



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 28, 2026 – 09:21 PM UTC

PDB ID : 5DFW / pdb_00005dfw
Title : CRYSTAL STRUCTURE OF HUMAN CD81 LARGE EXTRACELLULAR LOOP IN COMPLEX WITH SINGLE CHAIN FV FRAGMENT K13
Authors : Harris, S.F.; Villasenor, A.; Kuglstatter, A.
Deposited on : 2015-08-27
Resolution : 2.33 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4-5-2 with Phenix2.0
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

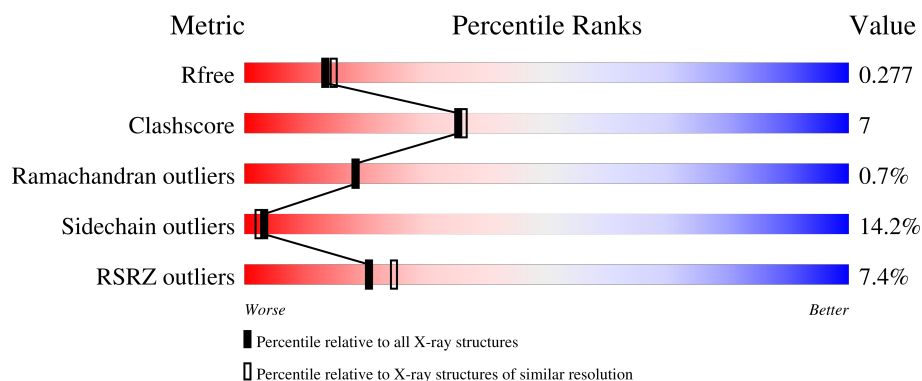
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.33 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}	180053	3031 (2.36-2.32)
Clashscore	190562	3127 (2.36-2.32)
Ramachandran outliers	187476	3095 (2.36-2.32)
Sidechain outliers	187428	3095 (2.36-2.32)
RSRZ outliers	180081	3033 (2.36-2.32)

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 $\leq 5\%$. 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	98	2% 65% 23% 8%
2	H	243	9% 64% 22% 5% 9%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2501 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 CD81 antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	90	Total	C	N	O	S	0	0	0
			693	430	118	141	4			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	110	GLY	-	expression tag	UNP P60033
A	111	SER	-	expression tag	UNP P60033
A	202	HIS	-	expression tag	UNP P60033
A	203	HIS	-	expression tag	UNP P60033
A	204	HIS	-	expression tag	UNP P60033
A	205	HIS	-	expression tag	UNP P60033
A	206	HIS	-	expression tag	UNP P60033
A	207	HIS	-	expression tag	UNP P60033

- Molecule 2 is a protein called SINGLE CHAIN FV FRAGMENT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	222	Total	C	N	O	S	0	0	0
			1725	1095	291	333	6			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	28	Total	O	0	0
			28	28		
3	H	55	Total	O	0	0
			55	55		

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	49.65Å 49.67Å 80.45Å 90.00° 105.30° 90.00°	Depositor
Resolution (Å)	41.83 – 2.33 41.83 – 2.33	Depositor EDS
% Data completeness (in resolution range)	68.2 (41.83-2.33) 68.3 (41.83-2.33)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.60 (at 2.34Å)	Xtriage
Refinement program	BUSTER 2.9.2	Depositor
R, R_{free}	0.190 , 0.264 0.194 , 0.277	Depositor DCC
R_{free} test set	577 reflections (5.14%)	wwPDB-VP
Wilson B-factor (Å ²)	29.9	Xtriage
Anisotropy	0.542	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 55.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.027 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	2501	wwPDB-VP
Average B, all atoms (Å ²)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.32% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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

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	0.86	0/702	1.56	8/948 (0.8%)
2	H	0.78	0/1771	1.24	13/2410 (0.5%)
All	All	0.81	0/2473	1.34	21/3358 (0.6%)

There are no bond length outliers.

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	137	ASP	CA-CB-CG	7.33	119.94	112.60
1	A	189	ASP	CA-CB-CG	7.17	119.77	112.60
2	H	651	GLY	CA-C-N	-6.78	110.69	122.36
2	H	651	GLY	C-N-CA	-6.78	110.69	122.36
2	H	807	VAL	N-CA-CB	6.36	118.62	110.99
2	H	703	PRO	CA-C-N	5.66	128.64	120.38
2	H	703	PRO	C-N-CA	5.66	128.64	120.38
2	H	408	THR	CB-CA-C	5.61	119.46	110.16
1	A	145	ALA	CA-C-N	5.54	127.65	120.56
1	A	145	ALA	C-N-CA	5.54	127.65	120.56
2	H	597	THR	CB-CA-C	5.42	120.25	112.12
1	A	176	PRO	CA-C-N	5.38	128.35	120.71
1	A	176	PRO	C-N-CA	5.38	128.35	120.71
2	H	398	HIS	N-CA-C	5.29	118.21	109.85
2	H	403	ASP	CA-CB-CG	5.27	117.87	112.60
1	A	162	LEU	CA-C-N	5.19	127.75	120.28
1	A	162	LEU	C-N-CA	5.19	127.75	120.28
2	H	796	THR	CB-CA-C	5.17	117.43	109.15
2	H	753	PHE	CA-CB-CG	5.10	118.90	113.80
2	H	103	ARG	CA-C-N	5.02	127.89	120.82
2	H	103	ARG	C-N-CA	5.02	127.89	120.82

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	693	0	672	9	0
2	H	1725	0	1652	30	0
3	A	28	0	0	0	0
3	H	55	0	0	3	0
All	All	2501	0	2324	35	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 7.

All (35) 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:173:ASN:HD21	2:H:396:GLY:H	1.22	0.85
2:H:199:HIS:HD2	2:H:212:TRP:HE1	1.23	0.84
2:H:104:LEU:HD21	2:H:332:PRO:HD3	1.63	0.79
2:H:398:HIS:NE2	3:H:901:HOH:O	2.21	0.64
2:H:199:HIS:CD2	2:H:212:TRP:HE1	2.12	0.63
2:H:353:TYR:N	3:H:901:HOH:O	2.32	0.63
2:H:303:THR:HB	2:H:316:HIS:HB2	1.83	0.60
2:H:203:LYS:HB2	2:H:213:ILE:HD11	1.83	0.59
2:H:521:ILE:HG12	2:H:805:THR:HG21	1.86	0.57
2:H:603:LEU:HB2	2:H:613:LEU:HD11	1.84	0.57
2:H:562:ILE:O	2:H:597:THR:HG22	2.06	0.56
1:A:201:LYS:O	1:A:202:HIS:ND1	2.42	0.51
2:H:104:LEU:H	2:H:104:LEU:HD12	1.76	0.49
2:H:305:THR:OG1	2:H:316:HIS:HE1	1.96	0.48
1:A:184:ASN:HD22	1:A:187:LYS:HD2	1.79	0.47
2:H:601:TRP:HD1	2:H:614:VAL:HG22	1.79	0.47
2:H:597:THR:HG21	2:H:754:TRP:HE3	1.80	0.47
1:A:170:LEU:O	2:H:352:GLY:O	2.33	0.46
2:H:352:GLY:N	3:H:901:HOH:O	2.47	0.46
2:H:106:GLN:HE21	2:H:402:GLY:HA3	1.80	0.46
2:H:551:ARG:HG3	2:H:716:GLN:HG2	1.98	0.45
2:H:703:PRO:HB2	2:H:705:ARG:HG2	1.98	0.45
2:H:710:GLY:HA3	2:H:717:TYR:HA	1.98	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:199:HIS:CD2	2:H:251:ARG:HB2	2.52	0.44
1:A:196:ASP:HB3	1:A:201:LYS:O	2.18	0.44
2:H:208:GLN:HE21	2:H:209:GLY:H	1.65	0.43
1:A:171:LYS:O	2:H:352:GLY:HA2	2.18	0.43
2:H:523:CYS:HB2	2:H:601:TRP:CH2	2.53	0.43
1:A:135:VAL:HG21	1:A:165:LEU:HD22	1.99	0.43
2:H:511:LEU:HB3	2:H:807:VAL:HG12	2.01	0.43
1:A:173:ASN:ND2	2:H:352:GLY:O	2.52	0.42
2:H:729:PHE:CZ	2:H:809:ILE:HG23	2.54	0.42
2:H:294:ASP:HB3	2:H:297:PHE:HD1	1.85	0.42
2:H:504:MET:HE3	2:H:752:HIS:CD2	2.55	0.42
1:A:131:LEU:HD11	1:A:169:VAL:HG22	2.01	0.41

There are no symmetry-related clashes.

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	88/98 (90%)	86 (98%)	2 (2%)	0	100	100
2	H	218/243 (90%)	206 (94%)	10 (5%)	2 (1%)	14	13
All	All	306/341 (90%)	292 (95%)	12 (4%)	2 (1%)	18	18

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	H	712	GLY
2	H	125	SER

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	80/86 (93%)	68 (85%)	12 (15%)	3	2
2	H	188/193 (97%)	162 (86%)	26 (14%)	3	2
All	All	268/279 (96%)	230 (86%)	38 (14%)	3	2

All (38) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	147	VAL
1	A	148	LYS
1	A	161	THR
1	A	163	THR
1	A	168	SER
1	A	181	ILE
1	A	185	LEU
1	A	188	GLU
1	A	189	ASP
1	A	193	LYS
1	A	201	LYS
1	A	202	HIS
2	H	111	LEU
2	H	112	VAL
2	H	113	GLN
2	H	125	SER
2	H	208	GLN
2	H	251	ARG
2	H	292	LYS
2	H	296	LYS
2	H	301	LYS
2	H	408	THR
2	H	411	SER
2	H	511	LEU
2	H	519	VAL
2	H	603	LEU
2	H	611	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	H	614	VAL
2	H	695	THR
2	H	696	THR
2	H	697	LEU
2	H	711	SER
2	H	715	THR
2	H	721	ILE
2	H	724	LEU
2	H	796	THR
2	H	808	GLU
2	H	809	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	118	GLN
1	A	133	GLN
1	A	172	ASN
1	A	173	ASN
1	A	184	ASN
2	H	105	HIS
2	H	106	GLN
2	H	199	HIS
2	H	208	GLN
2	H	316	HIS
2	H	733	HIS
2	H	752	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 oligosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

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, 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	90/98 (91%)	0.13	2 (2%) 62 67	23, 36, 54, 67	0
2	H	222/243 (91%)	0.74	21 (9%) 14 17	24, 49, 75, 107	0
All	All	312/341 (91%)	0.57	23 (7%) 20 25	23, 45, 73, 107	0

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	102	VAL	4.8
2	H	809	ILE	3.8
2	H	207	ALA	3.7
1	A	160	SER	3.6
2	H	105	HIS	3.0
2	H	104	LEU	3.0
2	H	515	VAL	3.0
2	H	807	VAL	2.9
2	H	729	PHE	2.8
1	A	202	HIS	2.7
2	H	810	LYS	2.7
2	H	513	VAL	2.6
2	H	408	THR	2.5
2	H	206	PRO	2.5
2	H	503	GLN	2.5
2	H	519	VAL	2.4
2	H	808	GLU	2.4
2	H	411	SER	2.3
2	H	651	GLY	2.3
2	H	520	THR	2.2
2	H	726	SER	2.2
2	H	103	ARG	2.1
2	H	511	LEU	2.0

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 oligosaccharides 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.