Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 08:21 am BST

PDB ID	:	6EK2
Title	:	CRYSTAL STRUCTURE OF HUMAN CD81 LARGE EXTRACELLULAR LOOP IN COMPLEX WITH SINGLE CHAIN FV FRAGMENT 10
Authors	:	Harris, S.F.; Villasenor, A.; Kuglstatter, A.
Deposited on	:	2017-09-25
Resolution	:	2.65 Å(reported)

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

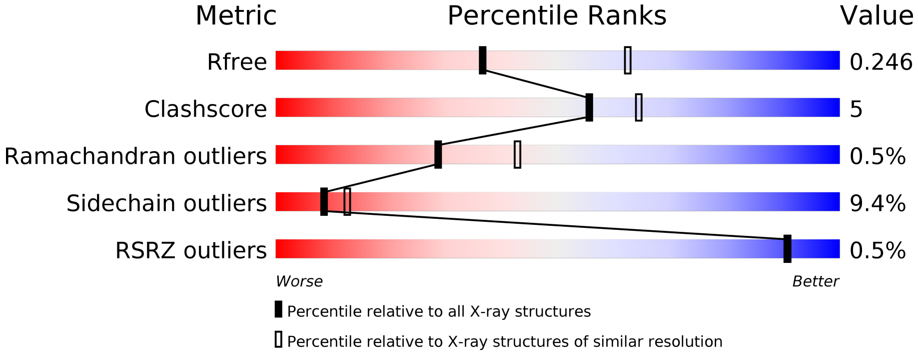
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance

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

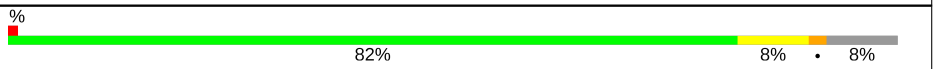
The reported resolution of this entry is 2.65 Å.

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	1332 (2.68-2.64)
Clashscore	141614	1374 (2.68-2.64)
Ramachandran outliers	138981	1349 (2.68-2.64)
Sidechain outliers	138945	1349 (2.68-2.64)
RSRZ outliers	127900	1318 (2.68-2.64)

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 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	98	
1	B	98	
2	H	247	
2	I	247	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5162 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
			Total	C	N	O	S			
1	A	90	693	430	118	141	4	0	0	0
1	B	89	683	424	115	140	4	0	0	0

There are 16 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
B	110	GLY	-	expression tag	UNP P60033
B	111	SER	-	expression tag	UNP P60033
B	202	HIS	-	expression tag	UNP P60033
B	203	HIS	-	expression tag	UNP P60033
B	204	HIS	-	expression tag	UNP P60033
B	205	HIS	-	expression tag	UNP P60033
B	206	HIS	-	expression tag	UNP P60033
B	207	HIS	-	expression tag	UNP P60033

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	H	227	1767	1120	293	346	8	0	0	0
2	I	226	1762	1117	292	345	8	0	0	0

- Molecule 3 is water.

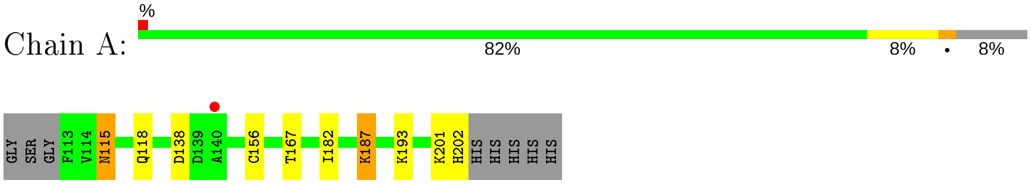
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	39	Total O 39 39	0	0
3	B	29	Total O 29 29	0	0
3	H	118	Total O 118 118	0	0
3	I	71	Total O 71 71	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA 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.

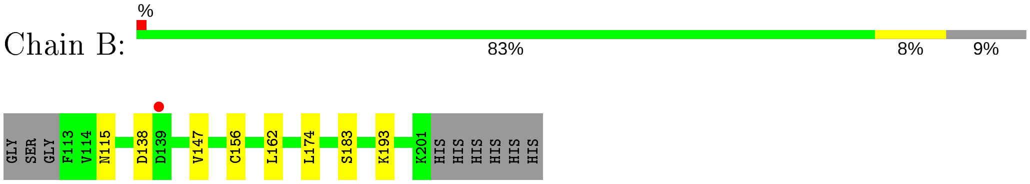
- Molecule 1: CD81 antigen

Chain A:



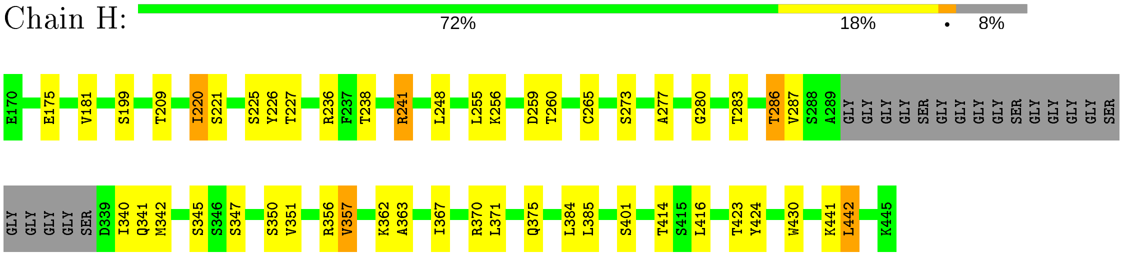
- Molecule 1: CD81 antigen

Chain B:



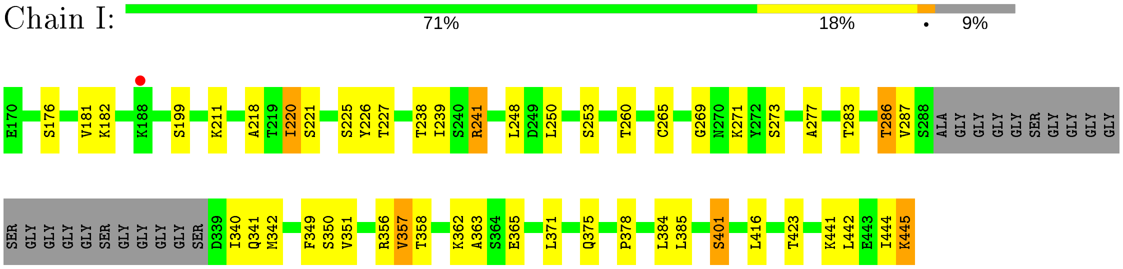
- Molecule 2: SINGLE CHAIN FV FRAGMENT

Chain H:



- Molecule 2: SINGLE CHAIN FV FRAGMENT

Chain I:



4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	52.32Å 118.06Å 130.13Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.72 – 2.65 43.72 – 2.65	Depositor EDS
% Data completeness (in resolution range)	72.6 (43.72-2.65) 73.2 (43.72-2.65)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.18	Depositor
$< I/\sigma(I) >$ [1]	1.72 (at 2.65Å)	Xtriage
Refinement program	BUSTER 2.9.2	Depositor
R, R_{free}	0.179 , 0.247 0.185 , 0.246	Depositor DCC
R_{free} test set	933 reflections (5.29%)	wwPDB-VP
Wilson B-factor (Å^2)	42.0	Xtriage
Anisotropy	0.166	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 51.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	5162	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 9.09% 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	0.50	0/702	0.74	0/948
1	B	0.50	0/691	0.74	0/933
2	H	0.48	0/1810	0.77	1/2449 (0.0%)
2	I	0.46	0/1805	0.77	1/2442 (0.0%)
All	All	0.48	0/5008	0.76	2/6772 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	I	273	SER	N-CA-C	-5.67	95.68	111.00
2	H	273	SER	N-CA-C	-5.26	96.81	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	693	0	672	4	0
1	B	683	0	665	4	0
2	H	1767	0	1698	19	0
2	I	1762	0	1693	24	0
3	A	39	0	0	0	0
3	B	29	0	0	0	0
3	H	118	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	I	71	0	0	2	0
All	All	5162	0	4728	49	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 5.

All (49) 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:H:236:ARG:NH1	2:H:259:ASP:OD2	2.19	0.73
2:I:220:ILE:HB	2:I:239:ILE:HG21	1.74	0.70
2:I:220:ILE:HB	2:I:239:ILE:CG2	2.28	0.63
2:H:181:VAL:HG11	2:H:255:LEU:HD13	1.80	0.63
2:I:378:PRO:HD3	3:I:508:HOH:O	2.01	0.60
1:B:156:CYS:HB3	1:B:193:LYS:HG3	1.87	0.57
2:I:260:THR:HG23	2:I:286:THR:HA	1.86	0.57
1:A:156:CYS:HB3	1:A:193:LYS:HG3	1.87	0.56
2:H:260:THR:HG23	2:H:286:THR:HA	1.86	0.56
1:B:162:LEU:HD11	2:I:271:LYS:HG2	1.88	0.56
1:B:162:LEU:CD1	2:I:271:LYS:HG2	2.37	0.54
2:H:220:ILE:HG12	2:H:241:ARG:HG2	1.91	0.53
2:I:220:ILE:HG12	2:I:241:ARG:HG2	1.92	0.52
2:I:239:ILE:HD11	2:I:250:LEU:HD13	1.91	0.51
2:I:342:MET:HE2	2:I:363:ALA:HA	1.93	0.51
2:I:218:ALA:HB1	2:I:239:ILE:HG13	1.93	0.49
2:I:340:ILE:HG22	2:I:342:MET:HE3	1.93	0.49
2:H:424:TYR:CE1	2:H:442:LEU:HD22	2.48	0.49
1:A:201:LYS:O	1:A:202:HIS:ND1	2.46	0.48
2:H:256:LYS:O	2:H:287:VAL:HG11	2.13	0.48
2:I:375:GLN:HB2	2:I:385:LEU:HD11	1.96	0.48
2:H:375:GLN:HB2	2:H:385:LEU:HD11	1.96	0.48
2:H:340:ILE:HG22	2:H:342:MET:HE3	1.96	0.47
2:H:357:VAL:HG22	2:H:416:LEU:HD11	1.97	0.46
2:H:236:ARG:HH12	2:H:259:ASP:CG	2.16	0.46
2:I:444:ILE:HG13	2:I:445:LYS:N	2.31	0.46
2:I:221:SER:HB3	2:I:226:TYR:HB2	1.98	0.45
2:I:423:THR:HA	2:I:441:LYS:HA	1.99	0.45
2:H:370:ARG:HD3	2:H:430:TRP:HD1	1.81	0.45
2:I:340:ILE:HG22	2:I:342:MET:CE	2.47	0.45
2:H:340:ILE:HG22	2:H:342:MET:CE	2.47	0.44
2:I:349:PHE:HB2	2:I:442:LEU:HD23	2.00	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:221:SER:HB3	2:H:226:TYR:HB2	2.00	0.43
2:I:181:VAL:HG13	2:I:287:VAL:HB	2.01	0.43
2:I:357:VAL:HG22	2:I:416:LEU:HD11	2.01	0.43
2:H:423:THR:HA	2:H:441:LYS:HA	2.00	0.43
2:H:277:ALA:HA	2:H:384:LEU:HD22	1.99	0.43
2:I:220:ILE:HD13	2:I:239:ILE:HG22	2.00	0.42
2:I:277:ALA:HA	2:I:384:LEU:HD22	2.02	0.42
1:A:182:ILE:HD13	1:A:187:LYS:HE2	2.02	0.42
2:H:342:MET:HE2	2:H:363:ALA:HA	2.01	0.41
2:I:401:SER:HB3	3:I:505:HOH:O	2.19	0.41
2:H:175:GLU:OE2	2:H:280:GLY:HA3	2.20	0.41
1:B:147:VAL:HG11	1:B:174:LEU:HD13	2.01	0.41
2:I:269:GLY:HA2	2:I:277:ALA:HB2	2.03	0.41
2:H:370:ARG:HD3	2:H:430:TRP:CD1	2.56	0.41
2:I:351:VAL:HG11	2:I:357:VAL:HG13	2.03	0.41
1:A:115:ASN:ND2	1:A:118:GLN:H	2.19	0.41
2:H:342:MET:HE1	2:H:367:ILE:HD11	2.03	0.40

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%)	83 (94%)	5 (6%)	0	100	100
1	B	87/98 (89%)	79 (91%)	6 (7%)	2 (2%)	6	8
2	H	223/247 (90%)	212 (95%)	10 (4%)	1 (0%)	34	48
2	I	222/247 (90%)	209 (94%)	13 (6%)	0	100	100
All	All	620/690 (90%)	583 (94%)	34 (6%)	3 (0%)	29	43

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	138	ASP
1	B	183	SER
2	H	345	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%)	76 (95%)	4 (5%)	24	38
1	B	79/86 (92%)	78 (99%)	1 (1%)	69	82
2	H	191/195 (98%)	169 (88%)	22 (12%)	5	7
2	I	191/195 (98%)	167 (87%)	24 (13%)	4	6
All	All	541/562 (96%)	490 (91%)	51 (9%)	8	13

All (51) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	115	ASN
1	A	138	ASP
1	A	167	THR
1	A	187	LYS
1	B	115	ASN
2	H	199	SER
2	H	209	THR
2	H	220	ILE
2	H	225	SER
2	H	227	THR
2	H	238	THR
2	H	241	ARG
2	H	248	LEU
2	H	265	CYS
2	H	283	THR
2	H	286	THR
2	H	341	GLN
2	H	347	SER
2	H	350	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	H	351	VAL
2	H	356	ARG
2	H	357	VAL
2	H	362	LYS
2	H	371	LEU
2	H	401	SER
2	H	414	THR
2	H	442	LEU
2	I	176	SER
2	I	182	LYS
2	I	199	SER
2	I	211	LYS
2	I	220	ILE
2	I	225	SER
2	I	227	THR
2	I	238	THR
2	I	241	ARG
2	I	248	LEU
2	I	253	SER
2	I	265	CYS
2	I	283	THR
2	I	286	THR
2	I	341	GLN
2	I	350	SER
2	I	356	ARG
2	I	357	VAL
2	I	358	THR
2	I	362	LYS
2	I	365	GLU
2	I	371	LEU
2	I	401	SER
2	I	445	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	115	ASN
1	B	115	ASN
2	H	243	ASN
2	H	251	GLN
2	H	270	ASN
2	H	341	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	I	243	ASN
2	I	251	GLN
2	I	341	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 carbohydrates 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 (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, 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%)	-0.35	1 (1%)	80	79	21, 36, 84, 96	0
1	B	89/98 (90%)	-0.41	1 (1%)	80	79	24, 38, 80, 100	0
2	H	227/247 (91%)	-0.52	0	100	100	22, 38, 57, 85	0
2	I	226/247 (91%)	-0.29	1 (0%)	92	93	29, 55, 75, 96	0
All	All	632/690 (91%)	-0.40	3 (0%)	91	91	21, 43, 75, 100	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	139	ASP	3.2
2	I	188	LYS	2.2
1	A	140	ALA	2.1

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 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.