



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 5, 2021 – 04:08 PM JST

PDB ID : 7F1T  
Title : Crystal structure of the human chemokine receptor CCR5 in complex with MIP-1a  
Authors : Zhang, H.; Chen, K.; Tan, Q.; Han, S.; Zhu, Y.; Zhao, Q.; Wu, B.  
Deposited on : 2021-06-09  
Resolution : 2.60 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.22
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.22

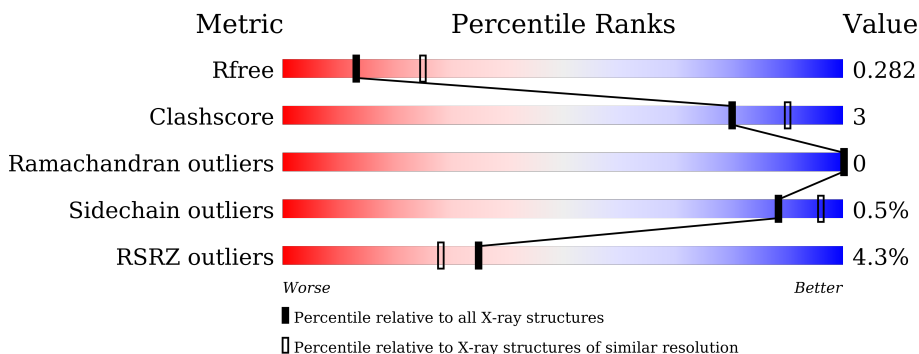
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	491	<div> <div>4%</div> <div>81%</div> <div>5%</div> <div>14%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3372 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called C-C motif chemokine 3,C-C chemokine receptor type 5,Rubredoxin,C-C chemokine receptor type 5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	423	Total	C	N	O	S	0	0	0
			3371	2224	524	599	24			

There are 62 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-81	CYS	THR	engineered mutation	UNP P10147
A	-22	GLY	-	linker	UNP P10147
A	-21	SER	-	linker	UNP P10147
A	-20	GLY	-	linker	UNP P10147
A	-19	SER	-	linker	UNP P10147
A	-18	GLY	-	linker	UNP P10147
A	-17	SER	-	linker	UNP P10147
A	-16	GLY	-	linker	UNP P10147
A	-15	SER	-	linker	UNP P10147
A	-14	GLY	-	linker	UNP P10147
A	-13	SER	-	linker	UNP P10147
A	-12	GLY	-	linker	UNP P10147
A	-11	SER	-	linker	UNP P10147
A	-10	GLY	-	linker	UNP P10147
A	-9	SER	-	linker	UNP P10147
A	-8	GLY	-	linker	UNP P10147
A	-7	SER	-	linker	UNP P10147
A	-6	GLY	-	linker	UNP P10147
A	-5	SER	-	linker	UNP P10147
A	-4	GLY	-	linker	UNP P10147
A	-3	SER	-	linker	UNP P10147
A	-2	GLY	-	linker	UNP P10147
A	-1	SER	-	linker	UNP P10147
A	0	GLY	-	linker	UNP P10147
A	1	SER	-	linker	UNP P10147
A	16	CYS	THR	engineered mutation	UNP P51681

*Continued on next page...*

*Continued from previous page...*

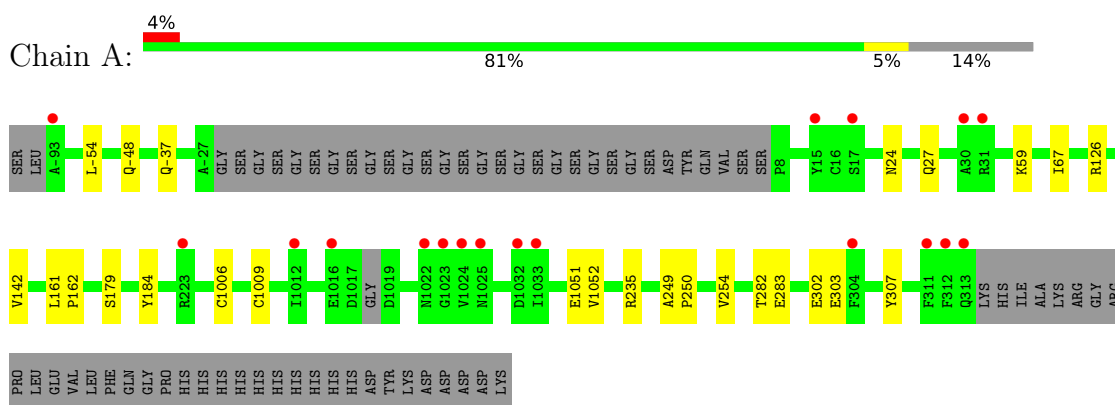
Chain	Residue	Modelled	Actual	Comment	Reference
A	58	TYR	CYS	engineered mutation	UNP P51681
A	64	ALA	MET	engineered mutation	UNP P51681
A	163	ASN	GLY	engineered mutation	UNP P51681
A	233	ASP	ALA	engineered mutation	UNP P51681
A	274	ALA	ARG	engineered mutation	UNP P51681
A	284	ALA	THR	engineered mutation	UNP P51681
A	303	GLU	LYS	engineered mutation	UNP P51681
A	320	GLY	-	expression tag	UNP P51681
A	321	ARG	-	expression tag	UNP P51681
A	322	PRO	-	expression tag	UNP P51681
A	323	LEU	-	expression tag	UNP P51681
A	324	GLU	-	expression tag	UNP P51681
A	325	VAL	-	expression tag	UNP P51681
A	326	LEU	-	expression tag	UNP P51681
A	327	PHE	-	expression tag	UNP P51681
A	328	GLN	-	expression tag	UNP P51681
A	329	GLY	-	expression tag	UNP P51681
A	330	PRO	-	expression tag	UNP P51681
A	331	HIS	-	expression tag	UNP P51681
A	332	HIS	-	expression tag	UNP P51681
A	333	HIS	-	expression tag	UNP P51681
A	334	HIS	-	expression tag	UNP P51681
A	335	HIS	-	expression tag	UNP P51681
A	336	HIS	-	expression tag	UNP P51681
A	337	HIS	-	expression tag	UNP P51681
A	338	HIS	-	expression tag	UNP P51681
A	339	HIS	-	expression tag	UNP P51681
A	340	HIS	-	expression tag	UNP P51681
A	341	ASP	-	expression tag	UNP P51681
A	342	TYR	-	expression tag	UNP P51681
A	343	LYS	-	expression tag	UNP P51681
A	344	ASP	-	expression tag	UNP P51681
A	345	ASP	-	expression tag	UNP P51681
A	346	ASP	-	expression tag	UNP P51681
A	347	ASP	-	expression tag	UNP P51681
A	348	LYS	-	expression tag	UNP P51681

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Zn 1 1	0	0

i

- Molecule 1: C-C motif chemokine 3,C-C chemokine receptor type 5,Rubredoxin,C-C chemokine receptor type 5



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	49.13Å 204.36Å 69.01Å 90.00° 105.53° 90.00°	Depositor
Resolution (Å)	50.00 – 2.60 34.06 – 2.60	Depositor EDS
% Data completeness (in resolution range)	94.9 (50.00-2.60) 90.3 (34.06-2.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.82 (at 2.61Å)	Xtriage
Refinement program	BUSTER v.2.8.0	Depositor
R, $R_{free}$	0.228 , 0.271 0.242 , 0.282	Depositor DCC
$R_{free}$ test set	888 reflections (4.90%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	96.1	Xtriage
Anisotropy	0.319	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 85.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3372	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	117.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.45% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	0.59	0/3463	0.79	0/4722

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	235	ARG	Sidechain

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3371	0	3302	18	0
2	A	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	3372	0	3302	18	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

The worst 5 of 18 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1051:GLU:HG2	1:A:1051:GLU:O	1.87	0.73
1:A:254:VAL:HG21	1:A:282:THR:HG21	1.77	0.66
1:A:1052:VAL:HG22	1:A:1052:VAL:O	2.06	0.56
1:A:24:ASN:O	1:A:27:GLN:HG2	2.06	0.56
1:A:126:ARG:HA	1:A:126:ARG:NE	2.21	0.55

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	417/491 (85%)	401 (96%)	16 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.



Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	369/431 (86%)	367 (100%)	2 (0%)	88 96

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	179	SER
1	A	283	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	-48	GLN
1	A	-37	GLN
1	A	21	GLN
1	A	194	GLN
1	A	280	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	423/491 (86%)	0.07	18 (4%) 35 28	66, 111, 176, 235	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1033	ILE	4.9
1	A	304	PHE	4.6
1	A	312	PHE	4.4
1	A	30	ALA	3.9
1	A	-93	ALA	3.7

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q < 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	ZN	A	1101	1/1	0.90	0.13	225,225,225,225	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.