PROTEINS: Structure, Function, and Bioinformatics 67:1193–1197 (2007)

STRUCTURE NOTE

Crystal Structure of Recombinant Human Stromal Cell-Derived Factor-1α

Eui Kyung Ryu,[1] Tae Gyun Kim,[1] Taek Hun Kwon,[1] In Duk Jung,[2] Dowook Ryu,[1] Yeong-Min Park,[2] Junhong Kim,[3] Kyo Han Ahn,[1] and Changill Ban[1]*

[1]*Department of Chemistry, Pohang University of Science and Technology (POSTECH), San 31, Hyojadong, Nam-gu, Pohang, Kyungpook 790-784, South Korea*
[2]*Department of Microbiology and Immunology and National Research Laboratory of Dendritic Cell Differentiation and Regulation, and Medical Research Institute, Pusan National University College of Medicine, Pusan 602-739, South Korea*
[3]*Department of Cardiology Internal Medicine, Pusan National University College of Medicine, Pusan 602-739, South Korea*

Introduction. Chemokines are a superfamily of cytokine molecules that mediate several cellular functions.[1–2] The chemokines superfamily has been classified into three species, CC, CXC, and CX_3C, based on the presence of an additional one or three amino acids between the first two cysteine residues.[3] Stromal cell-derived factor-1α (SDF-1α) is an unusual member of the CXC family that was first isolated from a mouse bone marrow stromal cell line.[4,5] SDF-1α has been implicated in a variety of physical functions, including hematopoiesis, cardiogenesis, B cell lymphopoiesis, and migration and cellular colonization of primordial germ cells in mammals.[6–8] SDF-1α interacts specifically with the physiological receptor CXCR4 for multicellular functions. A knockout experiment of either the *sdf*-1α or *cxcr*4 gene supported the claim that the SDF-1α/CXCR4 complex has an important role in embryonic development and proliferation.[5,9] Mice lacking the *sdf*-1α gene died in the perinatal period due to a severe cardiac ventricular system septum defect.[5]

Although various functions of SDF-1α have been reported, biochemical assays of its activity have been limited because the protein is not over-produced in a recombinant form. To overcome this obstacle, we expressed human SDF-1α in *E. coli* and obtained the refolded protein. The three-dimensional crystal structure of refolded SDF-1α displays some structural differences from previously reported structures. Also, the chemotaxis assay shows that refolded SDF-1α binds to receptors in the CCRF-CEM cell line, indicating that the refolded protein is an active form.

Materials and Methods. *Cloning, expression, and refolding of SDF-1α:* The *sdf-1*α gene was amplified from normal human fibroblast cDNA by polymerase chain reaction (PCR). The *sdf-1*α gene was inserted into pET28a containing a $(His)_6$-tag and a TEV protease site. This construction was transformed into *E. coli* strain BL21(DE3) for overexpression. The mutant K1A was constructed by site-directed mutagenesis using PCR. The transformed SDF-1α *E. coli* cell was cultured in Luria Bertani (LB) broth at 37°C until the culture reached an absorbance of 0.6 at 600 nm. Expression of SDF-1α was induced by addition of isopropyl-thio-β-D-galactopyranoside (IPTG) to a final concentration of 0.4 m*M*. After 4 h, the cultured cells were harvested by centrifuging. The pellet was resuspended in a PBS buffer (10 m*M* sodium phosphate, 150 m*M* NaCl, pH 7.4) and was disrupted by sonication. Inclusion bodies were solubilized in a denaturation buffer (8*M* urea, 20 m*M* Tris-HCl, pH 8.0, 500 m*M* NaCl, 0.5 m*M* β-mercaptoethanol, 3% glycerol) and allowed to stand for 2 h at room temperature. The denaturatant was centrifuged, and the supernatant was applied to a Ni-NTA column, which was preequilibrated with the denaturation buffer. On-column refolding[10] of the protein was induced by decreasing the urea concentration in a linear gradient from 8.0 to 0.0*M* over 6 h. After refolding, the protein was eluted by increasing the concentration of imidazole from 0 to 0.3*M* with elution buffer (20 m*M* Tris-HCl, pH 8.0, 500 m*M* NaCl, 0.5 m*M* β-mercaptoethanol, 3% glycerol, 300 m*M* imidazole). The refolded protein was pooled and concentrated. To remove the $(His)_6$-tag, the TEV protease was treated in pooled protein for 6 h at room temperature. After digestion, the

Grant sponsor: POSTECH Core Research Program; Grant number: 1RC0660201; Grant sponsor: Center for Integrated Molecular Systems at POSTECH; Grant number: R11-2000-070-070010; Grant sponsor: Korea Health Industry Development Institute; Grant number: A050426.

E. K. Ryu and T. G. Kim contributed equally to this work.

*Correspondence to: Changill Ban, Department of Chemistry, Pohang University of Science and Technology (POSTECH), San 31, Hyojadong, Nam-gu, Pohang, Kyungpook 790-784, South Korea. E-mail: ciban@postech.ac.kr

Received 23 October 2006; Revised 14 November 2006; Accepted 22 November 2006

Published online 13 March 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21350

TABLE I. Data and Refinement Statistics

	SDF-1α
Data statistics	
Number of reflections (unique/total)	15,982/177,130
Resolution range (Å)	50–1.81
Completeness (%)	98
R_{merge} (%)[a]	0.064
I/sigma	14.3
Unit cell parameter (Å)	$a = 35.22$, $b = 56.68$, $c = 71.63$
Space group	$P2_12_12_1$
Refinement statistics	
Resolution range (Å)	20–1.95
Number of reflections (unique/total)	9436/10,933
Number of protein atoms	1217
Number of water molecules	99
R/R_{free} (%)[b]	21.33/24.66
Average B-factor(Å^2)	38.3
r.m.s. deviation	
Bonds (Å)	1.65
Angles (°)	2.74

[a] $R_{merge} = \Sigma|I - <I>|/\Sigma<I>$, where I and $<I>$ are the measured and averaged intensities of multiple measurements of the same reflection, respectively. The summation is over all the observed reflections.
[b] $R = \Sigma|F_o - F_c|/\Sigma|F_o|$ calculated for all observed data. $R_{free} = \Sigma|F_o - F_c|/\Sigma|F_o|$ calculated for a specified number of randomly chosen reflections that were excluded from the refinement.

protein was applied to a desalting column, and a second Ni-NTA column was used to separate the target protein from the $(His)_6$-tag and undigested target protein. The crystallization and chemotaxis assays with refolded SDF-1α were performed after further purification through a Superdex peptide gel filtration column (Amersham Biosciences, GE Healthcare).

Crystallization, data collection, and structure determination: The purified SDF-1α was concentrated to 5 mg/mL. Crystals of refolded SDF-1α were grown using the hanging-drop vapor diffusion method at 20°C. The protein solution was mixed with an equal volume of reservoir solution (0.1*M* Hepes, pH 7.5, 1.4*M* Sodium citrate). The crystals were then transferred to the mother liquor, containing 15% ethylene glycol as a cryoprotectant of the fresh cooled system (100 K), at beamline 4A in the Pohang Acceleration Laboratory (Pohang, Korea). The diffraction data were processed, scaled, and reduced using the HKL2000 package.[11] The crystals belong to the orthorhombic space group $P2_12_12_1$, with cell dimensions $a = 35.22$ Å, $b = 56.68$ Å, and $c = 71.63$ Å and contained two molecules in the asymmetric unit. The crystal structure of the refolded SDF-1α was solved by the molecular replacement, using the Molrep program.[12] The previously reported crystal structure of SDF-1α (PDB code: 1QG7) was used as the search model.[13] The complete model was built manually using the Coot program.[14] Crystallographic refinement was carried out using CNS,[15] including simulated annealing, minimizing, and B-factor refinement. The geometries of the final models were checked with the program PROCHEK,[16] no residues were found in disallowed regions. Table I sets out data collection and refinement statistics.

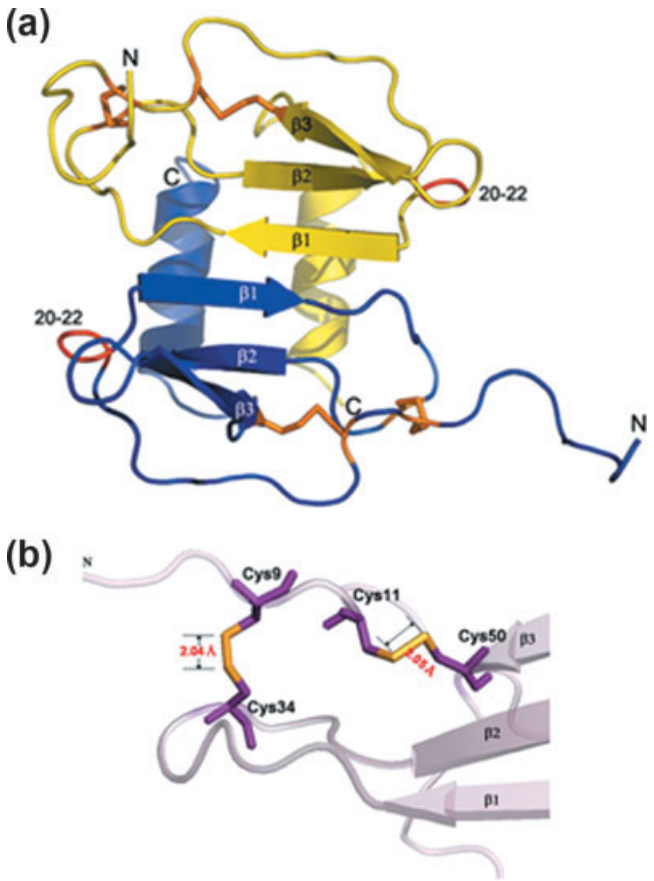

Fig. 1. (**a**) Refolded SDF-1α overall structure is a β-β-β-α topology with a three stranded antiparallel β-sheet, upon which lies a C-terminal α-helix. The A chain is shown in yellow and the B chain in blue. A 3_{10} helix region, found in other crystal structure, is shown to the long loop region (residues 20–22) in red. The disulfide bonds of two chains are shown in orange. (**b**) Disulfide bridges of refolded SDF-1α shown in purple. Disulfide bonds (orange) have 2.04 Å (Cys9–Cys34) and 2.05 Å (Cys11–Cys50) of length, respectively.

Chemotaxis assay of refolded SDF-1α: The CCRF-CEM cells, which naturally express CXCR4, were obtained from the KCLB (Seoul, Korea). The cells were grown and maintained in an RPMI 1640 medium with 25 m*M* Hepes and 10% FBS.[17] On the day of the experiment, the cells were harvested and washed once with migration assay buffer (RPMI 1640 with 25 m*M* Hepes, pH 7.5, 0.1% BSA). Also, both SDF-1α and its mutants were diluted into the same buffer. The migration assay was performed in a 48-well ChemoTx chamber (Neuroprobe, Gaithersburg, MD). A total of 27 μL of SDF-1α was added to the bottom well of the chamber. The wells were covered with a polycarbonate nucleopore filter (5-μm pore size), and 50 μL of a suspension of CCRF-CEM cells (4×10^6 cells/mL) was then added to the upper wells. After assembly, the chamber was incubated for 3 h at 37°C with 5% CO_2. The cells that had migrated into the bottom wells were then counted. Measurement of the activity of mutant K1A was also performed, as above. This experiment was performed in triplicate.

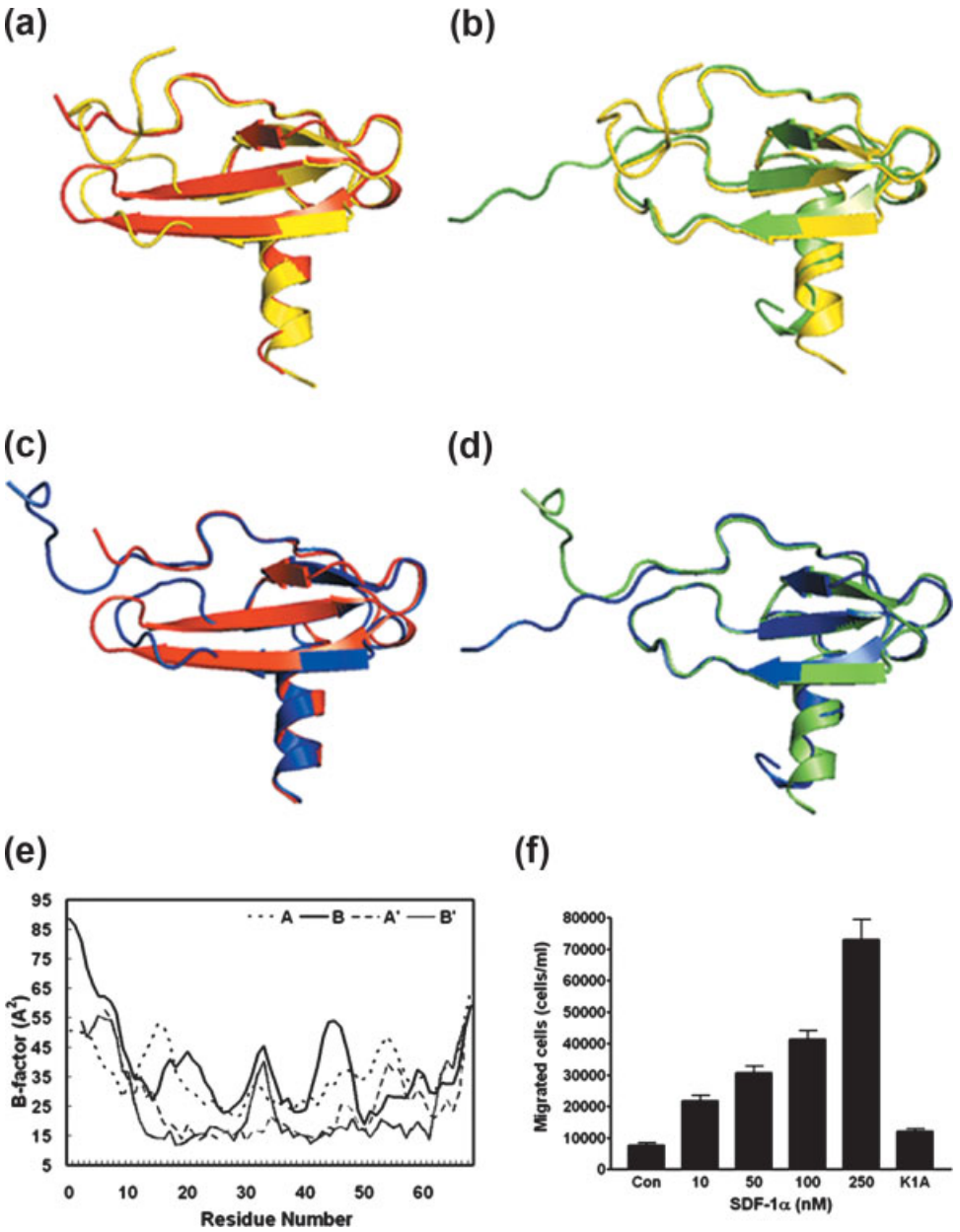

Fig. 2. The each two chains of refolded SDF-1α and mammalian SDF-1α are called A, B, A′, and B′ chain, respectively. Superposition of the two molecules was effected by optimizing the overlap of the C_α atom. (**a**) Superposition between A and A′, (**b**) superposition between A and B′, (**c**) superposition between B and A′, and (**d**) superposition between B and B′. Color code is: yellow (A), blue (B), red (A′), and green (B′). (**e**) According to the C_α atom of AA′, AB, BA′, and BB′, the RMS deviations calculated in (a) to (d) are 1.32, 4.45, 1.73, and 3.31, respectively. The B-factor profiles of β-strands are quite low, whereas the helix region appears higher than the strands; this indicates that the helix is more flexible than the strands. The loop regions of refolded SDF-1α and mammalian structure are more flexible than the secondary structures. Both the N- and C-terminal regions have an extremely high B factor value. (**f**) Chemotactic activity of refolded SDF-1α and its mutant K1A. Concentration-dependent chemotaxis of CCRF-CEM cells expressing CXCR4 is shown in response to refolded SDF-1α. The right side column of the graph shows that 250 n*M* of K1A mutant has no activity for chemotaxis to receptor CXCR4.

Results and Discussion. *Overall structure of refolded SDF-1α:* The crystal structure of refolded SDF-1α has a chemokine topology (β-β-β-α) very similar to the previously reported solution[18] and crystal structures[13,19] [Fig. 1(a)], consisting of a three-stranded antiparallel β-sheet followed by an α-helix. The structure of the refolded SDF-1α monomer looks like a forefinger protruding from a closed fist. Refolded SDF-1α exists in

TABLE II. Pairwise Comparison of Refolded SDF-1α With Reported SDF-1α Structures

	Refolded		1QG7		1A15		1SDF
	A(1–68)	B(1–68)	A(6–67)	B(2–67)	A(1–67)	B(9–64)	A(2(3)*–67)
β1	24–28	24–28	23–31	23–28	24–31	24–31	23–29
β2	38–42	38–42	35–42	38–42	35–42	35–42	37–42
β3	46–50	46–50	48–50	46–50	45–49	45–49	47–51
αH	57–66	56–66	56–65	56–61 63–65	56–64	56–64	58–65
3_{10}H	X	X	20–22	20–22	19–22	19–22	X

This means the number of residues through comparison of the secondary structures between the refolded SDF-1α and all other known SDF-1α. The chains each consist of three antiparallel β-strands (β1, β2, and β3) followed by a α-helix. The boldface is the refolded SDF-1α in this work. 1QG7, 1A15, and 1SDF are PDB code of mammalian structure, synthetic structure, and NMR structure, respectively. A parenthesis is the number of modeled residues. αH and 3_{10}H are the α-helix and 3_{10} helix, respectively. An asterisk is analogs (residues 2–67 and 3–67) of monomeric solution structure. SDF-1α for the previous NMR structure (1SDF) was synthesized by step-wise solid phase method using tBoc protection chemistry.[18] SDF-1α for the crystal structure of chemical synthesis (1A15) was generated by native chemical ligation of N-terminal peptide (1–33) and C-terminal peptide (34–67).[19] SDF-1α for crystal structure of mammalian (1QG7) was generated by using primary chicken embryo fibroblasts infected with a recombinant Sendai virus expressing human SDF-1α.[13] It is worth to notice that the different lengths of β1 and β2 in refolded SDF-1α are not due to the crystal lattice contacts.

three antiparallel β-strands (β1: residues 24–28, β2: residues 38–42, and β3: residues 46–50) in both the A and B chains. Also, the α-helix (residues 56–66) in the B chain is one residue longer (Lys56) than the α-helix (residues 57–66) in the A chain. A 3_{10} helix defined in previous crystal structures is not found. The dimer of the refolded SDF-1α structure is stabilized through interaction between the antiparallel β1 strands of the A and B chains.

It is well known that the extended N-terminal region plays an important role in interactions with receptor CXCR4.[18,20] The N-terminal region of refolded SDF-1α adopted an extended conformation. In the monomeric structure based on NMR spectroscopy,[18] SDF-1α was observed with a disordered N-terminal region (residues 1–8). Also, in the dimeric structure based on X-ray crystallography,[13,19] the side-chains of each residue in the N-terminal region of SDF-1α were not exactly assigned. However, in the present dimeric refolded SDF-1α structure, the side-chains of the entire residues 1–68 are clearly observed with corrected mapping through continuous electron density. In comparison between the present structure and previously reported structures, the N-terminal regions of chain A and B have different conformations. In case of the previous structure, the N-terminal regions of chain A and B are commonly interacting with neighboring SDF-1α molecule. In refolding SDF-1α, the N-terminal residues of chain A have also interaction with neighboring molecule. Especially, the Lys1 of chain A interacts with the neighboring His17 and Phe13 residues in crystal packing. However, the N-terminal region of chain B is snugly laid in the solvent layer in the crystal lattice without any interactions with neighboring molecules.

The superposition between the A and B chains shows that the N-terminal region of the A chain has a perpendicularly bent loop at the Cys9 residue and rest regions are very conserved. The Cys residue is involved in the formation of disulfide bridges [Fig. 1(b)]. Two disulfide bridges (Cys9–Cys34 and Cys11–Cys50) is a common structural feature of CXC chemokines.[1] These disulfide bridges provide strong geometric constraints on the surrounding residues.[13] The disulfide bridges provide good evidence that SDF-1α has refolded correctly.

Structural flexibility of refolded SDF-1α: The structure of the refolded SDF-1α refined in this work has a similar morphology to the mammalian SDF-1α (native SDF-1α expressed in mammalian cell), as described above. However, comparison of the refolded SDF-1α structure (A and B chains) with the mammalian SDF-1α structure (A′ and B′ chains) revealed several differences in local conformation. The differences in the lengths of the strand and the helix between the two structures are specified in Figure 2(a–d). β1 and β2 in the A chain are shorter by four residues and three residues, respectively, than β1′ (residues 23–28) and β2′ (residues 35–42) in the A′ chain. But β3 in the A chain is two residues longer than β3′ (residues 48–50) in the A′ chain [Fig. 2(a)]. The β1 in the A chain is shorter by one residue than the β1′ strand (residues 23–28) in the B′ chain, but β2 is the same length in the two structures, and β3 in the A chain is two residues longer than β3′ (residues 48–50) in the B′ chain [Fig. 2(b)]. The lengths of β1 and β2 in the B chain are shorter by four and three residues, respectively, than β1′ (residues 23–31) and β2′ (residues 35–42) in the A′ chain, but β3 in the B chain is two residues longer than β3′ (residues 48–50) in the A′ chain [Fig. 2(c)]. The length of β1 in the B chain is one residue shorter than β1′ (residues 23–28) in the B′ chain, but β2 is the same length in the two structures, and β3 in the B chain is longer by two residues than β3′ (residues 48–50) in the B′ chain [Fig. 2(d)]. The β-strand differences are affected by the length and direction of the non-overlapped region between the β-strands. In addition to differences in the β-strands, the α-helix of the refolded structure also shows some differences compared with the mammalian structure. Comparison of the α-helix between the B chain and the A′ chain shows the same length (residues 56–65), but the α-helix (residues 57–65) of the A chain is

one residue longer than the α′-helix of the A′ chain [Fig. 2(a,c)]. There are two α′-helices (residues 56–61 and 63–65) in the B′ chain, but the α-helix of the refolded SDF-1α has a single α-helix [Fig. 2(b,d)]. As a result of the conformational differences of the secondary structures, the C-terminal ends of the A and B chains in refolded SDF-1α run in the opposite direction. Other structural differences are found on in residue 20–22 region, which forms a single 3_{10} helix. This region of the refolded SDF-1α is composed of a very long loop formation. The differences in the lengths of the strands and the helices among mammalian,[13] NMR,[18] synthetic,[19] and refolded SDF-1α structures are summarized in Table II.

As shown in Figure 2(e), the B-factor profile of the refolded SDF-1α has uniformly higher values than those of the mammalian SDF-1α.[13] The N- and C-terminal regions of the refolded SDF-1α are very flexible, like mammalian SDF-1α.[13] In the B-factor profile for the refolded SDF-1α, the B-factor of the loop region between β1 and β2 is relatively higher than that of the other secondary structures. Because the β1′ of the A′ chain is five residues longer than that of the others, the length of the loop between β1′ and β2′ in the A′ chain is short and stable. The loop region (residues 20–22), composed of the 3_{10} helix of the mammalian SDF-1α, is also highly flexible. Based on several differences in the lengths of the secondary structures, it can be suggested that subtle differences have significant effects on the conformation and flexibility of the overall structure of SDF-1α.

Chemotaxis assay for SDF-1α and its mutants: Previous NMR studies[18] have demonstrated that the N-terminal residue of SDF-1α participates in the binding and activation of the target receptor. In particular, Lys1 and Pro2 are directly involved in interaction with CXCR4 receptor.[20] Also, the RFFESH motif (residues 12–17) is considered another region for CXCR4 receptor binding and activation.[18] To validate the activity of the refolded SDF-1α, we investigated the interactions of SDF-1α with the receptor using a CCRF-CEM cell line, which naturally expresses the CXCR4 [Fig. 2(f)]. The activity of the refolded SDF-1α (wild type) occurred with the same concentration-dependent pattern. A K1A mutant was used as a negative control. The chemotactic activity of the K1A mutant is similar to the basal level [Fig. 2(f)]. Thus, the structural differences between the refolded SDF-1α and mammalian SDF-1α are irrelevant to the chemotactic activity taking place through interaction between the ligand SDF-1α and the receptor CXCR4.

For the first time, a recombinant SDF-1α in *E. coli* has been obtained as a refolded form using on-column refolding. Structural differences between refolded SDF-1α and previously reported SDF-1α are found in various positions. The chemotactic properties of refolded SDF-1α interacting with CXCR4 receptor are similar to previous results. On the basis of the structural and functional analysis, the refolded SDF-1α proteins should provide a cornerstone for studying the structural and functional features of SDF-1α and its interaction with CXCR4.

Coordinates: The coordinates have been deposited in the PDB (ID: 2J7Z).

REFERENCES

1. Zlotnik A, Yoshie O. Chemokines: a new classification system and their role in immunity. Immunity 2000;12:121–127.
2. Dewan MZ, Ahmed S, Iwasaki Y, Ohba K, Toi M, Yamamoto N. Stromal cell-derived factor-1 and CXCR4 receptor interaction in tumor growth and metastasis of breast cancer. Biomed Pharmacother 2006;60:273–276.
3. Baggiolini M. Chemokines in pathology and medicine. J Int Med 2001;250:91–104.
4. Tashiro K, Tada H, Heilker R, Shirozu M, Nakano T, Honjo T. Signal sequence trap: a cloning stategy for secteted proteins and type I membrane proteins. Science 1993;261:600–603.
5. Nagasawa T, Hirota S, Tachibana K, Takakura N, Nishikawa S-I, Kitamura Y, Yoshida N, Kikutani H, Kishimoto T. Defects of B-cell lymphopoiesis and bone-marrow myelopoiesis in mice lacking the CXC chemokine PBSF/SDF-1. Nature 1996;382:635–638.
6. Doitsidou M, Reichman-Fried M, Stebler J, Koprunner M, Dorries J, Meyer D, Esguerra CV, Leung T, Raz E. Guidance of primordial germ cell migration by the chemokine SDF-1. Cell 2002;111:647–659.
7. Nagasawa T, Tachibana K, Kishimoto T. A novel CXC chemokine PBSF/SDF-1 and its receptor CXCR4: their functions in development, hematopoiesis and HIV infection. Semin Immunol 1998;10:179–185.
8. Ara T, Nakamura Y, Egawa T, Sugiyama T, Abe K, Kishimoto T, Matsui Y, Nagasawa T. Impaired colonization of the gonads by primodial germ cells in mice lacking a chemokine, stromal cell-derived factor-1 (SDF-1). Proc Natl Acad Sci USA 2003;100:5319–5323.
9. Ma Q, Jones D, Borghesani PR, Segal RA, Nagasawa T, Kishimoto T, Bronson RT, Springer TA. Impaired B-lymphopoiesi, myelopoiesis, and detailed cerebellar neuron migration in CXCR4- and SDF-1- deficient mice. Proc Natl Acad Sci USA 1998;95:9448–9453.
10. Ryu EK, Cho KJ, Kim JK, Harmer NJ, Blundell TL, Kim KH. Expression and purification of recombinant human fibroblast growth factor receptor in *Escherichia coli*. Protein Expr Purif 2006;49:15–22.
11. Otwinowski X, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;276:307–326.
12. Vagin AA, Teplyakov A. MOLREP: an automated program for molecular replacement. J Appl Crystallogr 1997;30:1022–1025.
13. Ohnishi Y, Senda T, Nandhagopap N, Sugimoto K, Shioda T, Nagai Y Mitsui T. Crystal structure of recombinant native SDF-1α with additional mutagenesis studies: an attempt at a more comprehensive interpretation of accumulated structure-activity relationship data. J Interferon Cytokine Res 2000;20:691–700.
14. Emsley P, Cowtan K. Coot: model building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 2004;60:2126–2132.
15. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 1998;54:905–921.
16. Laskowski RA, Macarthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality for assessing the accuracy of protein structures. J Appl Crystallogr 1993;26:283–291.
17. Sachpatzidis A, Benton BK, Manfredi JP, Wang H, Hamilton A, Dohlman HG, Lolis E. Identification of allostericpeptide agonists of CXCR4. J Biol Chem 2003;278:896–907.
18. Crump MP, Gong J-H, Loetscher P, Rajarathnam K, Amara A, Arenzana-Seisdedos F, Virelizier J-L, Baggiolini M, Sykes BD, Clark-Lewis I. Solution structure and basis for functional activity of stromal cell-derived factor-1: dissociation of CXCR4 activation form binding and inhibition of HIV-1. EMBO J 1997;16: 6996–7007.
19. Dealwis C, Fernandez EJ, Thompson DA, Simon RJ, Siani MA, Lolis E. Crystal structure of chemically synthesized [N33A] stromal cell-derived factor-1α, a potent ligand for the HIV-1 fusion coreceptor. Proc Natl Acad Sci USA 1998;95:6941–6946.
20. Huang X, Shen J, Cui M, Shen L, Luo X, Ling K, Pei G, Jiang H, Chen K. Molecular dynamics simulations on SDF-1α: binding with CXCR4 receptor. Biophys J 2003;84:171–184.