J Mol Model (2002) 8:217–222
DOI 10.1007/s00894-002-0089-6

ORIGINAL PAPER

**Xiu-Fan Shi · Shixi Liu · Jinggong Xiangyu
Yaping Zhang · Jingfei Huang · Shuqiu Liu
Ci-Quan Liu**

# Structural analysis of human CCR2b and primate CCR2b by molecular modeling and molecular dynamics simulation

Received: 5 April 2002 / Accepted: 21 May 2002 / Published online: 3 July 2002


**Abstract** CCR2b, a chemokine receptor for MCP-1, -2, -3, -4, plays an important role in a variety of diseases involving infection, inflammation, and/or injury, as well as being a coreceptor for HIV-1 infection. Two models of human CCR2b (hCCR2b) were generated by homology modeling and 1 ns restrained molecular dynamics (MD) simulation. In one only C113–C190 forms a disulfide bond (SS model); in another the potential C32–C277 disulfide bond was formed (2SS model). Analysis of the structures and averaged displacements of Cα atoms of the N-terminal residues shows that the main differences between the SS and 2SS models lie in a region D25YDYGAPCHKFD36; in the extracellular part of the 2SS model the accessible surfaces of N12, F23, Y26, Y28 and F35 are obviously raised and a more stable H-bond net is formed. The potential energy of the 2SS–water assembly finally fluctuated around –43,020 kJ $mol^{-1}$, which is about 302 kJ $mol^{-1}$ lower than that of the SS–water assembly. All these results suggest that the 2SS model is more favorable. The CCR2b genes of 17 primates were sequenced and four CCR2b models for primates *Ateles paniscus* (*A. pan*), *Hylobates leucogyneus* (*H. leu*), *Papio cynocephalus* (*P. cyn*) and *Trachypithecus francoist* (*T. fra*) were generated based on the 2SS model. A comparison of hCCR2b with primate CCR2b also supports the importance of the region D25YDYGAPCHKFD36. Electrostatic potential maps of human and primate CCR2b all display the dipolar characteristics of CCR2b with the negative pole located in the extracellular part and a strong positive pole in the cytoplasmic part. Based on the CCR2b model, we suggest that the main functional residues fall in the D25YDYGAPCHKFD36 region, and the negative electrostatic feature is a non-specific, but necessary, factor for ligands or gp120/CD4 binding.

**Keywords** CCR2 structure · Molecular dynamics · Electrostatic potentials

X.-F. Shi · J. Xiangyu · Y. Zhang · J. Huang · S. Liu (✉)
C.-Q. Liu
Cellular and Molecular Evolutionary Key Laboratory,
Kunming Institute of Zoology, Chinese Academy of Sciences,
Kunming, Yunnan, P.R. China, 650223
e-mail: tbrg@public.km.yn.cn
Tel.: +86 871 5195183, Fax: +86 871 5191823

S. Liu
Modern Biology Center, Yunnan University, Kunming, Yunnan,
P.R. China, 650091

## Introduction

Chemokines are a family of proinflammatory cytokines that attract and activate specific types of leukocytes via interaction with their specific receptors. This type of receptor belongs to the family of 7-transmembrane glycoproteins coupled to a G-protein signaling pathway. [1]

The CC chemokine receptor 2 (CCR2) is such a receptor for the chemokines monocyte chemoattractant-1, -2, -3 and -4 (MCP-1, -2, -3 and -4). [2] Two isoforms, CCR2a and CCR2b, produced by alternative splicing of the carboxyl terminal tails, have been identified. [3] MCP-1 and its receptor CCR2 have been shown to play a significant role in the pathology of inflammatory diseases, such as in resistance to Mycobacterium tuberculosis, in lung transplantation, in lipopolysaccharide-induced death and in delay-type hypersensitivity in the skin. CCR2 has also been demonstrated to be one of the coreceptors for human immunodeficiency virus type I (HIV-1) infection. [4]

Understanding the mechanism of ligand binding and the interaction between HIV-1 and CCR2 is essential for developing novel therapeutic agents against inflammatory diseases or HIV-1 infection. Experiments using chimeric chemokine receptors demonstrated that the amino terminal extracellular domain of CCR2 plays a crucial role for high affinity binding of MCP-1 and all three extracellular loops, especially the first one, are required for optimal signaling. [5, 6] Site-directed mutagenesis revealed the importance of N104 and E105 for high affinity agonist binding and the vital role of H100 for initiation of transmembrane signaling. [7] Studies on the ef-

fect of MCP-1 and of a panel of CCR2-specific monoclonal antibodies on the suppression of HIV-1 replication suggested that the binding site with HIV-1 is located in the N-terminus of CCR2. [4] Mirzadegan and coworkers indicated that E291 is the binding site for CCR2b antagonists. [8]

Characterization of structural and functional determinants of CCR2 for its ligand binding activity and HIV-1 coreceptor function is essential for understanding the mechanisms of inflammatory diseases, HIV-1 viral entry and developing novel therapeutic agents against inflammatory diseases or HIV-1 infection. Studies suggested that HIV may originate from SIV. Kuhmann et al. have used African green monkey as a highly informative model in research of immunodeficiency virus infections. [9, 10] In our laboratory Zhang and Xiangyu have sequenced the entire CCR5 and CCR2b genes for 17 primates, and studied these sequences from the evolutionary aspect. Based on these sequences and the known human CCR5 sequence, Yang has built 3-D structures of CCR5 of human and primates via molecular modeling, and discussed possible different features between human and primates in CCR5 serving as a coreceptor for HIV/SIV infection. [11] In this paper, we focus on the generation of a reasonable 3D structures of human CCR2b (hCCR2b) via molecular modeling and molecular dynamics (MD). Then based on the hCCR2b model we generate primate CCR2b (pCCR2b) models. Analyses of the extracellular characteristics of these models provides a structural basis for many experimental results and for research on new therapeutic agents against inflammatory diseases or HIV-1 infection.

## Material and methods

### Template-searching and prediction of transmembrane (TM) helices

The primary sequence of human CCR2b (hccr2b) was taken from Swiss-Prot (P41597). The crystal structures of bovine rhodopsin (1F88A, 1F88B) at 2.80 Å resolution was obtained as the modeling template of hCCR2b from the ExPdb database using Swiss-PdbViewer (spdbv 3.7b2, http://www.expasy.ch/spdbv/). [12] Blast results of hCCR2b and bovine rhodopsin (BLASTP V2.2.1, matrix: BLOSUM62) showed score: 166; identities: 21%; positives: 40% and gaps: 11%. Unfortunately, the sequence homologies between hCCR2b and bovine rhodopsin are not strong enough for building a 3D model straightforwardly. To help to define the TM regions, the secondary structure was predicted by Psi-Pred (http://globin.bio.warwick.ac.uk/psipred/) and the location of seven TM helices was predicted by TMHMM (v.2.0) (http://www.cbs.dtu.dk/services/TMHMM-2.0/) and HMMTOP (v.2.0) (http://www.enzim.hu/hmmtop/index.html). Finally, 3D-pssm (http://www.bmm.icnet.uk/servers/3dpssm/), which uses 1D and 3D sequence profiles coupled with secondary structure and solvation information to recognize the protein fold, [13] was used for alignment and fold recognition.

### Homology modeling

A homology model of hCCR2b was generated using spdbv 3.7b2 (http://www.expasy.ch/spdbv/). [12] The alignment was manually adjusted referring to the 3D-pssm alignment, prediction of secondary structure and TM helices. Loop-searching and primary energy optimization were performed with Swiss-Model, an automated homology modeling server developed at Glaxo Welcome Experimental Research in Geneva.

hCCR2b contains 13 cysteines, of which only C113–C190 forms a disulfide bond based on homological comparability with the template bovine rhodopsin. The model containing only a C113–C190 disulfide bond is denoted the SS model. By taking the potential disulfide bond C32–C277 into account, another model was built by connecting C32 and C277 of the SS model, and is denoted the 2SS model.

### Molecular mechanics and molecular dynamics simulations

The SS and 2SS models were first optimized using molecular mechanics, force field: Amber; van der Waals (VDW). cutoff: 9.5 Å; dielectric value: $1\times r$; summation method for electrostatic energy: cell-multipole; algorithms: steepest descent followed by conjugate gradient method; line-search precision: 0.100; final-convergence:10.0.

The extracellular and the cytoplasmic regions of the SS and 2SS models were embedded separately in 10 Å thick water layers, forming two assemblies: SS–water and 2SS–water, then further optimized by 1 ns molecular dynamics (MD) with TM helices and the outer 5 Å thick water layers fixed to prevent vaporization of the inner-layer water molecules in the MD process. The force field and non-bonded cutoff are the same as those in the molecular mechanics calculations except for the dielectric value, which was set to a constant 1. Rattle bond was adopted to allow for a 2 fs time step. The temperature was 300 K. 1 ns MD was run for each model.

Molecular mechanics and molecular dynamics simulations were performed using InsightII and Discover3 (v 98) software on a SGI O2 workstation.

### Generation of primate CCR2b models

Primate CCR2b (pCCR2b) models were generated based on the 2SS model (PDB 1KP1) of hCCR2b using sPdbv. Optimization of pCCR2bs was performed by molecular mechanics under the same conditions as that for hCCR2b.

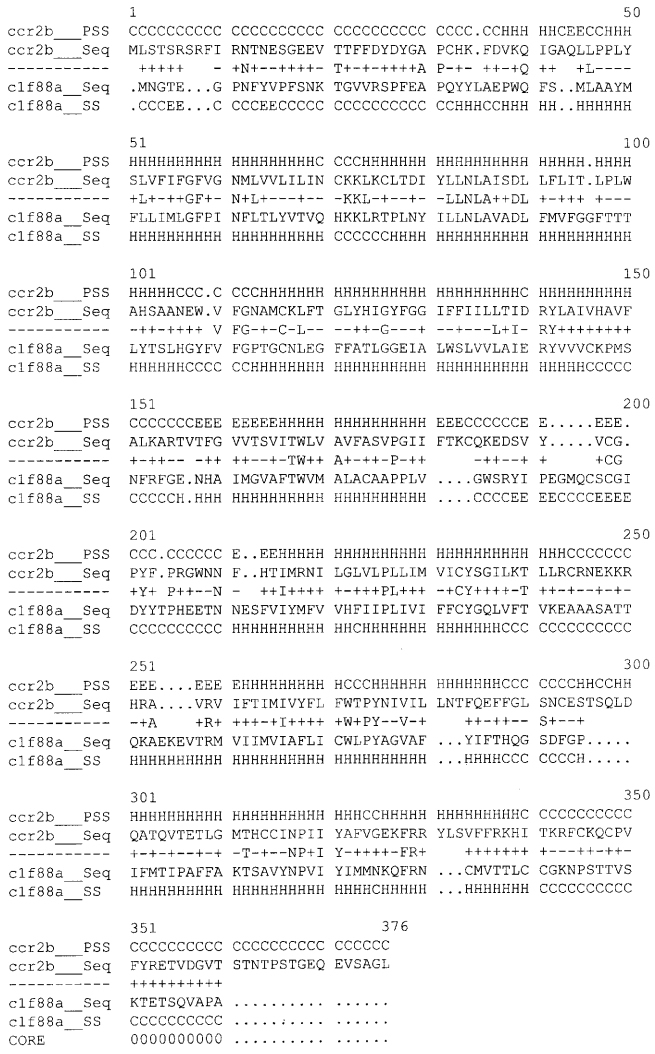


**Fig. 1** The 3D-pssm alignment of hCCR2b and bovine rhodopsin

## Results

### 7-TM helix regions

The alignment given by 3D-pssm (Fig. 1) shows that the predicted helix regions of CCR2b are essentially compatible with bovine rhodopsin, although the sequence homologies are not strong. So mainly based on this alignment ($E$=0.0027, highly confident) and referring to the other two TM predictions, the final 7-TM regions were obtained (see Table 1). Figure 1 and Table 1 show that the TM regions of our model are mainly in agreement with those predicted, and the alignment is essentially consistent with that in Fig. 1 except for the 7th TM helix. The main differences are in the starting positions of TM1, 3, and 5. According to the helix lengths of 1F88, the lengths of TM1, 3, and 5 in our model are a little longer than predicted, and consequently N-terminus and extracellular loop 1and 2 are slightly shorter.

### MD analysis

1 ns MD simulations were performed for SS–water and 2SS–water assemblies. In each case the potential energies of each assembly reached a plateau after ~500 ps, fluctuating around –42,718 and –43,020 kJ mol$^{-1}$, respectively. 50 frames were extracted from the last 500 ps for analysis.

For the SS and 2SS models, the displacements of each Cα of the N-terminal residues (1–39) with respect to the average structure of the SS and 2SS models, respectively, were calculated for each frame. The values averaged over their respective 50 frames are shown in Fig. 2a and b, respectively. In the SS model the displacements of most Cαs are between 0.3 and 0.9 Å, while in the 2SS model they are between 0.3 and 0.6 Å smaller than those of SS model (see Fig. 2a and b). This is expected from

**Table 1** The alignment of 7-TM helices and predicted TM

| TM no | Temp. model | Sequence[a] | Sequence no | Predicted TM[b] |
|---|---|---|---|---|
| 1 | 1F88 | WQFSM**L**AAYMF**L**LIML**GF**PI**N**F**L**TLYVTVQ | | |
| | CCR2B | IGAQL**L**PPLYS**L**VFIF**GF**VG**N**M**L**VVLILIN | 40–69 | 48–68, 46–68 |
| 2 | 1F88 | YI**LLNLA**VA**DL**FMVFGGFTTT | | |
| | CCR2B | IY**LLNLA**IS**DL**LFLITLPLWA | 79–99 | 79–99, 81–99 |
| 3 | 1F88 | **C**N**L**EGFFATL**G**GEIALWSLVV**L**AI**ERY**VV | | |
| | CCR2B | **C**K**L**FTGLYHI**G**YFGGIFFIIL**LTI**D**RY**LA | 113–141 | 126–143, 118–136 |
| 4 | 1F88 | NHAIMGVAF**TW**VM**A**LACAA**P**PL | | |
| | CCR2B | TFGVVTSVI**TW**LV**A**VFASV**P**GI | 155–176 | 154–177, 166–178 |
| 5 | 1F88 | **N**ESFV**I**YMFVVHFI**PL**IVIFF**CY**G | | |
| | CCR2B | **N**NFHT**I**MRNILGLVL**PL**LIMVI**CY**S | 199–223 | 208–225, 208–230 |
| 6 | 1F88 | **V**TRMVI**IM**VIA**FL**ICWLPYAGVAFYI**F** | | |
| | CCR2B | **V**RVIFT**IM**IVY**FL**FWTPYNIVILLNT**F** | 246–268 | 244–265, 243–265 |
| 7 | 1F88 | IF**MT**IPAFFAKTS**A**VYNPVIYI | | |
| | CCR2B | LG**MT**HCCINPIIY**A**FVGEKFRR | 293–314 | 292–309, ND |

[a] Identical residues are indicated in bold face
[b] ND Indicates no predicted data

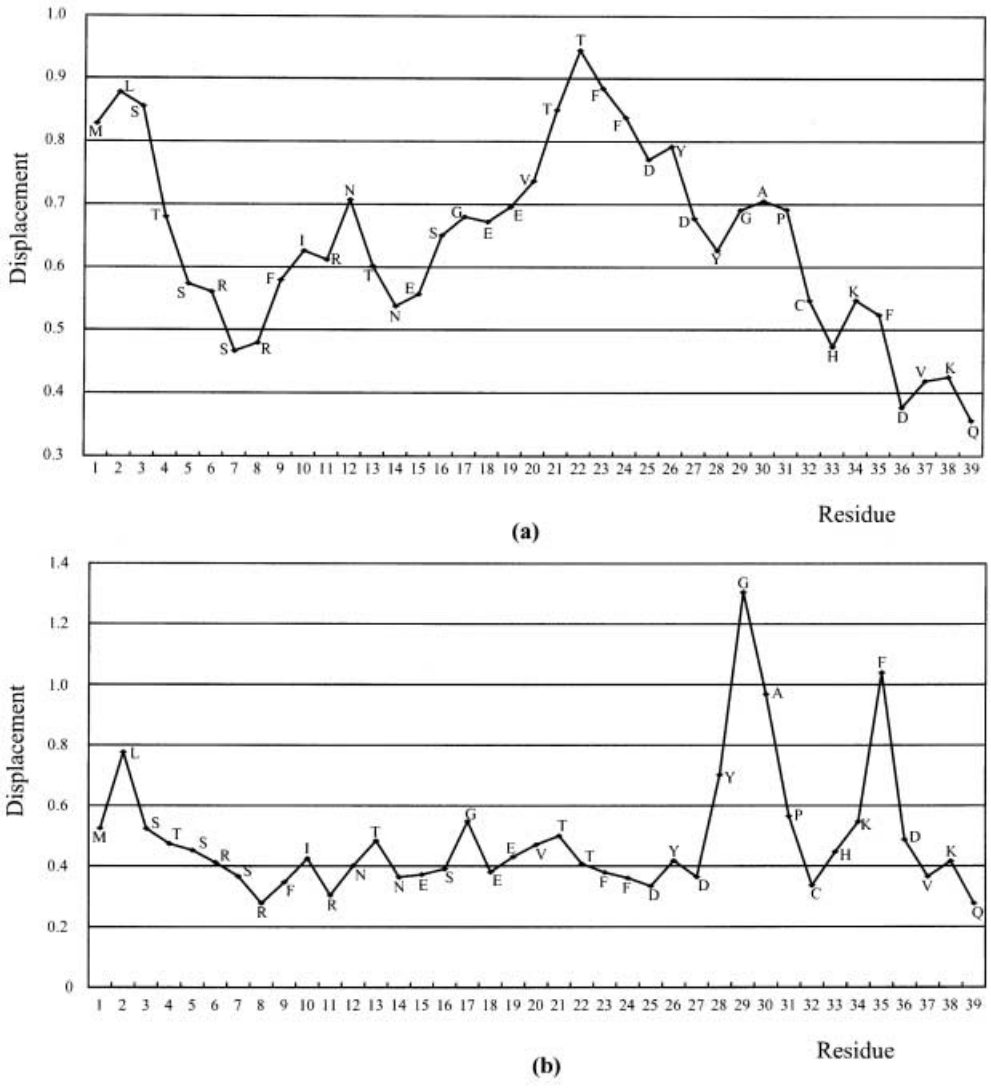


**Fig. 2** The average displacements of Cα atoms in the N-terminus of the SS model (**a**) and of the 2SS model (**b**). C32, which is labeled with a larger letter, forms a disulfide bond with C277 in the 2SS model

the restraint due to forming the C32–C277 disulfide bond. Unexpected, however, are the rising flexibilities of G29 and F35, forming two peaks with C32 at the valley in Fig. 2b. The reduction of most Cα's flexibility, especially of the ionic residues like R8, R11, D25, D27, of the N-terminus in the 2SS model reflects the possible formation of stable H-bond nets among ionic residues.

Examining the extracellular regions of the SS and 2SS models by superimposing their TM regions for all 50 frames of the SS– and 2SS–water assemblies, we found that the main differences were located in the region from D25 to D36 (DYDYGAPCHKFD) (see Fig. 3). The accessible surfaces measured for each N-terminus residue of each frame show significant increases for the following residues of the 2SS model compared with the SS model: N12 and F23 increase from 30–40% up to above 50%; Y26 from 22–25% up to 37–40%; Y28 from 26–32% up to 42–45%; F35 from 9–12% up to 44–50%.

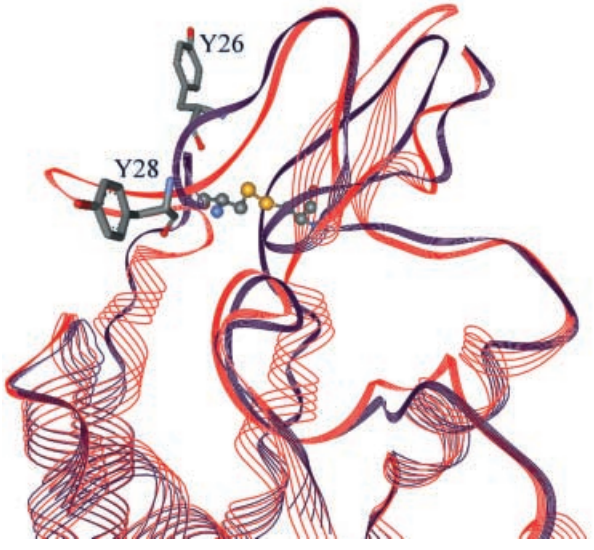


**Fig. 3** The extracellular region and partial TM region of the SS model (*red*) and the 2SS model (*black*), showing the conformational changes when the C32–C277 disulfide bond formed (circles and lines represent C32–C277 of the 2SS model). Sidechains of Y26 and Y28 of the 2SS model are shown

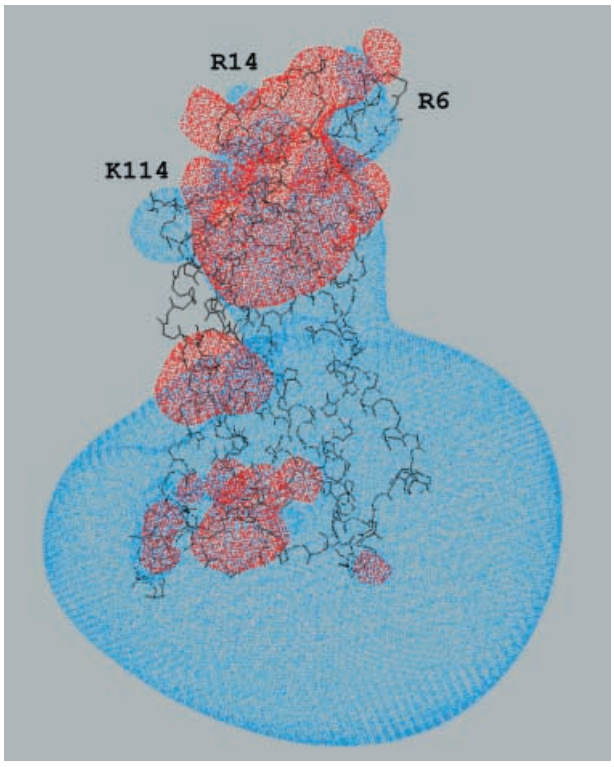


**Fig. 4** The electrostatic potential map of the hCCR2b 2SS model. *Red* indicates –3 and *blue* +3. K114, R11 and R6 form the three *blue* balls exposed on the *red* surface

### Evaluation of electrostatic properties of hCCR2b

hCCR2b consists of 360 residues, out of which 33 are basic residues (R+K) and 22 are acidic (D+E), so hCCR2b is a basic protein and shows a strong net positive charge (+11) in neutral solution (all His are considered as neutral). Interestingly, the charged residues show polar distributions in the hCCR2b molecule, there are 13 negatively charged and nine positively charged groups in or close to the extracellular region, whereas 21 positively charged and seven negatively charged groups are located in or close to the cytoplasmic part. Ten electrostatic potential maps were made for randomly selected ten frames of the 2SS model, one of which is shown in Fig. 4. The distribution characteristics of the charged groups, which are shared by the ten maps, can be seen on the electrostatic potential map (Fig. 4): the red color indicating –3 dominates over the extracellular part with only a few blue (indicating +3) balls caused by R6, R11 and K114, exposed on the red surface. In contrast, a big blue ball surrounds the cytoplasmic part, indicating the strong positive potential there.

### Comparison of hCCR2b with pCCR2b

The entire CCR2b genes for 17 primates were sequenced in our laboratory. The identities of the amino acid sequences between hCCR2b and the CCR2b of each primate are above 95%. There are altogether 32 residue mutation sites, including a deleted site 14: R6H, N14–, E15G, G17S, F23I, F35Y, I54T, C73S, E105G, M112V/A, I122V, F125L, V158A, V173I, I79V, K180E, C181S, K183E, V189I, N200S, I217V, V244L, I263V, N266S, G273S, S281R, V319M, T325A, R327H, V334I, G352A, A358V. There are seven, six and five mutation sites in the C- and N-termini and the 2nd extracellular loop, respectively; no mutation site found in the 2nd and 7th TM regions. Models for four primates: *A. pan*, *H. leu*, *P. cyn* and *T. fra*, which are relatively far from human beings from the evolutionary aspect, were built based on the 2SS model of hCCR2b. The loss of N14 in CCR2b of *A. pan* causes an obvious deviation of the backbone from the hCCR2b in the region from R11 to S16 (RANTES), making the exposure of R11 rise from 20% up to 40%, of Y28 from 44% to 48% due to N12 moving near to N14. In *H. leu*, *P. cyn* and *T. fra*, the mutation involved a few charged groups, but the basic features of hCCR2b in electrostatic potential maps, the strong dipole with R11, K38 and K114 exposed on the red surface, still remained.

## Discussion

It is difficult to build a very valid 3D model of a protein molecule if no suitable template with strong sequence homology can be found, as in the case of CCR2b. The definition of the 7-TM regions in this model does not seem accurate enough. However, considering the essential agreement of 7-TM alignment (Table 1) with the 3D-pssm alignment (Fig. 1) and the potential functional residues proposed by experiments, such as H100, N104, E105 and E291, are all contained in the extracellular part, the 2SS model is basically correct.

Experiments have demonstrated the importance of the N-terminus in high affinity binding with MCP-1 [5, 6] and in serving as a coreceptor for HIV-1 [4]. Prediction of the secondary structure of the N-terminal sequence suggests that there is no regular secondary structure, i.e. very flexible, the flexibility being reflected clearly in the MD simulations. The N-terminus contains five basic residues (R+K) and six acidic residues (D+E), which form a multiple H-bond network with the charged groups in the extracellular loops and reduce the average displacement of these charged groups in the course of the MD simulation, especially for the 2SS model (compare Fig. 2a with b.). This suggests that stronger H-bond interactions exist in the 2SS model. The potential energy of the 2SS–water assembly averaged over the last 500 ps in the 1 ns MD simulation is also much lower than that of the SS–water assembly, about 302 kJ $mol^{-1}$. The 2SS model is thus more favorable than the SS model from the angle of stability, supporting the mutation experiments: mutation of either C32 or C277 to A resulted in a complete loss of binding activity. [4]

Electrostatic analyses reveal the dipole characteristics of hCCR2b and pCCR2b, the negative pole is located in

the extracellular part, while a strong positive pole is found in the cytoplasmic region. Oligomeric modeling and electrostatic analysis of HIV gp120 revealed a basic region that faces away from the virus, toward the target cell membrane, and is conserved on core gp120. It is conjectured that nonspecific electrostatic interactions may play a role in the binding of viruses to polyanions. [14] We noticed that all the native ligands of CCR2, MCP-1, -2, -3, -4, are basic proteins, with +5, +5, +9, +8 electron net charges, respectively. In contrast, RENTAS, whose receptors are CCR1, CCR3 and CCR5, is a weakly acidic protein, with –1 electron net charge. The electrostatic potential maps of MCPs and RENTAS display different features although their 3D structures show similarity (not shown here). Electrostatic forces are long-range effective forces. We therefore guess that the extracellular acidic region may play a role in attracting certain kinds of chemokines and in the interaction with gp120. Residue-mutation experiments showed that two clusters of primarily basic residues (R24, K35, K38, K49, and Y13) of MCP-1 are crucial for binding to CCR2, [15] suggesting the role of the acidic region of receptors in chemokine binding. Of course, further exploration of the binding sites by theoretical modeling is necessary.

Tyrosine sulfation of proteins occurs widely in multicellular eukaryotic organisms. It was reported that the sulfation of N-terminal tyrosines of CCR5 and CCR2b is crucial for binding with their ligands or with HIV-1 gp120/CD4 complexes. [16, 17] Comparing the 2SS model with the SS model, we found that the accessible surfaces of Y26 and Y28 are clearly increased when the disulfide bond C32–C277 is formed. This may facilitate tyrosine sulfation and ligand contact.

The N-terminus of CCR2 is essential for ligand binding and serves as a coreceptor for HIV-1 infection. Analysis of frequent substitution polymorphisms in African green monkey (agm) CCR5 reveals that substitutions of D13N, Y14N, Q93R, and Q93K, of which the Y14N substitution eliminates a tyrosine sulfation site that is important for infections, all strongly inhibited infections by the SIVagm isolates in vitro. [10, 16] In the N-terminus of the 17 pCCR2b sequenced in our lab there are altogether six residue mutations, none of which, however, falls in the region from F24 to K34 (FDYDYGAPCHK). The similarity of the region F24 to K34 with the region D25 to D36, where the main differences between SS and 2SS models are shown (see section on MD analysis), further suggests the potential importance of the region D25 to D36 for function. However, the potentially functional residue E105 indicated by experiment [7] was found to be replaced by G in the primate *H. leu*, but we also noticed another mutation K183E in *H. leu* CCR2b, as well as E105 within 6 Å. We guess that E183 may function as E105 in binding with ligands.

The CCR2b model presented here provides some useful information. But considering the main limitation of lacking a suitable template, further modeling, e.g. with variable 7-TM regions, even their orientations, testing their functions by docking to see how well they agree with experimental results, may need to be done.

PDB access ID

1KAD is for the SS model and 1KP1 for the 2SS model, both were selected with the lowest potential energy among the 50 frames of SS–water and 2SS–water assemblies, respectively.

## References


1. Murphy PM (1994) Annu Rev Immunol 12:593–633
2. Balter M (1998) Science 280:825–826
3. Charo IF, Myers SJ, Herman A, Franci C, Connolly AJ, Coughlin SR (1994) Proc Natl Acad Sci USA 91:2752–2756
4. Frade JM, Llorente M, Mellado M, Alcami J, Gutierrez-Ramos JC, Zaballos A, Real G, Martinez-A C (1997) J Clin Invest 100:497–502
5. Monteclaro FS, Charo IF (1996) J Biol Chem 271:19084–19092
6. Monteclaro FS, Charo IF (1997) J Biol Chem 272:23186–23190
7. Han KH, Green SR, Tangirala RK, Tanaka S, Quehenberger O (1999) J Biol Chem 274:32055–32062
8. Mirzadegan T, Diehl F, Ebi B, Bhakta S, Polsky I, McCarley D, Mulkins M, Weatherhead GS, Lapierre JM, Dankwardt J, Morgans Jr D, Wilhelm R, Jarnagin K (2000) J Biol Chem 275:25562–25571
9. Kuhmann SE, Platt EJ, Kozak SL, Kabat D (1997) J Virol 71:8642–8656
10. Kuhmann SE, Madani N, Diop OM, Platt EJ, Morvan J, Muller-Trutwin MC, Barre-Sinoussi F, Kabat D (2001) J Virol 75:8449–8460
11. Yang J, Zhang Y-W, Huang J-F, Zhang Y-P, Liu C-Q (2000) Theor Chem 505:199–210
12. Guex N, Peitsch MC (1997) Electrophoresis 18:2714–2723
13. Kelley LA, MacCallum RM, Sternberg MJE (2000) J Mol Biol 299:499–520
14. Kwong PD, Wyatt R, Sattentau QJ, Sodroski J, Hendrickson WA (2000) J Virol 74:1961–1972
15. Hemmerich S, Paavola C, Bloom A, Bhakta S, Freedman R, Grunberger D, Krstenansky J, Lee S, McCarley D, Mulkins M, Wong B, Pease J, Mizoue L, Mirzadegan T, Polsky I, Thompson K, Handel TM, Jarnagin K (1999) Biochemistry 38:13013–13025
16. Farzan M, Mirzabekov T, Kolchinsky P, Wyatt R, Cayabyab M, Gerard NP, Gerard C, Sodroski J, Choe H (1999) Cell 96:667–676
17. Preobrazhensky AA, Dragan S, Kawano T, Gavrilin MA, Gulina IV, Chakravarty L, Kolattukudy PE (2000) J Immunol 165:5295–5303