wwPDB X-ray Structure Validation Summary Report (i)

Aug 7, 2023 – 06:26 PM EDT

PDB ID	:	1PVH
Title	:	Crystal structure of leukemia inhibitory factor in complex with gp130
Authors	:	Boulanger, M.J.; Bankovich, A.J.; Kortemme, T.; Baker, D.; Garcia, K.C.
Deposited on	:	2003-06-27
Resolution	:	2.50 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

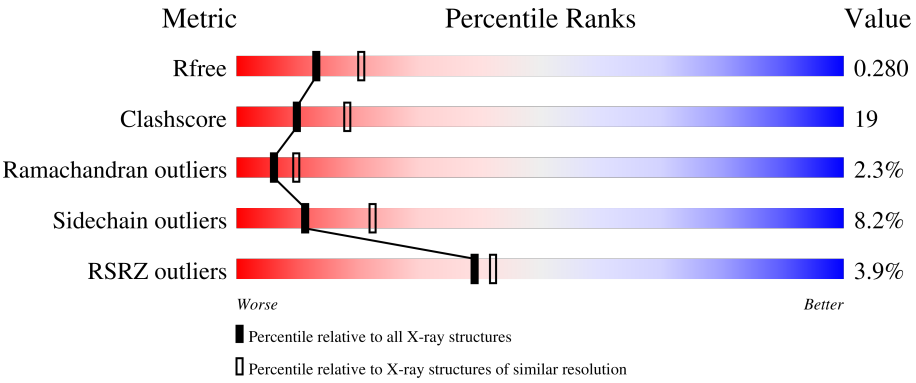
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 >=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 <=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	201	6% 68% 29% ·
1	C	201	3% 67% 29% ·
2	B	169	2% 60% 31% 9% ·
2	D	169	4% 56% 36% 8% ·

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 6026 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 Interleukin-6 receptor beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	201	Total	C	N	O	S	0	0	0
			1624	1031	269	317	7			
1	C	201	Total	C	N	O	S	0	0	0
			1624	1031	269	317	7			

- Molecule 2 is a protein called Leukemia inhibitory factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	169	Total	C	N	O	S	4	0	0
			1292	823	232	230	7			
2	D	169	Total	C	N	O	S	4	0	0
			1292	823	232	230	7			

- Molecule 3 is IODIDE ION (three-letter code: IOD) (formula: I).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	I	0	0
			1	1		
3	D	1	Total	I	0	0
			1	1		

- Molecule 4 is water.

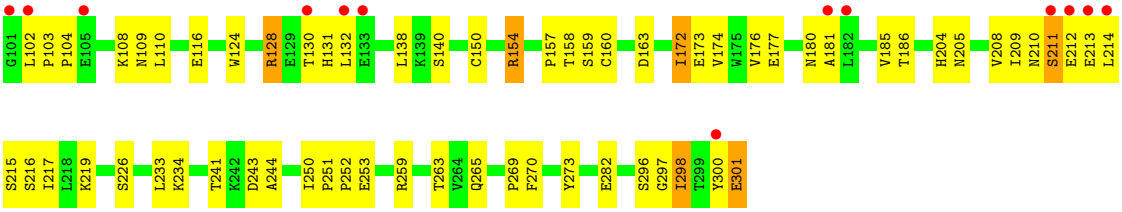
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	71	Total	O	0	0
			71	71		
4	B	33	Total	O	0	0
			33	33		
4	C	61	Total	O	0	0
			61	61		
4	D	27	Total	O	0	0
			27	27		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

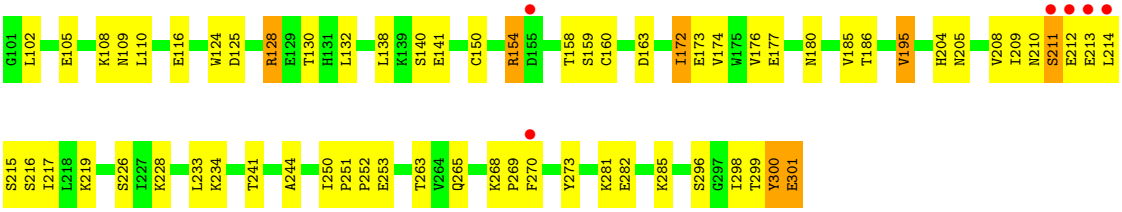
• Molecule 1: Interleukin-6 receptor beta chain

Chain A:



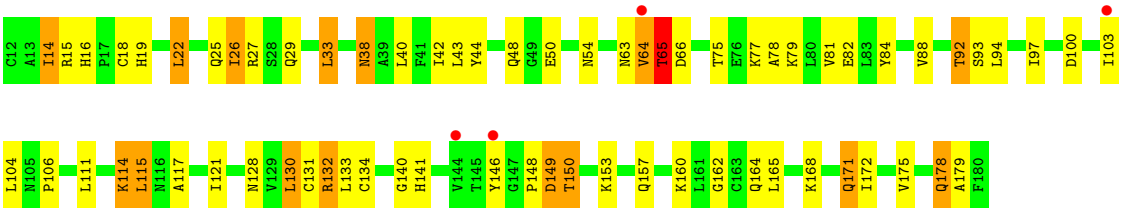
• Molecule 1: Interleukin-6 receptor beta chain

Chain C:



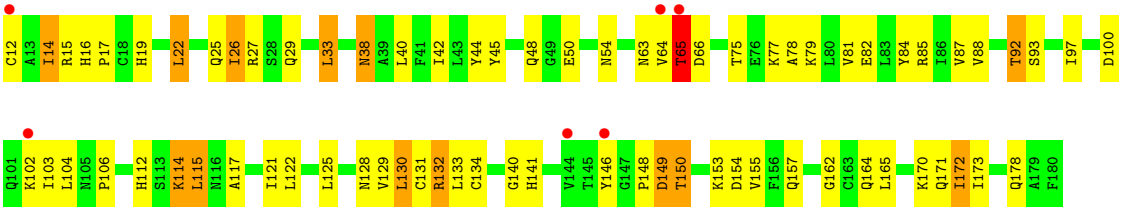
• Molecule 2: Leukemia inhibitory factor

Chain B:



• Molecule 2: Leukemia inhibitory factor

Chain D:


C12 A13 I14 R15 H16 P17 C18 H19 L22 Q25 I26 R27 S28 Q29 L33 M38 A39 L40 F41 I42 L43 Y44 Y45 Q48 G49 E50 N54 M63 V64 T65 D66 T75 E76 K77 A78 K79 L80 V81 E82 L83 Y84 R85 I86 V87 V88 T92 S93 I97 D100
Q101 K102 I103 L104 N105 P106 H112 S113 K114 L115 N116 A117 I121 L122 L125 N128 V129 L130 C131 R132 L133 C134 G140 H141 V144 T145 Y146 G147 P148 D149 T150 K153 D154 V155 F156 Q157 G162 C163 Q164 L165 K170 Q171 I172 I173 Q178 A179 F180

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	79.71Å 86.70Å 146.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.86 – 2.50 39.85 – 2.50	Depositor EDS
% Data completeness (in resolution range)	90.8 (39.86-2.50) 90.9 (39.85-2.50)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.87 (at 2.51Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.248 , 0.289 0.242 , 0.280	Depositor DCC
R_{free} test set	2616 reflections (8.03%)	wwPDB-VP
Wilson B-factor (Å^2)	50.9	Xtriage
Anisotropy	0.639	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 50.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	6026	wwPDB-VP
Average B, all atoms (Å^2)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.26% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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

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	0.54	0/1668	0.74	0/2267
1	C	0.54	0/1668	0.75	0/2267
2	B	0.50	0/1318	0.71	0/1783
2	D	0.51	0/1318	0.69	0/1783
All	All	0.53	0/5972	0.72	0/8100

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
2	D	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
2	D	45	TYR	Sidechain

5.2 Too-close contacts

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	1624	0	1571	60	0
1	C	1624	0	1571	68	0
2	B	1292	0	1304	46	0
2	D	1292	0	1304	50	0
3	A	1	0	0	0	0
3	D	1	0	0	0	0
4	A	71	0	0	5	0
4	B	33	0	0	0	0
4	C	61	0	0	2	0
4	D	27	0	0	4	0
All	All	6026	0	5750	221	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 19.

The worst 5 of 221 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:177:GLU:HG3	1:A:186:THR:HG22	1.46	0.96
1:C:177:GLU:HG3	1:C:186:THR:HG22	1.49	0.92
2:D:40:LEU:HD11	2:D:115:LEU:HD13	1.58	0.83
1:A:130:THR:HG21	1:A:180:ASN:ND2	1.94	0.83
2:B:38:ASN:O	2:B:42:ILE:HG12	1.80	0.81

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	199/201 (99%)	187 (94%)	9 (4%)	3 (2%)	10	18
1	C	199/201 (99%)	188 (94%)	8 (4%)	3 (2%)	10	18

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	B	167/169 (99%)	152 (91%)	10 (6%)	5 (3%)	4 6
2	D	167/169 (99%)	151 (90%)	10 (6%)	6 (4%)	3 4
All	All	732/740 (99%)	678 (93%)	37 (5%)	17 (2%)	6 10

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	65	THR
2	B	150	THR
2	B	153	LYS
1	C	211	SER
1	C	300	TYR

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	185/185 (100%)	176 (95%)	9 (5%)	25 47
1	C	185/185 (100%)	176 (95%)	9 (5%)	25 47
2	B	139/144 (96%)	121 (87%)	18 (13%)	4 8
2	D	139/144 (96%)	122 (88%)	17 (12%)	5 9
All	All	648/658 (98%)	595 (92%)	53 (8%)	11 22

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	154	ARG
1	C	301	GLU
2	D	157	GLN
1	C	159	SER
1	C	172	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	157	GLN
2	D	112	HIS
2	D	29	GLN
2	D	63	ASN
2	B	157	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 2 ligands modelled in this entry, 2 are 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

There are no such residues in this entry.

5.8 Polymer linkage issues

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, 95^{th} 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	201/201 (100%)	0.21	13 (6%)	18	19	29, 48, 78, 97	0
1	C	201/201 (100%)	0.06	6 (2%)	50	53	32, 47, 79, 98	0
2	B	169/169 (100%)	0.15	4 (2%)	59	62	34, 54, 87, 98	1 (0%)
2	D	169/169 (100%)	0.22	6 (3%)	42	46	36, 55, 87, 99	1 (0%)
All	All	740/740 (100%)	0.16	29 (3%)	39	42	29, 51, 82, 99	2 (0%)

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	211	SER	6.9
1	C	212	GLU	6.6
2	B	64	VAL	5.9
1	C	213	GLU	5.7
1	A	212	GLU	4.9

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

There are no monosaccharides in this entry.

6.4 Ligands

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^{th} 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^2)	Q<0.9
3	IOD	A	302	1/1	1.00	0.14	44,44,44,44	0
3	IOD	D	202	1/1	1.00	0.14	43,43,43,43	0

6.5 Other polymers (i)

There are no such residues in this entry.