



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 01:38 PM GMT

PDB ID : 3UF2
Title : Crystal structure of the human Colony-Stimulating Factor 1 (hCSF-1) cytokine
Authors : Elegheert, J.; Savvides, S.N.
Deposited on : 2011-10-31
Resolution : 2.75 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
A user guide is available at
<http://wwpdb.org/validation/2016/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
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

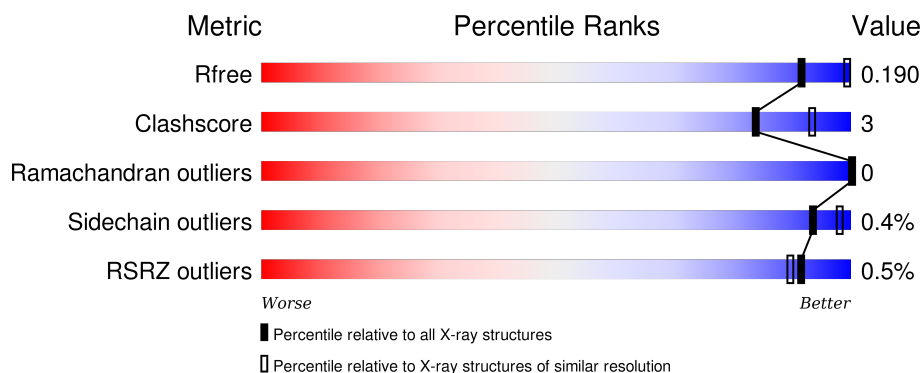
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.75 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}	91344	3340 (2.80-2.72)
Clashscore	102246	3829 (2.80-2.72)
Ramachandran outliers	100387	3767 (2.80-2.72)
Sidechain outliers	100360	3770 (2.80-2.72)
RSRZ outliers	91569	3352 (2.80-2.72)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	153	<div> <div>83%</div> <div>8% 8%</div> </div>
1	B	153	<div> <div>81%</div> <div>9% • 9%</div> </div>
1	C	153	<div> <div>86%</div> <div>8% 7%</div> </div>
1	D	153	<div> <div>86%</div> <div>• • 11%</div> </div>
1	E	153	<div> <div>%</div> <div>90% • 7%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	153	<div><div></div><div>82%</div><div>8%10%</div></div>
1	G	153	<div><div></div><div>84%</div><div>9%6%</div></div>
1	H	153	<div><div></div><div>%88%</div><div>10%</div></div>
1	I	153	<div><div></div><div>%86%</div><div>6%8%</div></div>
1	J	153	<div><div></div><div>2%86%</div><div>5%9%</div></div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 11607 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 Macrophage colony-stimulating factor 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	140	Total	C	N	O	S	0	1	0
			1145	722	193	219	11			
1	B	139	Total	C	N	O	S	0	1	0
			1140	718	193	218	11			
1	C	143	Total	C	N	O	S	0	1	0
			1183	748	196	228	11			
1	D	136	Total	C	N	O	S	0	2	0
			1107	701	182	213	11			
1	E	143	Total	C	N	O	S	0	1	0
			1158	729	193	225	11			
1	F	138	Total	C	N	O	S	0	2	0
			1138	718	191	218	11			
1	G	144	Total	C	N	O	S	0	0	0
			1167	732	195	229	11			
1	H	138	Total	C	N	O	S	0	0	0
			1124	711	184	218	11			
1	I	140	Total	C	N	O	S	0	2	0
			1140	718	193	218	11			
1	J	139	Total	C	N	O	S	0	1	0
			1136	716	191	218	11			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	EXPRESSION TAG	UNP P09603
A	-2	SER	-	EXPRESSION TAG	UNP P09603
A	-1	HIS	-	EXPRESSION TAG	UNP P09603
A	0	MET	-	EXPRESSION TAG	UNP P09603
B	-3	GLY	-	EXPRESSION TAG	UNP P09603
B	-2	SER	-	EXPRESSION TAG	UNP P09603
B	-1	HIS	-	EXPRESSION TAG	UNP P09603
B	0	MET	-	EXPRESSION TAG	UNP P09603
C	-3	GLY	-	EXPRESSION TAG	UNP P09603

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-2	SER	-	EXPRESSION TAG	UNP P09603
C	-1	HIS	-	EXPRESSION TAG	UNP P09603
C	0	MET	-	EXPRESSION TAG	UNP P09603
D	-3	GLY	-	EXPRESSION TAG	UNP P09603
D	-2	SER	-	EXPRESSION TAG	UNP P09603
D	-1	HIS	-	EXPRESSION TAG	UNP P09603
D	0	MET	-	EXPRESSION TAG	UNP P09603
E	-3	GLY	-	EXPRESSION TAG	UNP P09603
E	-2	SER	-	EXPRESSION TAG	UNP P09603
E	-1	HIS	-	EXPRESSION TAG	UNP P09603
E	0	MET	-	EXPRESSION TAG	UNP P09603
F	-3	GLY	-	EXPRESSION TAG	UNP P09603
F	-2	SER	-	EXPRESSION TAG	UNP P09603
F	-1	HIS	-	EXPRESSION TAG	UNP P09603
F	0	MET	-	EXPRESSION TAG	UNP P09603
G	-3	GLY	-	EXPRESSION TAG	UNP P09603
G	-2	SER	-	EXPRESSION TAG	UNP P09603
G	-1	HIS	-	EXPRESSION TAG	UNP P09603
G	0	MET	-	EXPRESSION TAG	UNP P09603
H	-3	GLY	-	EXPRESSION TAG	UNP P09603
H	-2	SER	-	EXPRESSION TAG	UNP P09603
H	-1	HIS	-	EXPRESSION TAG	UNP P09603
H	0	MET	-	EXPRESSION TAG	UNP P09603
I	-3	GLY	-	EXPRESSION TAG	UNP P09603
I	-2	SER	-	EXPRESSION TAG	UNP P09603
I	-1	HIS	-	EXPRESSION TAG	UNP P09603
I	0	MET	-	EXPRESSION TAG	UNP P09603
J	-3	GLY	-	EXPRESSION TAG	UNP P09603
J	-2	SER	-	EXPRESSION TAG	UNP P09603
J	-1	HIS	-	EXPRESSION TAG	UNP P09603
J	0	MET	-	EXPRESSION TAG	UNP P09603

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	24	Total O 24 24	0	0
2	B	18	Total O 18 18	0	0
2	C	19	Total O 19 19	0	0
2	D	20	Total O 20 20	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	E	16	Total 16	O 16	0	0
2	F	13	Total 13	O 13	0	0
2	G	17	Total 17	O 17	0	0
2	H	21	Total 21	O 21	0	0
2	I	9	Total 9	O 9	0	0
2	J	12	Total 12	O 12	0	0

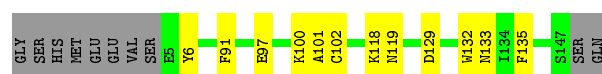
- Molecule 1: Macrophage colony-stimulating factor 1



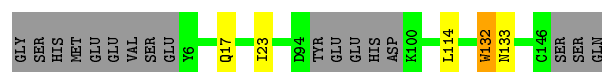
- Molecule 1: Macrophage colony-stimulating factor 1



- Molecule 1: Macrophage colony-stimulating factor 1



- Molecule 1: Macrophage colony-stimulating factor 1



- Molecule 1: Macrophage colony-stimulating factor 1



- Molecule 1: Macrophage colony-stimulating factor 1





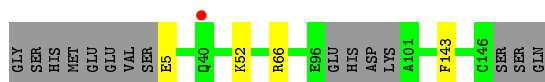
- Molecule 1: Macrophage colony-stimulating factor 1

Chain G: 84% 9% 6%



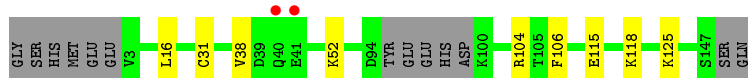
- Molecule 1: Macrophage colony-stimulating factor 1

Chain H: 88% 10% 2%



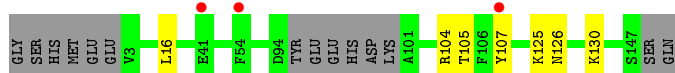
- Molecule 1: Macrophage colony-stimulating factor 1

Chain I: 86% 6% 8%



- Molecule 1: Macrophage colony-stimulating factor 1

Chain J: 86% 5% 9%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	61.15Å 104.49Å 116.24Å 90.00° 105.29° 90.00°	Depositor
Resolution (Å)	24.94 – 2.75 43.13 – 2.75	Depositor EDS
% Data completeness (in resolution range)	99.6 (24.94-2.75) 99.5 (43.13-2.75)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.13 (at 2.77Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_874)	Depositor
R, R_{free}	0.137 , 0.190 0.138 , 0.190	Depositor DCC
R_{free} test set	3659 reflections (10.00%)	DCC
Wilson B-factor (Å ²)	36.9	Xtriage
Anisotropy	0.072	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 32.3	EDS
Estimated twinning fraction	0.337 for h,-k,-h-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.26$	Xtriage
Outliers	0 of 36596 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	11607	wwPDB-VP
Average B, all atoms (Å ²)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 20.81 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 8.0139e-03.*

¹Intensities estimated from amplitudes.

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

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	0.22	0/1165	0.38	0/1566
1	B	0.22	0/1161	0.37	0/1562
1	C	0.23	0/1208	0.37	0/1628
1	D	0.22	0/1130	0.35	0/1522
1	E	0.22	0/1178	0.36	0/1585
1	F	0.22	0/1161	0.36	0/1560
1	G	0.23	0/1186	0.35	0/1598
1	H	0.22	0/1142	0.36	0/1538
1	I	0.22	0/1162	0.37	0/1562
1	J	0.22	0/1156	0.37	0/1555
All	All	0.22	0/11649	0.36	0/15676

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	1145	0	1133	8	1
1	B	1140	0	1114	12	0
1	C	1183	0	1149	10	1
1	D	1107	0	1081	2	0
1	E	1158	0	1133	6	0
1	F	1138	0	1125	9	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	G	1167	0	1126	10	0
1	H	1124	0	1092	3	0
1	I	1140	0	1128	5	0
1	J	1136	0	1120	6	0
2	A	24	0	0	1	1
2	B	18	0	0	2	1
2	C	19	0	0	1	0
2	D	20	0	0	0	0
2	E	16	0	0	1	0
2	F	13	0	0	0	0
2	G	17	0	0	0	0
2	H	21	0	0	0	0
2	I	9	0	0	0	0
2	J	12	0	0	1	0
All	All	11607	0	11201	67	2

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 67 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:B:112:GLN:OE1	2:B:162:HOH:O	2.07	0.72
1:E:9:HIS:O	2:E:159:HOH:O	2.09	0.68
1:B:82:GLU:O	1:B:86:ARG:NH1	2.27	0.68
1:C:129:ASP:OD1	2:C:160:HOH:O	2.13	0.67
1:H:5:GLU:N	1:H:5:GLU:OE1	2.28	0.66

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:21[B]:ARG:NH2	1:C:100:LYS:O[1_455]	1.94	0.26
2:A:160:HOH:O	2:B:154:HOH:O[2_546]	2.16	0.04

5.3 Torsion angles

5.3.1 Protein backbone

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	137/153 (90%)	134 (98%)	3 (2%)	0	100	100
1	B	136/153 (89%)	129 (95%)	7 (5%)	0	100	100
1	C	142/153 (93%)	136 (96%)	6 (4%)	0	100	100
1	D	134/153 (88%)	131 (98%)	3 (2%)	0	100	100
1	E	140/153 (92%)	134 (96%)	6 (4%)	0	100	100
1	F	136/153 (89%)	134 (98%)	2 (2%)	0	100	100
1	G	142/153 (93%)	137 (96%)	5 (4%)	0	100	100
1	H	134/153 (88%)	131 (98%)	3 (2%)	0	100	100
1	I	138/153 (90%)	133 (96%)	5 (4%)	0	100	100
1	J	136/153 (89%)	132 (97%)	4 (3%)	0	100	100
All	All	1375/1530 (90%)	1331 (97%)	44 (3%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains

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	134/145 (92%)	133 (99%)	1 (1%)	88	97
1	B	132/145 (91%)	131 (99%)	1 (1%)	86	96
1	C	137/145 (94%)	137 (100%)	0	100	100
1	D	128/145 (88%)	126 (98%)	2 (2%)	70	91

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	135/145 (93%)	135 (100%)	0	100	100
1	F	133/145 (92%)	133 (100%)	0	100	100
1	G	135/145 (93%)	134 (99%)	1 (1%)	88	97
1	H	130/145 (90%)	130 (100%)	0	100	100
1	I	133/145 (92%)	133 (100%)	0	100	100
1	J	133/145 (92%)	133 (100%)	0	100	100
All	All	1330/1450 (92%)	1325 (100%)	5 (0%)	93	98

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	142	SER
1	B	86	ARG
1	D	17	GLN
1	D	132	TRP
1	G	22	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	133	ASN
1	F	112	GLN
1	H	70	ASN
1	H	133	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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, 95th 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 ²)	Q<0.9
1	A	140/153 (91%)	-0.47	0	100 100	28, 51, 87, 112	0
1	B	139/153 (90%)	-0.53	0	100 100	24, 49, 86, 109	0
1	C	143/153 (93%)	-0.54	0	100 100	25, 47, 73, 96	0
1	D	136/153 (88%)	-0.48	0	100 100	29, 49, 87, 106	0
1	E	143/153 (93%)	-0.31	1 (0%)	89 86	31, 58, 100, 131	0
1	F	138/153 (90%)	-0.43	0	100 100	27, 55, 99, 116	0
1	G	144/153 (94%)	-0.51	0	100 100	31, 55, 85, 103	0
1	H	138/153 (90%)	-0.50	1 (0%)	89 86	24, 50, 87, 124	0
1	I	140/153 (91%)	-0.35	2 (1%)	78 73	36, 65, 101, 132	0
1	J	139/153 (90%)	-0.29	3 (2%)	65 59	36, 65, 105, 137	0
All	All	1400/1530 (91%)	-0.44	7 (0%)	91 90	24, 54, 94, 137	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	I	40	GLN	4.2
1	I	41	GLU	3.7
1	E	4	SER	2.3
1	J	41	GLU	2.3
1	J	54	PHE	2.2

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.