5, STAT1, STAT2, STAT3, STAT4, STAT5B, STIM1, STK4, STX11, STXBP2, TAP1, TAP2, TAPBP, TAZ, TBK1, TBX1, TCF3, TCIRG1, TCN2, TERC, TERT, THBD, TICAM1, TINF2, TIRAP, TLR3, TMC6, TMC8, TMEM173, TNFRSF11A, TNFRSF13B, TNFRSF13C, TNFRSF1A, TNFRSF4, TNFRSF6B, TNFSF12, TNFSF4, TRAF3, TRAF3IP2, TREX1, TRNT1, TTC37, TTC7A, TYK2, UNC119, UNC13D, UNC93B1, UNG, USB1, VIPAS39, VPS13B, VPS33B, VPS45, WAS, WIPF1, WRAP53, XIAP, XRCC4, ZAP70, ZBTB24
P4	Targeted resequencing	128	*ADA, AICDA, AKT1, AKT3, APCS, BCL2L1, BCL6, BLNK, BTK, CASP8, CCL5, CCR6, CD19, CD27, CD274, CD28, CD40, CD40LG, CD79A, CD79B, CD80, CD81, CD86, CDX1, CECR1, CLEC16A, CORO1B, CR2, CTLA4, CXCL12, CXCR4, CXCR5, DCLRE1C, DUSP2, FCGRT, FOXP3, GATA2, GRAP, GRB2, HDAC4, HSPA5, ICOS, ICOSLG, IGHM, IGLL1, IKBKB, IKBKE, IKBKG, IKZF1, IL21, IL21R, IL4I1, IRF2BP2, IRF4, KCNC4, KCNN4, KIDINS220, LRBA, LRRC32, LRRK2, MLH1, MS4A1, MSH2, MSH5, MTA3, NBN, NFKB1, NFKB2, NFKBIA, NLRP12, NOTCH1, NOTCH2, P2RX7, PDCD1, PDCD1LG2, PIK3AP1, PIK3CD, PIK3R1, PRDM1, PRKCD, PTEN, PTPN1, PTPN6, RAD50, RAG1, RAG2, RELA, RORA, RPS6KB2, RPTOR, SEC61A1, SEC61A2, SEC61G, SH2D1A, SH3KBP1, SOCS1, STAT1, STAT3, STK11, TCF3, TFRC, TGFB1, TGFB2, TGFB3, TMEM173, TNFAIP1, TNFRSF10A, TNFRSF13B, TNFRSF13C, TNFRSF17, TNFRSF18,*

	TNFRSF4, TNFSF10, TNFSF13, TNFSF13B, USP8, VAV1, VAV2, WNT5A, XCL1, XIAP, +7 undisclosed candidate genes

B

Oligonucleotide name		**DNA sequence (5'-3')**
R334X	F	GAT CCT CTC CAA AGG AAA GTG AGG TGG ACA TTC ATC TG
R334X	R	CAG ATG AAT GTC CAC CTC ACT TTC CTT TGG AGA GGA TC
E343K	F	GGA CAT TCA TCT GTT TCC ACT AAG TCT GAG TCT TCA AG
E343K	R	CTT GAA GAC TCA GAC TTA GTG GAA ACA GAT GAA TGT CC

Supplementary Table 2. Summary of immunophenotyping data for the 4 patients harboring *CXCR4*^{D84H}. (A) Peripheral blood lymphocyte subsets by flow cytometry in P1, P2, P3, and P4. (B) Leukocyte subset analysis of the mother of P2.

A

	P1	P2	P3	P4	Reference ranges		
Age, y	40	12	3, died	67	**3-7**	**12-18**	**≥18**
Lymphocyte subsets							
CD3$^+$ T, cells/µL (%Lym)	1378 **(88.6)**	1660 (68.6)	**295** (67.2)	1815 (75.6)	1400-2700 (60-75)	900-1800 (65-75)	700-2100 (55-83)
CD4$^+$ T, cells/µL (%Lym)	664 (42.7)	685 **(28.3)**	**14.80**	932 (38.9)	600-1700 (30-55)	550-1050 (30-47)	300-1400 (28-57)
CD4$^+$CD45RA naive T, cells/µL (%CD4$^+$)	741 (53.8)	**21.9** (of CD3)	**5.3** (of CD3)	26.6	400-1100 (30-60)	340-680 (28-50)	NA (13-56)
CD4$^+$CD45RO$^+$ memory T, cells/µL (%CD4$^+$)	311 **(22.6)**	(19.2 of CD3)	**(20.7 of CD3)**	71.6	100-530 (10-16)	150-370 (15-25)	NA (35-82)
CD8$^+$ T, cells/µL (%Lym)	757 **(48.7)**	**837 (34.6)**	**227 (51.7)**	699 (29.1)	500-1100 (15-30)	220-760 (17-30)	200-900 (10-39)
CD8$^+$ CD45RA$^+$ naive T, cells/µL (%CD8$^+$)	•	**(35.3 of CD3)**	**(52.5 of CD3)**	31.73	260-850 (15-28)	150-520 (17-29)	NA (7-58)
CD8$^+$ CD45RO$^+$ memory T, cells/µL (%CD8$^+$)	•	**(13.8 of CD3)**	**(21.5 of CD3)**	23.72	40-180 (3-10)	40-250 (4-12)	NA (26-58)
CD3$^-$ CD56$^+$ NK, cells/µL (%Lym)	**53 (3.4)**	140 (5.8)	**62** (14.2)	179 (7.5)	170-650 (5-20)	100-550 (5-25)	90-600 (7-31)
CD19$^+$ B, cells/µL (%Lym)	**104 (6.7)**	535 (22.1)	74 (17)	**14.6**	250-1100 (15-29)	170-500 (10-20)	100–500 (6-19)
CD19$^+$ CD27$^+$ IgD$^{-/+}$IgM$^+$ unswitched memory B cells (%CD19$^+$)	•	**(4.3)**	(9.8)	19.92	25-140 (4-12)	13-50 (5-29)	NA (3.1-20.8)
CD19$^+$ CD27$^+$ IgD$^-$ IgM$^-$ switched memory B cells (%CD19$^+$)	•	**(3.7)**	**(21.7)**	11.37	30-160 (4-15)	15-80 (6-20)	NA (5.1-31.8)
CD19$^+$ CD27$^-$ IgD$^+$ IgM$^+$ naive B cells (%CD19$^+$)	•	**(85.5)**	(59.1)	3.01	170-800 (60-85)	110-390 (50-80)	NA (36.4-84)

$CD19^+$ $IgM^{+\backslash-}$ $CD38^{++}$ Plasmablasts (%$CD19^+$)	•	•	•	**0.05**	(0.2-28)

CD, cluster of differentiation; Ig, immunoglobulin; NA, not available; NK, natural killer.

• = no data.

Values not within the normal range are shown in bold font.

[a]Reference ranges are based on values provided by Mayo Clinic Laboratories, Schatorjé EJH, et al[24] and Schatorjé EJH et al.[25]

B

Leukocyte subset analysis	P2 mother	Reference range
$CD3^+$ T cell, cells/µL (%Lym)	1333 (65.9)	700-2100 (55-83)
$CD4^+$ T cell, cells/µL (%Lym)	959 (47.4)	300-1400 (28-57)
$CD4^+$ $CD45RA^+$ naive T cell (%CD3)	(21.9)	(13-56)
$CD4^+$ $CD45RO^+$ memory T cell (%CD3)	(48.5)	(35-82)
$CD8^+$ T cell, cells/µL (%Lym)	376 (18.6)	200-900 (10-39)
$CD8^+$ $CD45RA^+$ naive T cell (%CD3)	(14.7)	(7-58)
$CD8^+$ $CD45RO^+$ memory T cell (%CD3)	(13.4)	(26-58)
$CD3^-CD56^+$ NK cell, cells/µL (%Lym)	362 (17.9)	90-600 (7-31)
$CD19^+$ B cell, cells/µL (%Lym)	271 (13.4)	100-500 (6-19)
$CD19^+$ $CD27^+$ memory B cell, cells/µL (%Lym)	(4.9)	(3.1-20.8)
$CD19^+$ $CD27^+$ $CD38^+$ IgD^+ unswitched memory B cell (%$CD19^+$)	(2.3)	(5.1-31.8)
$CD19^+$ $CD27^+$ $CD38^+$ IgD^- switched memory B cell (%$CD19^+$)	(2.6)	(4.4-26.2)
$CD19^+$ $CD27^-$ B cell (%$CD19^+$)	(95.1)	(36.4-84)
$CD19^+$ $CD27^-$ IgD^+ IgM^+ transitional/naive B cell (%$CD19^+$)	(87.3)	NA
$CD19^+$ $CD27^-$ IgD^- IgM^+ immature B cell (%$CD19^+$)	(7.8)	NA

CD, cluster of differentiation; Ig, immunoglobulin; NA, not available; NK, natural killer.

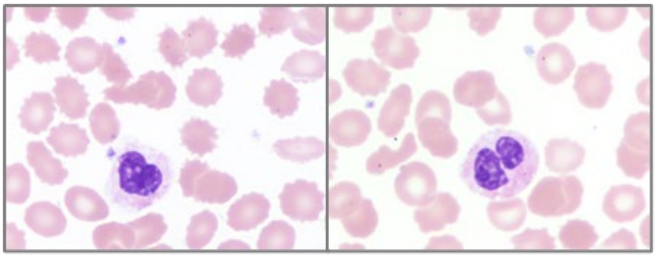
Supplementary Table 3. Allele frequency for c.250G>C variant on *CXCR4* transcript. Allele frequency for c.250G>C variant on *CXCR4* transcript ENST00000241393.4 that leads to amino acid change p.D84H, in 3 genomic databases. ESP, Exome Sequencing Project; gnomAD, Genome Aggregation Database; NHLBI, National Heart, Lung, and Blood Institute; TOPMed, Trans-Omics for Precision Medicine.

Source	Allele frequency
gnomAD (v2.1)	8.0E-06
NHLBI's ESP	7.7E-05
TOPMed (Freeze 8)	2.3E-05
Merged estimate	**3.59E-05**

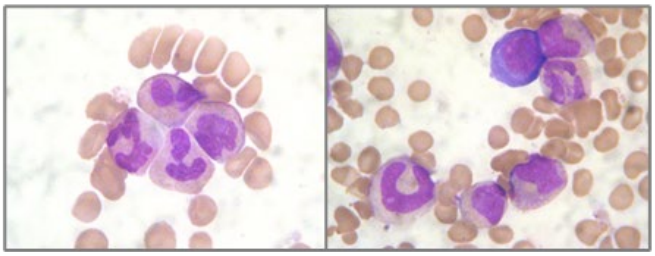
Supplementary Figures

Supplementary Figure 1. Morphology of D84H granulocytes and location of CXCR4 variants. (A) PB smears from P1 showing dysmorphic neutrophils with frequent nuclear bilobation and monolobation (left), rather than elongated thin internuclear chromatin strands typical of WHIM syndrome (original magnification ×100). Myelokathectic neutrophils were observed in PB smears and accounted for 2% of total neutrophils. (B) BM aspirates from P2 showing granulocytes with irregular nuclear contours, an abnormal morphology but not typical of myelokathexis morphology (original magnification ×100). Myelokathectic neutrophils accounted for 3% and 1% of total neutrophils observed in the BM aspirated from P2 and P3 (not shown), respectively. Images were obtained on an Olympus BX41 microscope equipped with Olympus UPlanFL N optic at the 100× oil magnification and 1.30 numerical aperture and were digitalized using Olympus DP23 camera and Olympus cellSens Entry software. (C) Position of variant p.D84H (red circle) in the TM7 model of CXCR4 (from GPCRdb). The locations of most frequent variant (p.R334X) and the only known missense variant (p.E343K) reported in WHIM syndrome are also highlighted (blue circles). Figure adapted from gpcrdb.org. (D) The amino acid residue D84 is found in a highly evolutionary conserved region in the second TM helix of the CXCR4 receptor. Figure adapted from genecards.org. BM, bone marrow; CXCR4, C-X-C chemokine receptor 4; GPCR, G protein–coupled receptor; PB, peripheral blood; WHIM, Warts, Hypogammaglobulinemia, Infections, and Myelokathexis.

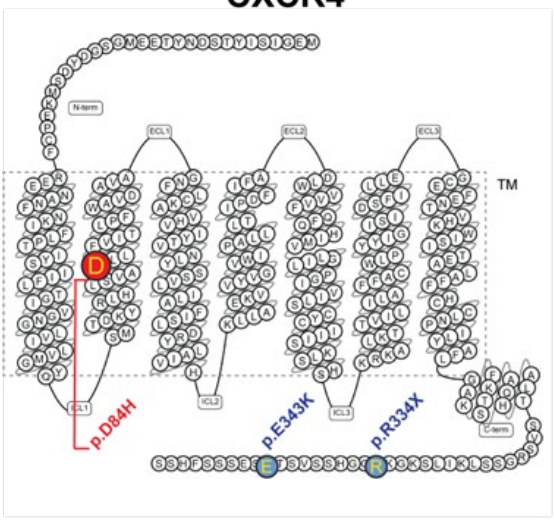
A



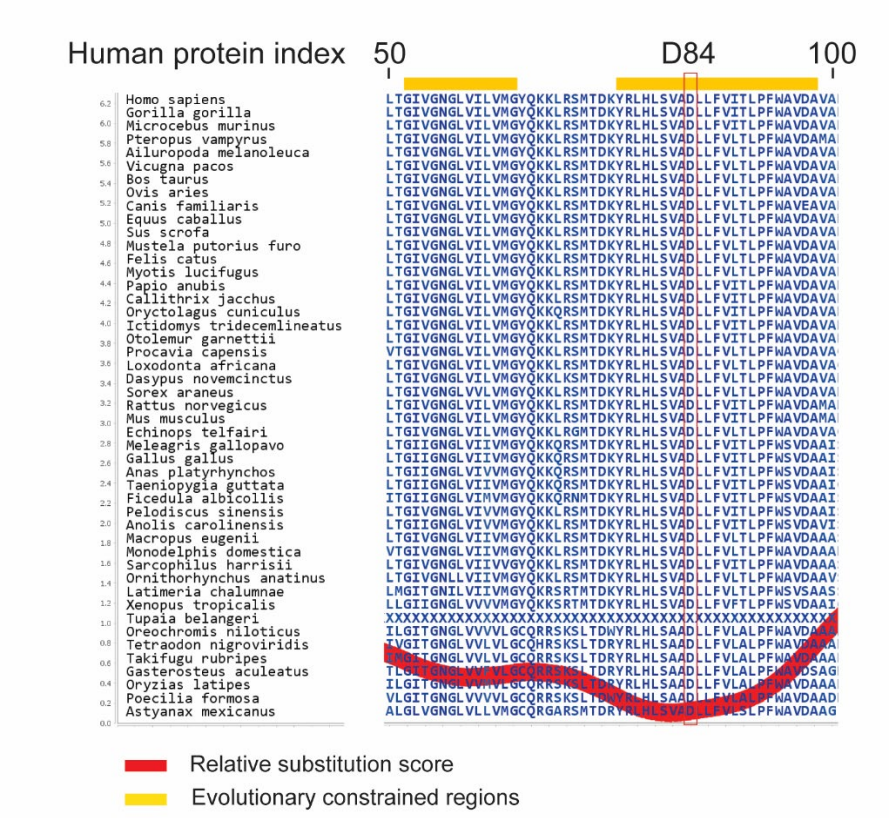
B



C

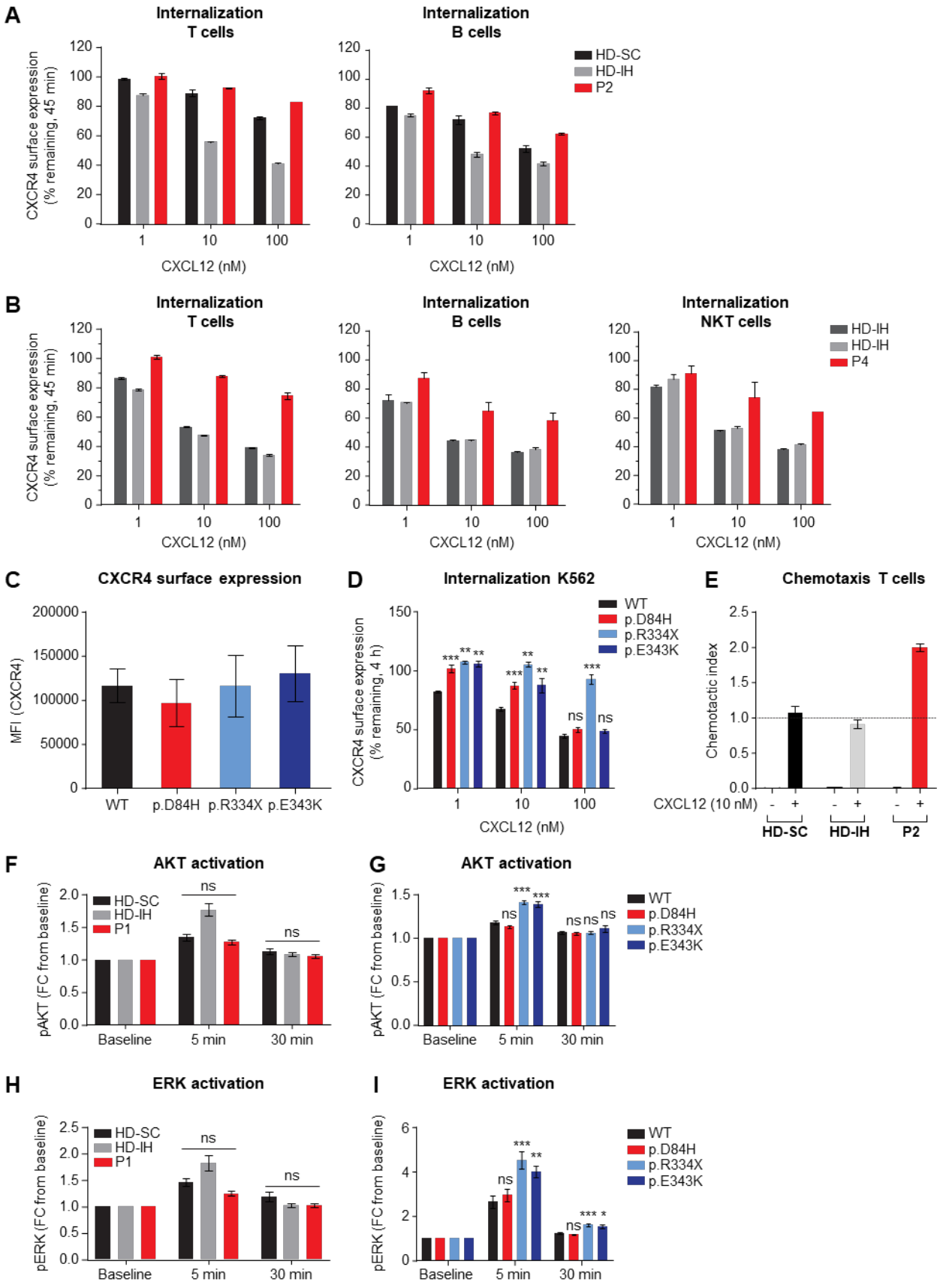

CXCR4
TM
p.D84H
p.E343K
p.R334X

D

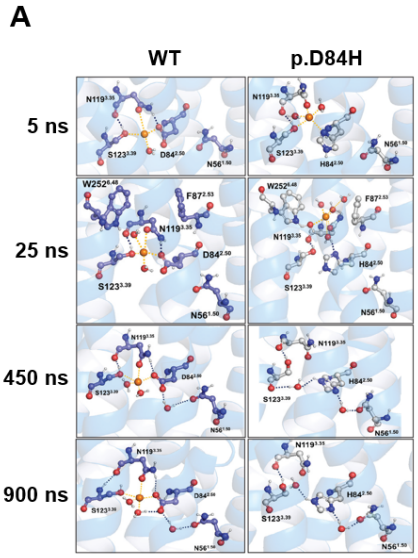

Human protein index
50
D84
100
Homo sapiens
Gorilla gorilla
Microcebus murinus
Pteropus vampyrus
Ailuropoda melanoleuca
Vicugna pacos
Bos taurus
Ovis aries
Canis familiaris
Equus caballus
Sus scrofa
Mustela putorius furo
Felis catus
Myotis lucifugus
Papio anubis
Callithrix jacchus
Oryctolagus cuniculus
Ictidomys tridecemlineatus
Otolemur garnettii
Procavia capensis
Loxodonta africana
Dasypus novemcinctus
Sorex araneus
Rattus norvegicus
Mus musculus
Echinops telfairi
Meleagris gallopavo
Gallus gallus
Anas platyrhynchos
Taeniopygia guttata
Ficedula albicollis
Pelodiscus sinensis
Anolis carolinensis
Macropus eugenii
Monodelphis domestica
Sarcophilus harrisii
Ornithorhynchus anatinus
Latimeria chalumnae
Xenopus tropicalis
Tupaia belangeri
Oreochromis niloticus
Tetraodon nigroviridis
Takifugu rubripes
Gasterosteus aculeatus
Oryzias latipes
Poecilia formosa
Astyanax mexicanus
Relative substitution score
Evolutionary constrained regions

Supplementary Figure 2. Effects of CXCR4^{D84H} in *in vitro* functional assays. (A, B) PBMCs isolated from HDs (HD-SC, HD-IH), P2, and P4 were stimulated with CXCL12 (vehicle, 1 nM, 10 nM, 100 nM) for 45 minutes. CXCR4 receptor internalization was assessed by surface expression of CXCR4 via flow cytometry. Values are expressed as % remaining CXCR4 compared with vehicle-treated cells. PBMCs were subtyped using fluorescent mAbs specific for CD3 (T cells), CD19 (B cells), and CD3 and CD56 (NKT cells). Statistical analysis was not performed owing to low sample numbers. Mean +/– SD of n=2. (C) K562 cells were transiently transfected with indicated CXCR4 constructs, and CXCR4 receptor surface expression was measured via flow cytometry. Values are expressed as MFI of CXCR4 staining. Values represent mean +/– SEM of n=4-7. (D) K562 cells were transiently transfected with indicated CXCR4 constructs and stimulated with CXCL12 (vehicle, 1 nM, 10 nM, 100 nM) for 4 hours. CXCR4 receptor internalization was assessed by surface expression of CXCR4 via flow cytometry. Statistical significance was determined by Mann-Whitney tests comparing variants with the WT at each respective concentration of ligand. Values are expressed as % remaining CXCR4 compared with vehicle-treated cells. Values represent mean +/– SEM of n = 3-17. **P*<.05, ***P*<.01, ****P*<.001. (E) PBMCs isolated from HDs (HD-SC, HD-IH) and P2 were subjected to Transwell chemotactic assay. Cells migrated toward 10 nM CXCL12 or medium only for 2.5 hours. The total number of migrated cells was normalized to the HD-SC. PBMCs were subtyped using fluorescent mAbs specific for CD3 (T cells). Statistical analysis was not performed owing to low sample numbers. Mean +/– SD of n=2. (F) PBMCs isolated from HDs (HD-SC, HD-IH) and P1 were stimulated with 10 nM CXCL12 for 5 and 30 min; fixed and the median fluorescence intensity of p-S473 AKT staining were measured by flow cytometry. Values are expressed as FC compared with unstimulated sample. Mean +/– SEM of n=5. Statistical analysis by Mann-Whitney tests compared the P1 response with the respective time point of HD-SC PBMCs. (G) K562 cells with stable CXCR4 expression were stimulated with 10 nM CXCL12 for 5 and 30 min; fixed and the median fluorescence intensity of p-S473 AKT staining were measured by flow cytometry. Statistical significance determined by Mann-Whitney tests compared variants with the WT at each respective timepoint. Values are expressed as FC compared with unstimulated sample. Mean +/– SEM of n=17-18. **P*<.05, ***P*<.01, ****P*<.001. (H) PBMCs isolated from HDs (HD-SC, HD-IH) and P1 were stimulated with 10 nM CXCL12 for 5 and 30 min; fixed and the median fluorescence intensity of p-T202/Y204 ERK staining was measured by flow cytometry. Values are expressed as FC compared with unstimulated sample. Mean +/– SEM of n=5. Statistical analysis by Mann-Whitney tests compared the P1 response with the respective time point of HD-SC PBMCs. (I) K562 cells with stable CXCR4 expression

were stimulated with 10 nM CXCL12 for 5 and 30 min; fixed and the median fluorescence intensity of p-T202/Y204 ERK staining were measured by flow cytometry. Statistical significance was determined by Mann-Whitney tests comparing variants with the WT at each respective timepoint. Values are expressed as FC compared with unstimulated sample. Mean +/– SEM of n=18. *P<.05, **P<.01, ***P<.001. AKT, protein kinase B; CD, cluster of differentiation; CXCL12, C-X-C chemokine ligand 12; CXCR4, C-X-C chemokine receptor 4; ERK, extracellular signal-regulated kinase; FC, fold change; HD, healthy donor; mAb, monoclonal antibody; MFI, mean fluorescence intensity; NK, natural killer; ns, not significant; PBMC, peripheral blood mononuclear cell; SEM, standard error of mean; WT, wild-type.


A
Internalization
T cells
Internalization
B cells
CXCR4 surface expression
(% remaining, 45 min)
CXCL12 (nM)
HD-SC
HD-IH
P2
B
Internalization
T cells
Internalization
B cells
Internalization
NKT cells
HD-IH
HD-IH
P4
C
CXCR4 surface expression
MFI (CXCR4)
WT
p.D84H
p.R334X
p.E343K
D
Internalization K562
CXCR4 surface expression
(% remaining, 4 h)
WT
p.D84H
p.R334X
p.E343K
E
Chemotaxis T cells
Chemotactic index
CXCL12 (10 nM)
HD-SC
HD-IH
P2
F
AKT activation
pAKT (FC from baseline)
HD-SC
HD-IH
P1
Baseline
5 min
30 min
G
AKT activation
WT
p.D84H
p.R334X
p.E343K
H
ERK activation
pERK (FC from baseline)
I
ERK activation

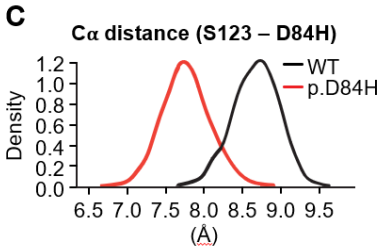
Supplementary Figure 3. Impact of CXCR4^{D84H} on receptor structure in molecular dynamics simulation. (A) Snapshots of allosteric binding pocket taken at indicated time points of MD simulations. Orange, sodium cation; red, oxygen atom; dark blue, nitrogen atom. (B) Average RRCSs measured between indicated residue pairs during the time course of MD simulations. RRCSs from CXCR4 crystal structures are included for reference. (C) Distributions of D84H$^{2\times50}$-S123$^{3\times39}$ Cα distance measured throughout the MD simulations. (D) K562 cells with stable CXCR4 expression were stimulated with forskolin +/– 100 nM CXCL12 for 30 minutes. cAMP production was measured by ELISA. Values represent mean +/– SEM of n=8-9. Statistical significance was determined by Mann-Whitney tests comparing the indicated samples. *P<.05. cAMP, cyclic adenosine monophosphate; ELISA, enzyme-linked immunosorbent assay; MD, molecular dynamic; ns, not significant; RRCS, residue-residue contact score; SEM, standard error of mean.

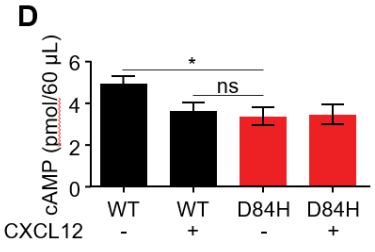


B

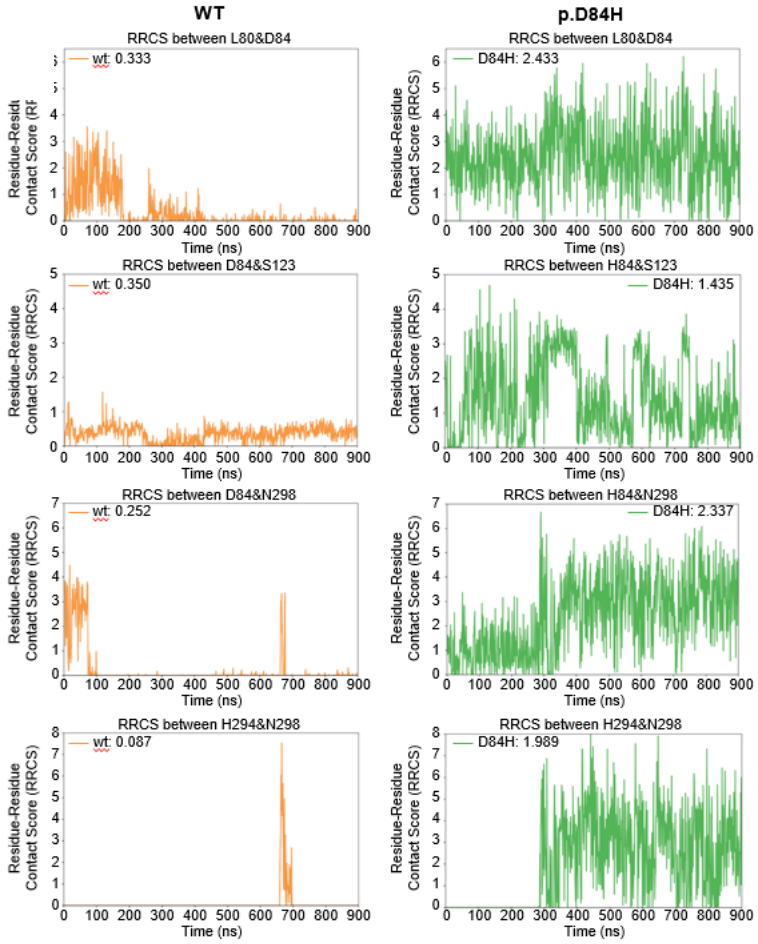
RRCS measurements

Ballesteros-Weinstein numbering	CXCR4	3ODU	3OE6	Average value	
				WT	p.D84H
2×46 : 2×50	L80 : D84	0.884	0.431	0.333	2.433
2×50 : 3×39	D84 : S123	0	0.265	0.350	1.435
2×50 : 7×49	D84 : N298	0	0	0.252	2.337
7×45 : 7×49	H294 : N298	0	0	0.087	1.989





Supplementary Figure 4. RRCS between indicated residue pairs in CXCR4^{WT} and CXCR4^{D84H} during MD simulations. RRCSs were measured between residue pairs during the time course of MD simulations. The average values are indicated in the top right corner. MD, molecular dynamic; RRCS, residue–residue contact score.

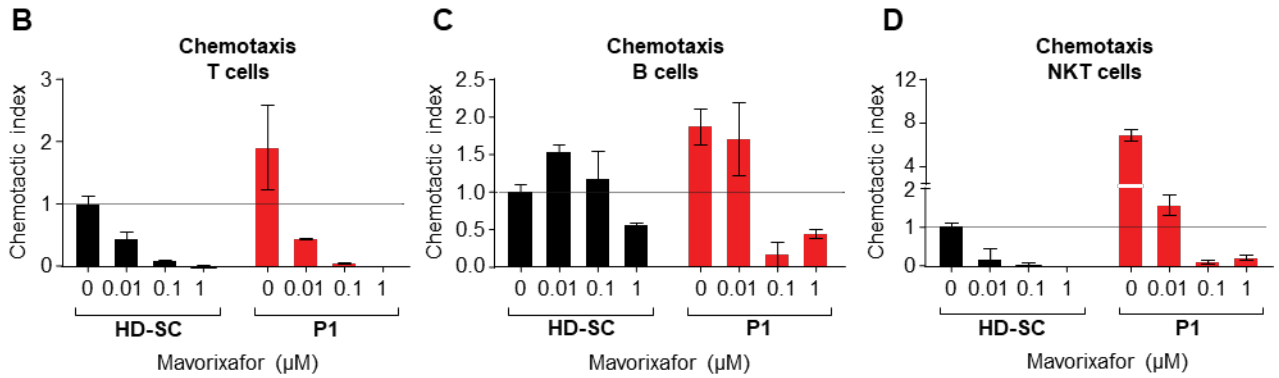


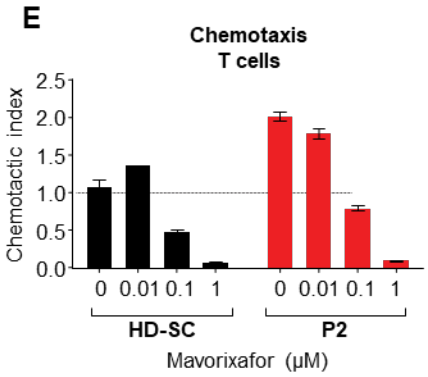
Supplementary Figure 5. Effect of mavorixafor on PBMCs isolated from P1, P2, and transfected K562 cells. (A) IC_{50} of mavorixafor determined in ligand binding inhibition assays in K562 cells with stable CXCR4 expression. (B-D) PBMCs isolated from P1 were preincubated with vehicle, 0.01, 0.1, or 1 µM mavorixafor for 30 minutes and subjected to migration toward 10 nM CXCL12. The total number of migrated cells was normalized with respect to the vehicle-pretreated HD-SC. PBMCs were subtyped using fluorescent mAbs specific for CD3 (T cells), CD19 (B cell), and CD3 and CD56 (NKT cells). Statistical analysis was not performed owing to low sample numbers. Mean +/– SD, n=2. (E) PBMCs collected from P2 were preincubated with vehicle, 0.01, 0.1, or 1 µM mavorixafor for 30 minutes and subjected to migration toward 10 nM CXCL12. The total number of migrated cells was normalized with respect to the vehicle-pretreated HD-SC. PBMCs were subtyped using fluorescent mAbs specific for CD3 (T cells). Statistical analysis was not performed owing to low sample numbers. Mean +/– SD of n = 2. CD, cluster of differentiation; CXCL12, C-X-C chemokine ligand 12; CXCR4, C-X-C chemokine receptor 4; HD, healthy donor; IC_{50}, half maximal inhibitory concentration; mAb, monoclonal antibody; NK, natural killer; PBMC, peripheral blood mononuclear cell; WT, wild-type.

A

Mavorixafor IC_{50}

CXCR4 variant	CXCL12 binding inhibition IC_{50} (nM)
WT	4.9
D84H	3.9
R334X	4.5
E343K	5.6





Supplementary references

1. Katritch V, Fenalti G, Abola EE, Roth BL, Cherezov V, Stevens RC. Allosteric sodium in class A GPCR signaling. *Trends Biochem Sci*. 2014;39(5):233-244.
2. Zhou Q, Yang D, Wu M, et al. Common activation mechanism of class A GPCRs. *eLife*. 2019;8:e50279.
3. Biebermann H, Krude H, Elsner A, Chubanov V, Gudermann T, Grüters A. Autosomal-dominant mode of inheritance of a melanocortin-4 receptor mutation in a patient with severe early-onset obesity is due to a dominant-negative effect caused by receptor dimerization. *Diabetes*. 2003;52(12):2984-2988.
4. Sadeghi H, Robertson GL, Bichet DG, Innamorati G, Birnbaumer M. Biochemical basis of partial nephrogenic diabetes insipidus phenotypes. *Mol Endocrinol*. 1997;11(12):1806-1813.
5. Ballesteros JA, Weinstein H. [19] Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. In: Sealfon SC, ed. *Methods in Neurosciences*. Vol 25. Academic Press; 1995:366-428.
6. Isberg V, de Graaf C, Bortolato A, et al. Generic GPCR residue numbers - aligning topology maps while minding the gaps. *Trends Pharmacol Sci*. 2015;36(1):22-31.
7. Cong X, Golebiowski J. Allosteric Na(+)-binding site modulates CXCR4 activation. *Phys Chem Chem Phys*. 2018;20(38):24915-24920.
8. Zhang WB, Navenot JM, Haribabu B, et al. A point mutation that confers constitutive activity to CXCR4 reveals that T140 is an inverse agonist and that AMD3100 and ALX40-4C are weak partial agonists. *J Biol Chem*. 2002;277(27):24515-24521.

9. Qin L, Kufareva I, Holden LG, et al. Structural biology. Crystal structure of the chemokine receptor CXCR4 in complex with a viral chemokine. *Science*. 2015;347(6226):1117-1122.
10. Wu B, Chien EY, Mol CD, et al. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science*. 2010;330(6007):1066-1071.
11. Zmajkovicova K, Pawar S, Maier-Munsa S, et al. Genotype-phenotype correlations in WHIM syndrome: a systematic characterization of CXCR4(WHIM) variants. *Genes Immun*. 2022;23(6):196-204.
12. Schoofs G, Van Hout A, D'huys T, Schols D, Van Loy T. A flow cytometry-based assay to identify compounds that disrupt binding of fluorescently-labeled CXC chemokine ligand 12 to CXC chemokine receptor 4. *J Vis Exp*. 2018;(133):e57271.
13. Schott-Verdugo S, Gohlke H. PACKMOL-Memgen: a simple-to-use, generalized workflow for membrane-protein-lipid-bilayer system building. *J Chem Inf Model*. 2019;59(6):2522-2528.
14. Abraham MJ, Murtola T, Schulz R, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 2015;1-2:19-25.
15. Tian C, Kasavajhala K, Belfon KAA, et al. ff19SB: amino-acid-specific protein backbone parameters trained against quantum mechanics energy surfaces in solution. *J Chem Theory Comput*. 2020;16(1):528-552.
16. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*. 1983;79(2):926-935.
17. Joung IS, Cheatham TE, 3rd. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J Phys Chem B*. 2008;112(30):9020-9041.

18. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. *J Chem Phys*. 1995;103(19):8577-8593.

19. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: A linear constraint solver for molecular simulations. *J Comput Chem*. 1997;18(12):1463-1472.

20. gnomAD v2.1. gnomAD Browser. October 17, 2018. Accessed December 29, 2022. https://gnomad.broadinstitute.org/news/category/releases/.

21. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-443.

22. Taliun D, Harris DN, Kessler MD, et al. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *Nature*. Feb 2021;590(7845):290-299.

23. Exome variant server. LHLBI Exome Sequencing Project. Accessed December 29, 2022. https://evs.gs.washington.edu/EVS/.

24. Schatorjé EJ, Gemen EF, Driessen GJ, et al. Age-matched reference values for B-lymphocyte subpopulations and CVID classifications in children. *Scand J Immunol*. 2011;74(5):502-510.

25. Schatorjé EJ, Gemen EF, Driessen GJ, Leuvenink J, van Hout RW, de Vries E. Paediatric reference values for the peripheral T cell compartment. *Scand J Immunol*. 2012;75(4):436-444.