Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 07:19 PM GMT

PDB ID	:	4OJF
Title	:	Humanised 3D6 Fab complexed to amyloid beta 1-8
Authors	:	Miles, L.A.; Crespi, G.A.N.; Parker, M.W.
Deposited on	:	2014-01-21
Resolution	:	2.00 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

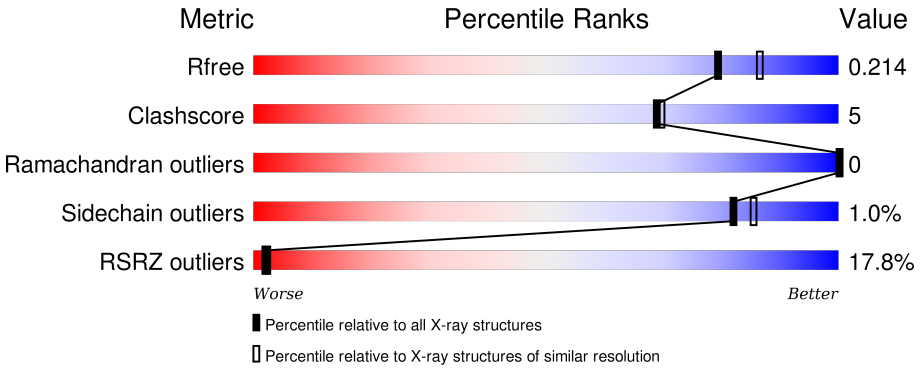
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.00 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}	91344	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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. 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	H	227	18% 89% 7%
2	L	219	17% 88% 12%
3	A	8	13% 75% 25%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 6984 atoms, of which 3304 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 Humanised 3D6 Fab Heavy Chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	H	219	Total	C	H	N	O	S	0	0	0
			3234	1027	1599	281	321	6			

- Molecule 2 is a protein called Humanised 3D6 Fab Light Chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	L	219	Total	C	H	N	O	S	0	0	0
			3347	1062	1658	283	338	6			

- Molecule 3 is a protein called Amyloid beta A4 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	6	Total	C	H	N	O	0	0	0
			101	33	47	11	10			

- Molecule 4 is water.

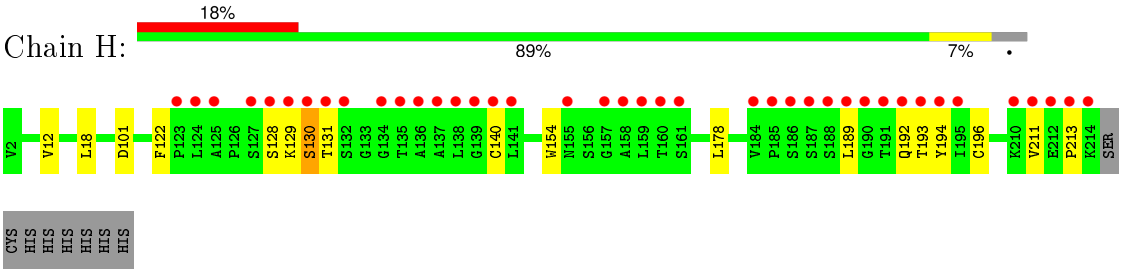
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	H	158	Total	O	0	0
			158	158		
4	L	141	Total	O	0	0
			141	141		
4	A	3	Total	O	0	0
			3	3		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

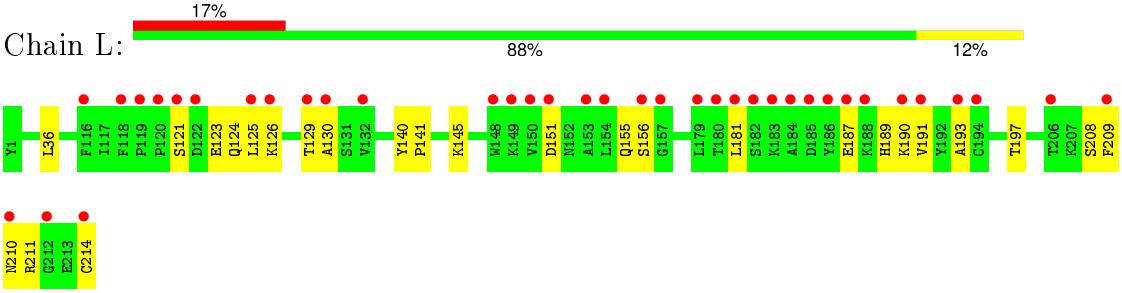
- Molecule 1: Humanised 3D6 Fab Heavy Chain

Chain H:



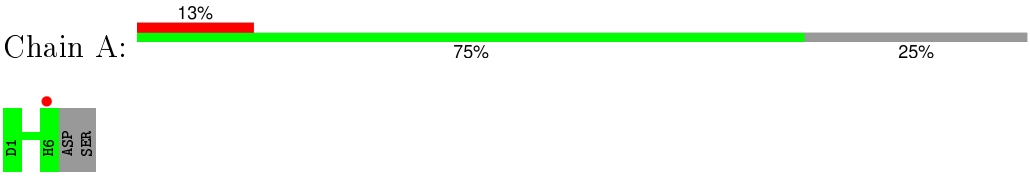
- Molecule 2: Humanised 3D6 Fab Light Chain

Chain L:



- Molecule 3: Amyloid beta A4 protein

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 2 21	Depositor
Cell constants a, b, c, α, β, γ	60.16Å 83.43Å 88.25Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.70 – 2.00 42.70 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.8 (42.70-2.00) 98.4 (42.70-2.00)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.50 (at 2.00Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8_1069)	Depositor
R, R_{free}	0.172 , 0.215 0.171 , 0.214	Depositor DCC
R_{free} test set	1527 reflections (5.04%)	DCC
Wilson B-factor (Å^2)	17.2	Xtriage
Anisotropy	0.339	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 51.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 30738 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6984	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.10% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	H	0.42	0/1674	0.63	1/2276 (0.0%)
2	L	0.43	0/1727	0.60	0/2345
3	A	0.51	0/55	0.53	0/72
All	All	0.43	0/3456	0.61	1/4693 (0.0%)

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
1	H	0	1

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	H	101	ASP	CB-CG-OD1	5.99	123.69	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	H	130	SER	Peptide

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	H	1635	1599	1596	11	0
2	L	1689	1658	1655	21	0
3	A	54	47	46	0	0
4	A	3	0	0	0	0
4	H	158	0	0	1	0
4	L	141	0	0	1	0
All	All	3680	3304	3297	31	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 (31) 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:H:129:LYS:N	1:H:130:SER:OG	2.04	0.91
2:L:187:GLU:OE2	2:L:211:ARG:NH2	2.12	0.81
2:L:190:LYS:HA	2:L:211:ARG:HD3	1.64	0.77
2:L:189:HIS:O	2:L:211:ARG:NH1	2.22	0.73
2:L:189:HIS:O	2:L:211:ARG:HD2	1.89	0.73
2:L:189:HIS:O	2:L:211:ARG:CD	2.42	0.68
2:L:190:LYS:HA	2:L:211:ARG:CD	2.29	0.62
2:L:121:SER:O	2:L:125:LEU:HB2	2.04	0.58
2:L:151:ASP:HA	2:L:191:VAL:HB	1.86	0.58
1:H:192:GLN:OE1	1:H:193:THR:N	2.31	0.56
2:L:190:LYS:O	2:L:210:ASN:HA	2.05	0.56
1:H:192:GLN:HG3	1:H:194:TYR:CE1	2.42	0.55
2:L:187:GLU:HA	2:L:211:ARG:NH2	2.22	0.53
2:L:126:LYS:NZ	4:L:440:HOH:O	2.42	0.52
2:L:208:SER:O	2:L:209:PHE:CD1	2.65	0.50
2:L:36:LEU:C	2:L:36:LEU:HD12	2.34	0.47
1:H:211:VAL:O	4:H:440:HOH:O	2.20	0.47
2:L:140:TYR:CG	2:L:141:PRO:HA	2.50	0.46
1:H:128:SER:C	1:H:130:SER:OG	2.54	0.46
1:H:128:SER:HA	1:H:129:LYS:HA	1.65	0.45
2:L:145:LYS:HB2	2:L:197:THR:HB	1.98	0.45
2:L:155:GLN:O	2:L:156:SER:OG	2.30	0.44
2:L:193:ALA:HA	2:L:208:SER:OG	2.18	0.43
2:L:193:ALA:CB	2:L:208:SER:OG	2.66	0.43
2:L:123:GLU:O	2:L:126:LYS:HG2	2.18	0.43
1:H:122:PHE:CE2	2:L:124:GLN:HG3	2.54	0.42
2:L:130:ALA:N	2:L:181:LEU:O	2.42	0.42
1:H:12:VAL:CG2	1:H:18:LEU:HD13	2.48	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:H:178:LEU:C	1:H:178:LEU:HD12	2.41	0.41
1:H:189:LEU:HD11	1:H:213:PRO:HG2	2.01	0.41
1:H:154:TRP:CH2	1:H:196:CYS:HB3	2.55	0.41

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	H	217/227 (96%)	206 (95%)	11 (5%)	0	100	100
2	L	217/219 (99%)	211 (97%)	6 (3%)	0	100	100
3	A	4/8 (50%)	4 (100%)	0	0	100	100
All	All	438/454 (96%)	421 (96%)	17 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	H	182/190 (96%)	180 (99%)	2 (1%)	80	83
2	L	195/195 (100%)	193 (99%)	2 (1%)	82	85
3	A	5/7 (71%)	5 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	382/392 (97%)	378 (99%)	4 (1%)	82 85

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	131	THR
1	H	140	CYS
2	L	129	THR
2	L	214	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	H	148	GLN

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

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	H	219/227 (96%)	0.83	40 (18%) 2 2	6, 18, 71, 85	0
2	L	219/219 (100%)	0.80	38 (17%) 2 2	8, 22, 66, 91	0
3	A	6/8 (75%)	0.15	1 (16%) 2 3	8, 9, 19, 40	0
All	All	444/454 (97%)	0.80	79 (17%) 2 2	6, 20, 69, 91	0

All (79) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	L	214	CYS	13.3
2	L	154	LEU	7.9
2	L	210	ASN	7.7
2	L	125	LEU	7.5
1	H	128	SER	6.7
1	H	195	ILE	6.1
1	H	130	SER	6.1
1	H	189	LEU	6.0
1	H	193	THR	5.8
1	H	211	VAL	5.3
2	L	153	ALA	5.1
1	H	159	LEU	4.7
2	L	209	PHE	4.7
1	H	212	GLU	4.7
2	L	126	LYS	4.6
1	H	190	GLY	4.5
1	H	194	TYR	4.5
2	L	157	GLY	4.5
1	H	187	SER	4.4
1	H	138	LEU	4.4
1	H	192	GLN	4.3
2	L	156	SER	4.2
2	L	181	LEU	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	L	184	ALA	4.0
2	L	129	THR	3.9
1	H	129	LYS	3.9
1	H	127	SER	3.8
2	L	119	PRO	3.5
1	H	131	THR	3.4
1	H	186	SER	3.4
2	L	122	ASP	3.4
2	L	191	VAL	3.4
1	H	132	SER	3.3
2	L	121	SER	3.2
1	H	140	CYS	3.1
2	L	183	LYS	3.1
1	H	191	THR	3.0
1	H	124	LEU	2.9
2	L	187	GLU	2.9
3	A	6	HIS	2.8
1	H	139	GLY	2.8
2	L	132	VAL	2.8
1	H	158	ALA	2.8
1	H	184	VAL	2.8
1	H	125	ALA	2.7
1	H	123	PRO	2.7
1	H	213	PRO	2.7
1	H	210	LYS	2.6
2	L	190	LYS	2.6
2	L	179	LEU	2.6
2	L	150	VAL	2.6
1	H	214	LYS	2.5
1	H	157	GLY	2.5
2	L	151	ASP	2.5
2	L	130	ALA	2.4
1	H	135	THR	2.4
2	L	186	TYR	2.4
2	L	116	PHE	2.4
2	L	149	LYS	2.4
2	L	194	CYS	2.4
1	H	155	ASN	2.4
1	H	161	SER	2.4
1	H	188	SER	2.4
2	L	182	SER	2.4
2	L	185	ASP	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	H	185	PRO	2.3
1	H	134	GLY	2.3
2	L	118	PHE	2.3
2	L	212	GLY	2.2
1	H	137	ALA	2.2
1	H	160	THR	2.1
2	L	180	THR	2.1
2	L	148	TRP	2.1
2	L	206	THR	2.1
2	L	188	LYS	2.0
1	H	136	ALA	2.0
1	H	141	LEU	2.0
2	L	120	PRO	2.0
2	L	193	ALA	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 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.