



wwPDB X-ray Structure Validation Summary Report ⓘ

May 26, 2020 – 09:16 pm BST

PDB ID : 4Q6I
Title : Crystal structure of murine 2D5 Fab, a potent anti-CD4 HIV-1-neutralizing antibody in complex with CD4
Authors : Boyington, J.C.; Nabel, G.J.; Mascola, J.R.
Deposited on : 2014-04-22
Resolution : 3.65 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 ⓘ symbol.

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

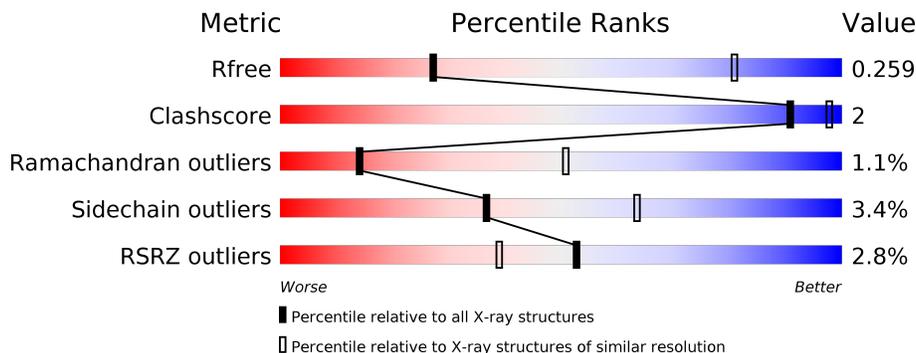
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.65 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	1557 (3.82-3.50)
Clashscore	141614	1037 (3.80-3.52)
Ramachandran outliers	138981	1004 (3.80-3.52)
Sidechain outliers	138945	1002 (3.80-3.52)
RSRZ outliers	127900	1441 (3.82-3.50)

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 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 $\leq 5\%$. 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	218	
1	D	218	
1	F	218	
1	L	218	
2	B	230	
2	E	230	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	230	 2% 87% 7%
2	H	230	 2% 87% 8%
3	C	208	 2% 79% 9% 13%
3	I	208	 2% 77% 11% 13%
3	J	208	 2% 79% 8% 13%
3	K	208	 2% 77% 11% 13%

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 37900 atoms, of which 18708 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 Light chain of murine 2D5 Fab.

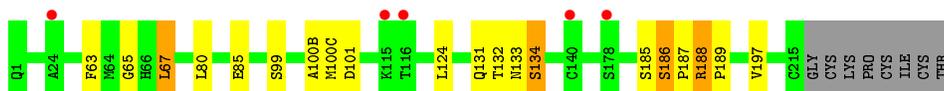
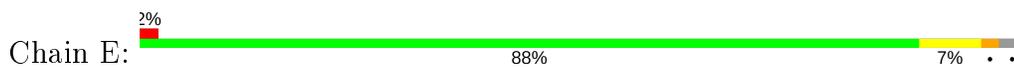
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	218	Total 3294	C 1054	H 1598	N 287	O 346	S 9	0	0	0
1	D	218	Total 3293	C 1054	H 1597	N 287	O 346	S 9	0	0	0
1	F	218	Total 3292	C 1054	H 1596	N 287	O 346	S 9	0	0	0
1	L	218	Total 3294	C 1054	H 1598	N 287	O 346	S 9	0	0	0

- Molecule 2 is a protein called Heavy chain of murine 2D5 Fab.

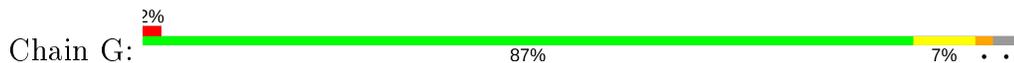
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	B	222	Total 3315	C 1066	H 1632	N 276	O 332	S 9	0	0	0
2	E	222	Total 3315	C 1066	H 1632	N 276	O 332	S 9	0	0	0
2	G	222	Total 3325	C 1068	H 1638	N 277	O 333	S 9	0	0	0
2	H	222	Total 3326	C 1068	H 1639	N 277	O 333	S 9	0	0	0

- Molecule 3 is a protein called T-cell surface glycoprotein CD4.

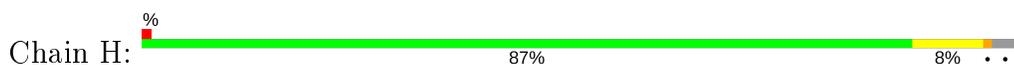
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
3	C	182	Total 2862	C 888	H 1445	N 248	O 277	S 4	0	0	0
3	I	182	Total 2862	C 888	H 1445	N 248	O 277	S 4	0	0	0
3	J	182	Total 2862	C 888	H 1445	N 248	O 277	S 4	0	0	0
3	K	182	Total 2860	C 888	H 1443	N 248	O 277	S 4	0	0	0



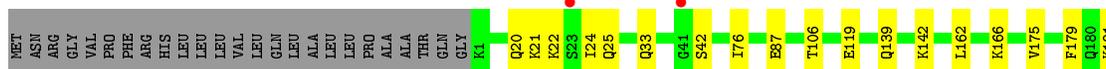
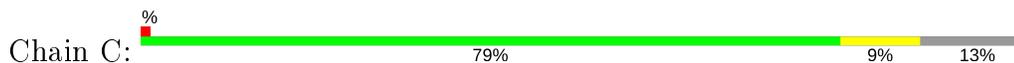
- Molecule 2: Heavy chain of murine 2D5 Fab



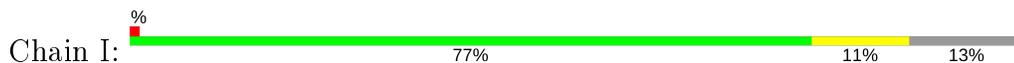
- Molecule 2: Heavy chain of murine 2D5 Fab



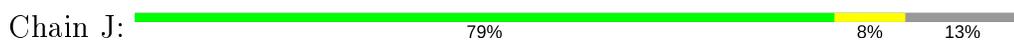
- Molecule 3: T-cell surface glycoprotein CD4



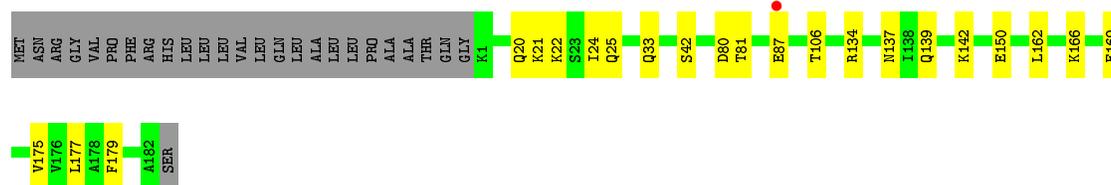
- Molecule 3: T-cell surface glycoprotein CD4



- Molecule 3: T-cell surface glycoprotein CD4



- Molecule 3: T-cell surface glycoprotein CD4



4 Data and refinement statistics

Property	Value	Source
Space group	P 4	Depositor
Cell constants a, b, c, α , β , γ	149.55Å 149.55Å 183.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.91 – 3.65 30.02 – 3.65	Depositor EDS
% Data completeness (in resolution range)	96.8 (29.91-3.65) 96.9 (30.02-3.65)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.35 (at 3.65Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, R_{free}	0.226 , 0.257 0.237 , 0.259	Depositor DCC
R_{free} test set	2210 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	129.9	Xtrriage
Anisotropy	0.298	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 80.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.024 for h,-k,-l	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	37900	wwPDB-VP
Average B, all atoms (Å ²)	146.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.02% 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 [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	0.23	0/1737	0.41	0/2359
1	D	0.23	0/1737	0.41	0/2359
1	F	0.24	0/1737	0.41	0/2359
1	L	0.23	0/1737	0.41	0/2359
2	B	0.23	0/1728	0.41	0/2359
2	E	0.23	0/1728	0.42	0/2359
2	G	0.23	0/1732	0.41	0/2364
2	H	0.24	0/1732	0.42	0/2364
3	C	0.24	0/1437	0.44	0/1937
3	I	0.24	0/1437	0.44	0/1937
3	J	0.23	0/1437	0.44	0/1937
3	K	0.24	0/1437	0.44	0/1937
All	All	0.23	0/19616	0.42	0/26630

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	1696	1598	1609	1	0
1	D	1696	1597	1610	5	0
1	F	1696	1596	1610	7	0
1	L	1696	1598	1610	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1683	1632	1637	4	0
2	E	1683	1632	1638	9	0
2	G	1687	1638	1644	9	0
2	H	1687	1639	1644	8	0
3	C	1417	1445	1449	4	0
3	I	1417	1445	1449	7	0
3	J	1417	1445	1449	4	0
3	K	1417	1443	1449	5	0
All	All	19192	18708	18798	64	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 2.

The worst 5 of 64 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:E:134:SER:O	2:E:185:SER:N	2.21	0.74
1:A:27(D):THR:O	1:A:29:ARG:N	2.22	0.72
1:D:27(D):THR:O	1:D:29:ARG:N	2.23	0.71
1:L:27(D):THR:O	1:L:29:ARG:N	2.23	0.70
1:F:27(D):THR:O	1:F:29:ARG:N	2.24	0.70

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	216/218 (99%)	202 (94%)	13 (6%)	1 (0%)	29 66
1	D	216/218 (99%)	204 (94%)	11 (5%)	1 (0%)	29 66
1	F	216/218 (99%)	203 (94%)	12 (6%)	1 (0%)	29 66

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	L	216/218 (99%)	202 (94%)	13 (6%)	1 (0%)	29	66
2	B	220/230 (96%)	196 (89%)	21 (10%)	3 (1%)	11	45
2	E	220/230 (96%)	196 (89%)	20 (9%)	4 (2%)	8	41
2	G	220/230 (96%)	196 (89%)	20 (9%)	4 (2%)	8	41
2	H	220/230 (96%)	196 (89%)	20 (9%)	4 (2%)	8	41
3	C	180/208 (86%)	159 (88%)	19 (11%)	2 (1%)	14	51
3	I	180/208 (86%)	157 (87%)	21 (12%)	2 (1%)	14	51
3	J	180/208 (86%)	159 (88%)	19 (11%)	2 (1%)	14	51
3	K	180/208 (86%)	158 (88%)	20 (11%)	2 (1%)	14	51
All	All	2464/2624 (94%)	2228 (90%)	209 (8%)	27 (1%)	14	51

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	28	SER
2	B	100(C)	MET
3	C	21	LYS
1	D	28	SER
2	E	100(C)	MET

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	192/192 (100%)	190 (99%)	2 (1%)	76	86
1	D	192/192 (100%)	188 (98%)	4 (2%)	53	73
1	F	192/192 (100%)	189 (98%)	3 (2%)	62	79
1	L	192/192 (100%)	188 (98%)	4 (2%)	53	73
2	B	189/197 (96%)	182 (96%)	7 (4%)	34	61
2	E	189/197 (96%)	183 (97%)	6 (3%)	39	64
2	G	190/197 (96%)	184 (97%)	6 (3%)	39	64

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	H	190/197 (96%)	184 (97%)	6 (3%)	39	64
3	C	164/185 (89%)	155 (94%)	9 (6%)	21	53
3	I	164/185 (89%)	155 (94%)	9 (6%)	21	53
3	J	164/185 (89%)	156 (95%)	8 (5%)	25	56
3	K	164/185 (89%)	153 (93%)	11 (7%)	16	48
All	All	2182/2296 (95%)	2107 (97%)	75 (3%)	37	63

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	G	124	LEU
2	H	190	SER
3	K	177	LEU
2	G	186	SER
2	H	99	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
2	E	164	HIS
1	L	198	HIS
1	F	198	HIS
3	C	107	HIS
2	G	164	HIS

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

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 [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	218/218 (100%)	0.62	25 (11%) 4 3	124, 199, 336, 490	0
1	D	218/218 (100%)	0.19	6 (2%) 53 39	113, 179, 275, 319	0
1	F	218/218 (100%)	0.21	5 (2%) 60 46	92, 184, 297, 351	0
1	L	218/218 (100%)	0.13	2 (0%) 84 74	84, 159, 254, 372	0
2	B	222/230 (96%)	0.38	13 (5%) 22 14	107, 207, 298, 382	0
2	E	222/230 (96%)	0.05	5 (2%) 60 46	106, 166, 245, 420	0
2	G	222/230 (96%)	0.16	5 (2%) 60 46	79, 167, 244, 394	0
2	H	222/230 (96%)	-0.00	3 (1%) 75 63	79, 143, 220, 391	0
3	C	182/208 (87%)	0.02	2 (1%) 80 70	101, 137, 213, 238	0
3	I	182/208 (87%)	0.08	2 (1%) 80 70	93, 134, 213, 275	0
3	J	182/208 (87%)	0.04	1 (0%) 91 85	119, 163, 230, 314	0
3	K	182/208 (87%)	0.09	1 (0%) 91 85	124, 160, 224, 267	0
All	All	2488/2624 (94%)	0.17	70 (2%) 53 39	79, 163, 278, 490	0

The worst 5 of 70 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	158	GLY	6.6
1	A	152	GLY	5.9
1	A	191	SER	5.7
1	A	134	CYS	5.1
1	A	118	PHE	4.9

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.