Full wwPDB X-ray Structure Validation Report (i)

Dec 5, 2022 – 02:24 pm GMT

PDB ID	:	7ZMZ
Title	:	Engineered Interleukin 2 bound to CD25 receptor
Authors	:	Fyfe, P.K.; Moraga, I.; Gaggero, S.; Mitra, S.
Deposited on	:	2022-04-20
Resolution	:	3.20 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

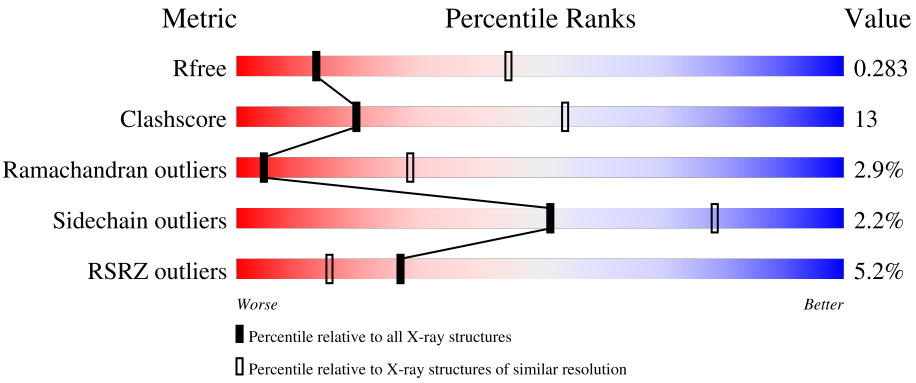
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.31.3
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

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 3.20 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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	146	6% 61% 23% 13%
2	D	258	2% 31% 16% 52%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2028 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-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	127	Total	C	N	O	S	0	0	0
			1031	664	169	191	7			

There are 19 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP P60568
A	37	HIS	THR	engineered mutation	UNP P60568
A	38	LEU	ARG	engineered mutation	UNP P60568
A	41	SER	THR	engineered mutation	UNP P60568
A	42	TYR	PHE	engineered mutation	UNP P60568
A	43	GLY	LYS	engineered mutation	UNP P60568
A	77	PHE	ASN	conflict	UNP P60568
A	78	ASN	PHE	conflict	UNP P60568
A	134	ALA	-	expression tag	UNP P60568
A	135	ALA	-	expression tag	UNP P60568
A	136	ALA	-	expression tag	UNP P60568
A	137	HIS	-	expression tag	UNP P60568
A	138	HIS	-	expression tag	UNP P60568
A	139	HIS	-	expression tag	UNP P60568
A	140	HIS	-	expression tag	UNP P60568
A	141	HIS	-	expression tag	UNP P60568
A	142	HIS	-	expression tag	UNP P60568
A	143	HIS	-	expression tag	UNP P60568
A	144	HIS	-	expression tag	UNP P60568

- Molecule 2 is a protein called Interleukin-2 receptor subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	125	Total	C	N	O	S	0	0	0
			997	620	180	182	15			

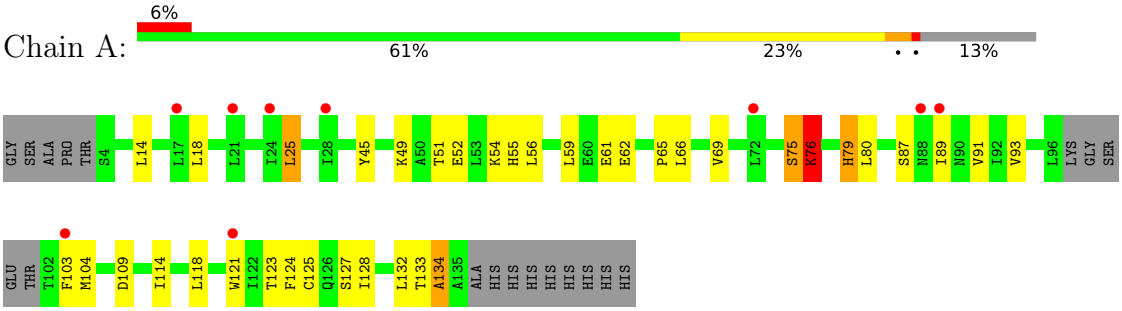
There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-1	GLY	-	expression tag	UNP P01589
D	218	ALA	-	expression tag	UNP P01589
D	219	ALA	-	expression tag	UNP P01589
D	220	ALA	-	expression tag	UNP P01589
D	221	LEU	-	expression tag	UNP P01589
D	222	GLU	-	expression tag	UNP P01589
D	223	VAL	-	expression tag	UNP P01589
D	224	LEU	-	expression tag	UNP P01589
D	225	PHE	-	expression tag	UNP P01589
D	226	GLN	-	expression tag	UNP P01589
D	227	GLY	-	expression tag	UNP P01589
D	228	PRO	-	expression tag	UNP P01589
D	229	GLY	-	expression tag	UNP P01589
D	230	ALA	-	expression tag	UNP P01589
D	231	ALA	-	expression tag	UNP P01589
D	232	GLY	-	expression tag	UNP P01589
D	233	GLY	-	expression tag	UNP P01589
D	234	GLY	-	expression tag	UNP P01589
D	235	LEU	-	expression tag	UNP P01589
D	236	ASN	-	expression tag	UNP P01589
D	237	ASP	-	expression tag	UNP P01589
D	238	ILE	-	expression tag	UNP P01589
D	239	PHE	-	expression tag	UNP P01589
D	240	GLU	-	expression tag	UNP P01589
D	241	ALA	-	expression tag	UNP P01589
D	242	GLN	-	expression tag	UNP P01589
D	243	LYS	-	expression tag	UNP P01589
D	244	ILE	-	expression tag	UNP P01589
D	245	GLU	-	expression tag	UNP P01589
D	246	TRP	-	expression tag	UNP P01589
D	247	HIS	-	expression tag	UNP P01589
D	248	GLU	-	expression tag	UNP P01589
D	249	HIS	-	expression tag	UNP P01589
D	250	HIS	-	expression tag	UNP P01589
D	251	HIS	-	expression tag	UNP P01589
D	252	HIS	-	expression tag	UNP P01589
D	253	HIS	-	expression tag	UNP P01589
D	254	HIS	-	expression tag	UNP P01589
D	255	HIS	-	expression tag	UNP P01589
D	256	HIS	-	expression tag	UNP P01589

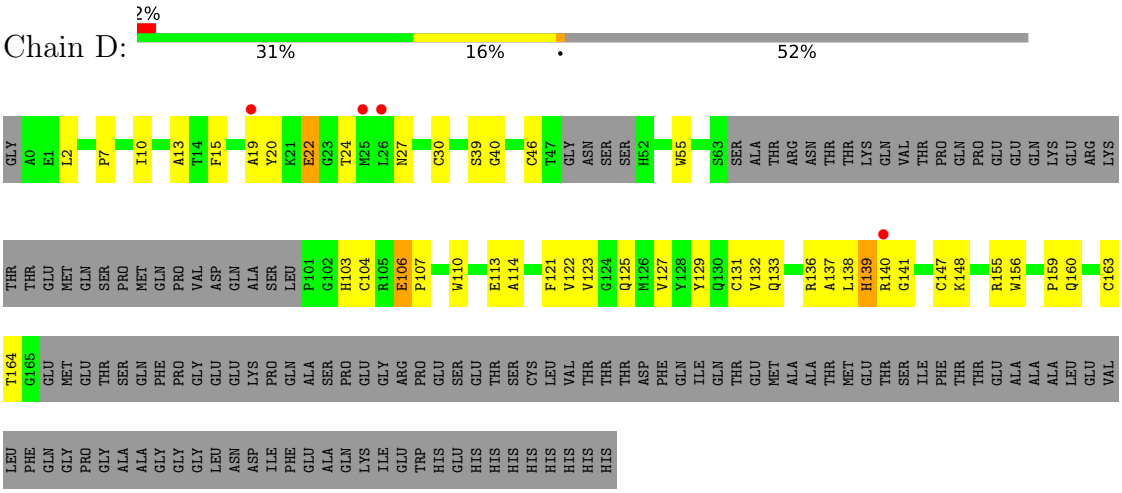
3 Residue-property plots (i)

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-2



- Molecule 2: Interleukin-2 receptor subunit alpha



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	89.05Å 89.05Å 116.68Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	64.30 – 3.20 64.34 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.3 (64.30-3.20) 99.8 (64.34-3.20)	Depositor EDS
R_{merge}	0.23	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.26 (at 3.19Å)	Xtriage
Refinement program	PHENIX 1.17.1-3660	Depositor
R, R_{free}	0.237 , 0.258 0.242 , 0.283	Depositor DCC
R_{free} test set	433 reflections (4.71%)	wwPDB-VP
Wilson B-factor (Å^2)	116.4	Xtriage
Anisotropy	0.596	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.041 for -h,-k,l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	2028	wwPDB-VP
Average B, all atoms (Å^2)	146.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.60% 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

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	0.35	0/1048	0.62	1/1416 (0.1%)
2	D	0.41	0/1024	0.67	0/1382
All	All	0.38	0/2072	0.64	1/2798 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	25	LEU	CB-CG-CD2	-6.24	100.39	111.00

There are no chirality outliers.

There are no planarity outliers.

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	1031	0	1057	24	0
2	D	997	0	938	30	0
All	All	2028	0	1995	51	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 13.

All (51) 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)
2:D:20:TYR:HB3	2:D:24:THR:HG21	1.53	0.88
1:A:51:THR:HG23	1:A:52:GLU:HG3	1.57	0.84
1:A:25:LEU:HB2	1:A:118:LEU:HD23	1.61	0.81
1:A:123:THR:O	1:A:127:SER:OG	2.08	0.71
1:A:61:GLU:OE1	2:D:39:SER:N	2.19	0.71
2:D:10:ILE:HB	2:D:13:ALA:HB3	1.73	0.70
2:D:139:HIS:O	2:D:141:GLY:N	2.29	0.66
2:D:15:PHE:HD2	2:D:127:VAL:HG21	1.62	0.64
1:A:61:GLU:HG2	2:D:39:SER:OG	1.98	0.63
1:A:66:LEU:HD11	1:A:118:LEU:HD11	1.83	0.61
1:A:114:ILE:H	1:A:114:ILE:HD12	1.65	0.61
2:D:7:PRO:HG3	2:D:156:TRP:CZ2	2.39	0.58
2:D:55:TRP:CE2	2:D:107:PRO:HG3	2.39	0.57
2:D:121:PHE:CD1	2:D:125:GLN:HG2	2.41	0.55
1:A:59:LEU:HD12	1:A:121:TRP:CD2	2.42	0.55
1:A:109:ASP:N	1:A:109:ASP:OD1	2.39	0.54
1:A:18:LEU:HD13	1:A:125:CYS:HB2	1.90	0.54
1:A:59:LEU:HD12	1:A:121:TRP:CG	2.46	0.51
1:A:103:PHE:CG	1:A:104:MET:N	2.79	0.51
2:D:15:PHE:HB2	2:D:127:VAL:HG22	1.91	0.51
2:D:140:ARG:NE	2:D:140:ARG:HA	2.26	0.49
2:D:20:TYR:O	2:D:103:HIS:HB3	2.12	0.49
2:D:136:ARG:N	2:D:164:THR:O	2.42	0.48
1:A:14:LEU:HD21	1:A:128:ILE:HB	1.95	0.48
2:D:20:TYR:HE2	2:D:106:GLU:HG2	1.80	0.46
2:D:132:VAL:HG12	2:D:133:GLN:N	2.30	0.46
1:A:45:TYR:HD2	1:A:62:GLU:OE1	1.98	0.46
2:D:131:CYS:SG	2:D:137:ALA:HB2	2.56	0.46
2:D:30:CYS:O	2:D:114:ALA:HB2	2.15	0.46
1:A:56:LEU:HD11	1:A:124:PHE:CE2	2.52	0.45
2:D:19:ALA:O	2:D:125:GLN:NE2	2.21	0.45
1:A:69:VAL:HG11	1:A:114:ILE:HG12	1.98	0.45
2:D:110:TRP:NE1	2:D:113:GLU:HB3	2.33	0.44
2:D:136:ARG:O	2:D:163:CYS:HA	2.18	0.44
2:D:141:GLY:HA3	2:D:160:GLN:HB2	2.00	0.44
1:A:49:LYS:HA	1:A:49:LYS:HD2	1.67	0.44
1:A:89:ILE:O	1:A:93:VAL:HG23	2.17	0.44
2:D:123:VAL:HA	2:D:147:CYS:HB3	2.00	0.44
1:A:132:LEU:O	1:A:134:ALA:N	2.50	0.44
1:A:54:LYS:HB2	1:A:103:PHE:CD2	2.53	0.44
2:D:55:TRP:NE1	2:D:104:CYS:HB3	2.33	0.43
2:D:2:LEU:HA	2:D:122:VAL:HA	1.99	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:22:GLU:HA	2:D:46:CYS:SG	2.58	0.43
2:D:137:ALA:C	2:D:138:LEU:HD12	2.38	0.43
1:A:87:SER:O	1:A:91:VAL:HG23	2.18	0.43
1:A:52:GLU:H	1:A:55:HIS:CE1	2.37	0.43
2:D:148:LYS:NZ	2:D:155:ARG:HD2	2.34	0.43
2:D:129:TYR:CE1	2:D:159:PRO:HG3	2.54	0.42
1:A:65:PRO:HG3	2:D:40:GLY:O	2.21	0.41
2:D:15:PHE:CD1	2:D:15:PHE:N	2.88	0.41
1:A:75:SER:C	1:A:76:LYS:HG3	2.42	0.40

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	123/146 (84%)	106 (86%)	11 (9%)	6 (5%)	2 17
2	D	119/258 (46%)	109 (92%)	9 (8%)	1 (1%)	19 58
All	All	242/404 (60%)	215 (89%)	20 (8%)	7 (3%)	4 28

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	76	LYS
1	A	80	LEU
1	A	133	THR
1	A	75	SER
1	A	79	HIS
2	D	22	GLU
1	A	134	ALA

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	119/134 (89%)	117 (98%)	2 (2%)	60 83
2	D	109/222 (49%)	106 (97%)	3 (3%)	43 74
All	All	228/356 (64%)	223 (98%)	5 (2%)	52 79

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

Mol	Chain	Res	Type
1	A	76	LYS
1	A	79	HIS
2	D	27	ASN
2	D	106	GLU
2	D	139	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	D	27	ASN
2	D	120	HIS

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 (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, 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	127/146 (86%)	0.63	9 (7%) 16 9	93, 128, 219, 268	0
2	D	125/258 (48%)	0.21	4 (3%) 47 31	101, 154, 199, 239	0
All	All	252/404 (62%)	0.42	13 (5%) 27 15	93, 138, 207, 268	0

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	103	PHE	4.0
2	D	140	ARG	3.8
1	A	21	LEU	2.9
1	A	89	ILE	2.5
1	A	17	LEU	2.5
1	A	88	ASN	2.4
1	A	72	LEU	2.4
2	D	26	LEU	2.4
2	D	19	ALA	2.3
1	A	24	ILE	2.3
2	D	25	MET	2.1
1	A	121	TRP	2.1
1	A	28	ILE	2.0

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.