Supplementary Table 2. Relative binding energies (in kcal/mol) for G protein binding to US28 Toledo variants bound different chemokines and gp120.

| Variant | CX3CL1 ΔΔG | CCL2 ΔΔG | CCL3 ΔΔG | CCL4 ΔΔG | CCL5 ΔΔG | CCL13 ΔΔG | gp120 Indonesian ΔΔG | gp120 Australian ΔΔG |
|---------|------------|----------|----------|----------|----------|-----------|----------------------|----------------------|
| D15E    | 0          | 0        | 0        | 0        | 0        | 0         | 0                    | 0                    |
| E18L    | 0          | 0        | 0        | 0        | 0        | 0         | 0                    | 0                    |
| A19D    | 0          | 0        | 0        | 0        | 0        | 0         | 0                    | 0                    |
| T21A    | 0          | 0        | 0        | 0        | 0        | 0         | 0                    | 0                    |
| F25L    | 0          | 0        | 0        | 0        | 0        | 0         | -0.01                | 0                    |
| Y40N    | -0.07      | -0.19    | -0.04    | -0.16    | -0.08    | -0.14     | -0.03                | 0                    |
| G50C    | -0.13      | -0.14    | -0.02    | +7.98    | +1.05    | +0.03     | -0.16                | +0.23                |
| N170D   | 0          | 0        | 0        | 0        | 0        | 0         | 0                    | 0                    |
| R267K   | 0          | +0.01    | 0        | 0        | 0        | -0.01     | 0                    | 0                    |
| US28 Toledo 4XT1_A | 1   | MTPTTTTAELTEFDYDEAATPCVFTDVNLQSKPVTFLFGSIG |
|---------------------|-----|---------------------------------------------|
|                     | 15  | ~~~~~~~~~~~~~~~~~~~~DYDEDATPCVFTDVNLQSKPVTFLFGSIG |
|                     |     | ****  ****************************************** |
| US28 Toledo 4XT1_A | 51  | NFLVIFTITWRRRIQCSGDVVFYINLAAADLLFVCCTPLWMQYLLDHNSLA |
|                     | 51  | NFLVIFTITWRRRIQCSGDVVFYINLAAADLLFVCCTPLWMQYLLD~~~ |
|                     |     | ********************************************* |
| US28 Toledo 4XT1_A | 101 | SVPCTLLTACFYVAMFASLCFITEIALDRYYAIVYMRYRPVKQACLFSIF |
|                     | 101 | SVPCTLLTACFYVAMFASLCFITEIALDRYYAIVYMRYRPVKQACLFSIF |
|                     |     | ********************************************* |
| US28 Toledo 4XT1_A | 151 | WWIFAVIIAIPHFMVVTKKNQCMTDYDYLEVSYPILIINVELMLGAFLVIP |
|                     | 151 | WWIFAVIIAIPHFMVVTKKNQCMTDYDYLEVSYPILIINVELMLGAFLVIP |
|                     |     | ********************************************* |
| US28 Toledo 4XT1_A | 201 | LSVISYCYYRISRIVAQSRSRHGRIVRVLIAVVLFIIFWLPYHLLTLFV |
|                     | 201 | LSVISYCYYRISRIVAQSRSRHGRIVRVLIAVVLFIIFWLPYHLLTLFV |
|                     |     | ********************************************* |
| US28 Toledo 4XT1_A | 251 | DTLKLKLISSCEFERSLKRALILTSEAFCHCCLNHLYFVGTFKFRQ |
|                     | 251 | DTLKLKLISSCEFERSLKRALILTSEAFCHCCLNHLYFVGTFKFRQ |
|                     |     | ********************************************* |
| US28 Toledo 4XT1_A | 301 | ELHCLLAEFQRFLFSDVSWYHSMFSRSSRSPRRETSSDTLSDEVCRVS |
|                     | 301 | ELHCLLAEFR~~~~~~~~~~~~~~~~~~~~~~~~~ *** |
|                     |     | ********************************************* |
| US28 Toledo 4XT1_A | 361 | QIIP |
|                     |     | ~~~~ |
## Supplementary Table 4. Sequence alignment for CCL2.

| CCL2   | 1DOK_A | 4XT1_B |
|--------|--------|--------|
| 1      | MKVSAALLCLLLLIAATFIPQGLAQPDAINAPVTCCYNTKISVQRLAS | ~~~~~~~~MKVSAALLCLLLLIAATFIPQGLAQPDAINAPVTCCYNTKISVQRLAS |
| 23     | ~~~~~~~~MQPDAINAPVTCCYNTKISVQRLAS | XHHGTVKCAITCSKMTS~KIPVALLIH |
| 25     | ~~~~~~~~XHHGTVKCAITCSKMTS~KIPVALLIH | **: ****: *:*:*:*:****** |

*The regions highlighting are the regions of the respective templates used to model the bound chemokine. Sequence identity/similarity is indicated between the chemokine and the region highlighted. X in the 4XT1_B sequence is pyroglutamic acid.*

## Supplementary Table 5. Sequence alignment for CCL3.

| CCL3   | 3FPU_B | 4XT1_B |
|--------|--------|--------|
| 1      | MQVSTAALAVLLCTMALCNQFSASLAADTPAC~~~CFSYTSRQIPQNF | ~~~~~~~~MQVSTAALAVLLCTMALCNQFSASLAADTPAC~~~CFSYTSRQIPQNF |
| 24     | ~~~~~~~~CFSYTSRQIPQNF | ~~~~~~~~XHHGTVKCAITCSKMTS~KIPVALLIH |
| 25     | ~~~~~~~~XHHGTVKCAITCSKMTS~KIPVALLIH | **: ****: *:*:*:*:****** |

*The regions highlighting are the regions of the respective templates used to model the bound chemokine. Sequence identity/similarity is indicated between the chemokine and the region highlighted. X in the 4XT1_B sequence is pyroglutamic acid.*

## Supplementary Table 6. Sequence alignment for CCL4.

| CCL4   | 3TN2_A | 4XT1_B |
|--------|--------|--------|
| 1      | MKLCVTVLSLMLVAAFCSPALSMGPMSDPPTACCFSYTAKLPRNFVVD | ~~~~~~~~APMGSDPATACCFSYTAKLPRNFVVD |
| 24     | ~~~~~~~~APMGSDPATACCFSYTAKLPRNFVVD | XHHGTVKCAITCSKMTS~KIFVALLIH |
| 25     | ~~~~~~~~XHHGTVKCAITCSKMTS~KIFVALLIH | **: ****: *:*:*:*:****** |

*The regions highlighting are the regions of the respective templates used to model the bound chemokine. Sequence identity/similarity is indicated between the chemokine and the region highlighted. X in the 4XT1_B sequence is pyroglutamic acid.*
### Supplementary Table 7. Sequence alignment for CCL5.\(^a\)

|        | 1 | MKVSAAALAVILIATALCAPASASPYSSDTTPCCFAYIARPLPRAHIKEY |
|--------|---|------------------------------------------------------|
| 5COY_A | 27 | ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ |
| 4XT1_B | 25 | ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~XHHGVTCAITCSKMTSK~1PVALLIHY |

\(^a\)The regions highlighting are the regions of the respective templates used to model the bound chemokine. Sequence identity/similarity is indicated between the chemokine and the region highlighted. X in the 4XT1_B sequence is pyroglutamic acid.

### Supplementary Table 8. Sequence alignment for CCL13.\(^a\)

|        | 1 | MKVSAVLLCLLLMTAAFNPGQAQPDALNVSTCCFTSSKISLQRLKS |
|--------|---|---------------------------------------------------|
| 2RA4_B | 26 | ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ |
| 4XT1_B | 25 | ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~XHHGVTCAITCSKMTSK~1PVALLIHY |

\(^a\)The regions highlighting are the regions of the respective templates used to model the bound chemokine. Sequence identity/similarity is indicated between the chemokine and the region highlighted. X in the 4XT1_B sequence is pyroglutamic acid.
### Supplementary Table 9. Sequence alignment for US28 Toledo modelled for gp120 binding.\textsuperscript{a}

| US28 Toledo | 1 | MTTTTTAELTTEFDYDEAATPCVTDVLNQSKPVTFLFLYGVFLFGSIG |
| 6MEO_B | 1 | MDYQVSSPIXDIX~~~YTSEPQCQINVKQIAARLLLPPLYSLVIFGFVGF |
| 4XT1_A | 15 | ~~~~DYDEATPCVTDVLNQSKPVTFLFLYGVFLFGSIG | * | :: | * | ** | * | ** | * | : | :: | ** | : | : | ** | : | : | ** | : | : | ** |
| US28 Toledo | 51 | NFLVIFITITWRRRISGQVDVVFINLAAADLFLVCTPLWMLQYLLDDHNSLA |
| 6MEO_B | 51 | NFLVILINCKRLKSMTDIYLLNLAISDLFFLLTVPFWAHYYAAQWDFG |
| 4XT1_A | 48 | *** | : | : | ** | ** | : | : | ** | : | : | ** | : | ** | ** | : | : | ** | : | ** | : | : | ** | : | : | ** |
| US28 Toledo | 101 | SVPCTLLACFYMAMFASLCFTFEIADRYYAIY~~~MRYRPVQAACL |
| 6MEO_B | 101 | NFLVIFTI |
| 4XT1_A | 101 | NFLVIFAIIAIPHMVVTKDND~~~QCMTDYDEYPSYIIILN~VE |
| US28 Toledo | 147 | FSIFWIFIAIIAIPFMVVTKKN~~~QCMTDYDYLEVPSYIIILN~VE |
| 6MEO_B | 148 | TSVTWVVAVFAVFSRSTSQKEGLHYTCSHPSYQQFWKPNFQTLK |
| 4XT1_A | 147 | FSIFWIFIAIIAIPHMVVTKDN~~~QCMTDYDEYPSYIIILN~VE |
| US28 Toledo | 192 | LMLGAFLPVLPLSVPISYCQYRISRIAVSQS~RHKGRIIVRVLAVVVLVFIIF |
| 6MEO_B | 192 | IVIILVLPLVMVICSGIKCLMLR<=RCKRHRKAVRLIFTMIVYLF |
| 4XT1_A | 192 | LMLGAFLPVLPLSVPISYCQYRISRIAVSQS~RHKGRIIVRVLAVVVLVFIIF |
| US28 Toledo | 241 | WLPYHTLTFDTVTLKLLLKWISSSCEFERSKRALILTESLAFCHCLNPLL |
| 6MEO_B | 248 | WAPYNIVLHTNFQEFGLNN~CSSNRLQAMQVTETLMTHCCINPIL |
| 4XT1_A | 241 | WLPYHTLTFDTVTLKLLLKWISSSCEFERSKRALILTESLAFCHCLNPLL |
| US28 Toledo | 291 | YVFVTGKFRQELHCLLAEFRLRSDRDVSYSFDSRFRSRRSRRSRSRRRTSSD |
| 6MEO_B | 297 | YAVFVGKFRQNLVVFQ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ |
| 4XT1_A | 291 | YVFVTGKFRQELHCLLAEPR|~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ |
| US28 Toledo | 341 | TLSDEVCRVSQIIP |
| 6MEO_B | ~~~~~~~~~~~~~~~~~~ |
| 4XT1_A | ~~~~~~~~~~~~~~~~~~ |

\textsuperscript{a}The regions highlighting are the regions of the respective templates used to model the gp120-bound conformation of US28 Toledo. Sequence identity/similarity is indicated between the receptor and the region highlighted. X in the 6MEO_B sequence indicates an unknown residue.
**Supplementary Table 10. Sequence alignment for gp120 Indonesian.**

| gp120_ID | 6MEO_G | Sequence_alignment |
|----------|--------|--------------------|
| 1        | MRVKKTQMNWLSWKGTLILGVMCANSNWLVTIVYGVPVWKEATT | 1-29 |
| 51       | LFCSADAKAHEVHNWATHACVPTQNLPLKNVTENFNWMKNNPM | 30-50 |
| 101      | EQMHEDVISLWQTHKLPCVKLPCVTLNCTNACLTVNDFSDTNEPT | 51-100 |
| 151      | ESTTRNNTETDEVKCNFTFVTTELTDMTKQVHALFYKLDIVQINDRSVN | 101-150 |
| 201      | NNSSSGKYMVCNTSVIKQACPQKSIDIPHIPHYCAPAYAILKCKDKKF | 151-200 |
| 251      | NGIGPCNNVSVQCTHGIRPVVSTQPLLNLGSLAEELIIRSINITNNAK | 201-250 |
| 301      | IIVHLNKSEISCHARYKNTRTSTHMVPGRTRFRTGIGDIKAHCEIN | 251-300 |
| 351      | GTTWINETLEQVKWLYQFPKPIIFQPQPGGDPEITMHHFNCQGEFF | 301-350 |
| 401      | YCNTTDLFNNNRTDG~~~VKLPCKIKFVPNKTIIFQPQPHAGDPEITMHHFNCQGEFF | 351-400 |
| 448      | CTSNITGIILTRDGATNTRNETSSNTEIFRPQGEDMDRDNWRELK | 401-440 |
| 498      | YKVVQIEPLGIAPTRAKR | 441-490 |

*Note: The sequences are aligned with a focus on highlighting conserved and variable regions.*
**Supplementary Table 11. Sequence alignment for gp120 Australian.**

| gp120 AU 6MEO_G | 1 | MRVKEKYQHLWRGWGRTMLLGMICSAEKLLWVTVYGVFPWKEATT |
| gp120 AU 6MEO_G | 29 | ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~DNLWTVTVYGVFPWKEATT |
| gp120 AU 6MEO_G | 51 | TLFCSADAKAYDTEVHNWATHACVPTDPNQFQEVVLNVNVTENFNWMKKNDM |
| gp120 AU 6MEO_G | 49 | TLFCSADAKAYDTEVHNWATHACVPTDPNQFQEVVLNVNVTENFNWMKKNDM |
| gp120 AU 6MEO_G | 101 | VEQMHEDIISLWDQSLKPCVKLPTLCVSLKCTDLKNDTNNSGSMIME |
| gp120 AU 6MEO_G | 99 | VEQMHEDIISLWDQSLKPCVKLPTLCVTINCDN~~~~~~~~~~~~~~~ |
| gp120 AU 6MEO_G | 151 | KGEIKNCFSNSISIRGKVQKEYAFFYKLDIIPIDNTDSYKLTSCNTSV |
| gp120 AU 6MEO_G | 187 | ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~TSYRLTSCNTSV |
| gp120 AU 6MEO_G | 201 | ITQACPQVSFEPHIPHIYCAPAGFAILKCNKTFNGTGCTNVVSTVQCTHG |
| gp120 AU 6MEO_G | 199 | ITQACPQVSFEPHIPHIYCTPAGAILKCNKTFNGTGCTNVVSTVQCTHG |
| gp120 AU 6MEO_G | 251 | IRPVVSTQLLNGSLAEVEEVRSVFNSTDNAKTIIVQLNTSVEINCTRPN |
| gp120 AU 6MEO_G | 249 | IKPVVSTQLLNGSLAEEDIVSERSENLNTNNAKTTIVQLKDPVINTCRPN |
| gp120 AU 6MEO_G | 301 | NNTRKRIRIQRPGPAVFVTIG~KIGNMRQAHCNISRASKWNNLTKAQASKL |
| gp120 AU 6MEO_G | 299 | NNTRKSIHIGPGR~~AFYATGDIIGDIRQAHCNLSRAQWNDTSLKIVTLK |
| gp120 AU 6MEO_G | 350 | REQFQSKNTIIFFKQSGGDPEITVHSNCGEFFYNSTQFLNSTWFNST |
| gp120 AU 6MEO_G | 347 | REQF~ENKTIKFQPSGGDPEITFHSNCGEFFYNCTQLNFLSTWNTNT |
| gp120 AU 6MEO_G | 400 | WSTEBSNNTGDSDIITLPCRIKIQIINWQKVGKAMYAPPISGQIRCSSNI |
| gp120 AU 6MEO_G | 396 | ~~~~~~~~~~~~~EDITLPCRIKIQVNLWQEVKVGKAMYAPPINKIKCSSNI |
| gp120 AU 6MEO_G | 450 | TGLLITRDGGNSN~NSEEIFRPGGDMRDNWRSELYKVKVIEPLGVAP |
| gp120 AU 6MEO_G | 435 | TGLLITRDGGNNTTEIFRPGGDMRDNWRSELYKVKVRIEPLG~~ |
| gp120 AU 6MEO_G | 499 | TKAKR |
| gp120 AU 6MEO_G | 499 | ~~~~ |