wwPDB X-ray Structure Validation Summary Report (i)

Jun 1, 2023 – 10:26 am BST

PDB ID	:	8C7V
Title	:	Phage display derived serum albumin binding knob domain engineered within a novel VH framework 3 bispecific antibody format
Authors	:	Adams, R.; Macpherson, A.
Deposited on	:	2023-01-17
Resolution	:	1.61 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 (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 (1)) were used in the production of this report:

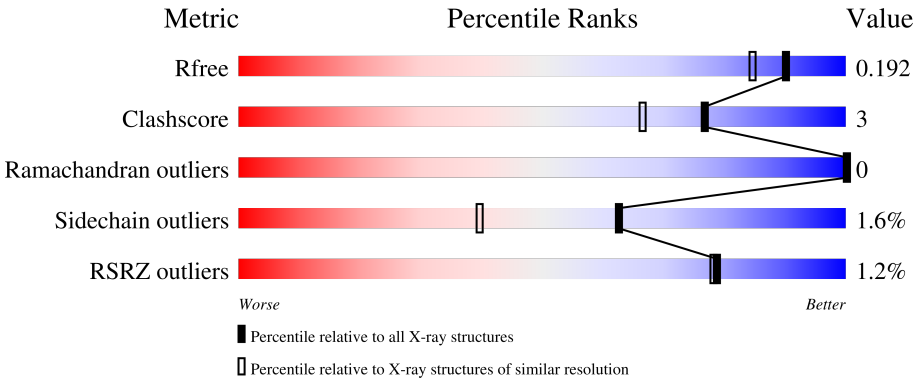
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.33
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.33

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 1.61 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	4693 (1.64-1.60)
Clashscore	141614	5002 (1.64-1.60)
Ramachandran outliers	138981	4888 (1.64-1.60)
Sidechain outliers	138945	4887 (1.64-1.60)
RSRZ outliers	127900	4609 (1.64-1.60)

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	222	96% .
1	C	222	94% 6%
1	E	222	5% 88% 12%
2	B	214	94% 5%
2	D	214	% 93% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	214	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 21212 atoms, of which 9714 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 Fab heavy chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	222	Total	C	H	N	O	S	0	0	0
			3326	1068	1645	281	325	7			
1	C	222	Total	C	H	N	O	S	0	0	0
			3326	1068	1645	281	325	7			
1	E	222	Total	C	H	N	O	S	0	0	0
			3326	1068	1645	281	325	7			

- Molecule 2 is a protein called Fab light chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	213	Total	C	H	N	O	S	0	0	0
			3235	1028	1593	272	337	5			
2	D	213	Total	C	H	N	O	S	0	0	0
			3235	1028	1593	272	337	5			
2	F	213	Total	C	H	N	O	S	0	0	0
			3235	1028	1593	272	337	5			

- Molecule 3 is water.

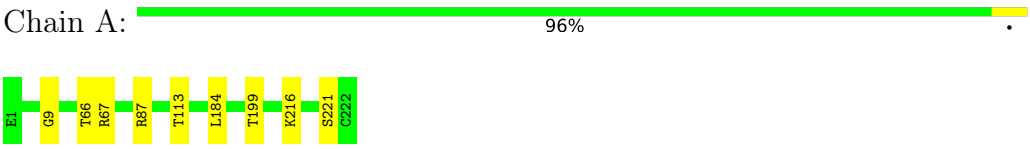
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	309	Total	O	0	0
			309	309		
3	B	220	Total	O	0	0
			220	220		
3	C	305	Total	O	0	0
			305	305		
3	D	248	Total	O	0	0
			248	248		
3	E	226	Total	O	0	0
			226	226		
3	F	221	Total	O	0	0
			221	221		

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.

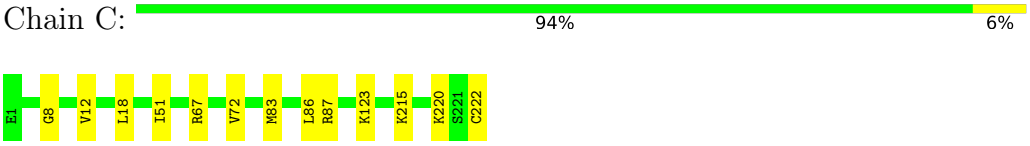
• Molecule 1: Fab heavy chain

Chain A:



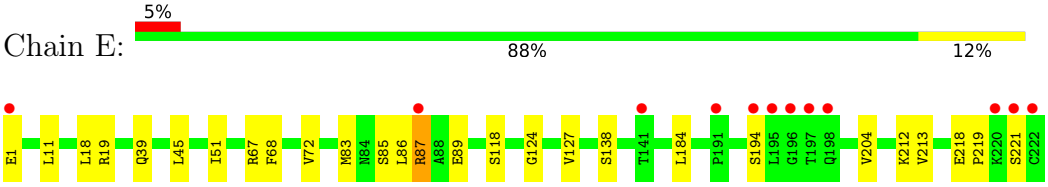
• Molecule 1: Fab heavy chain

Chain C:



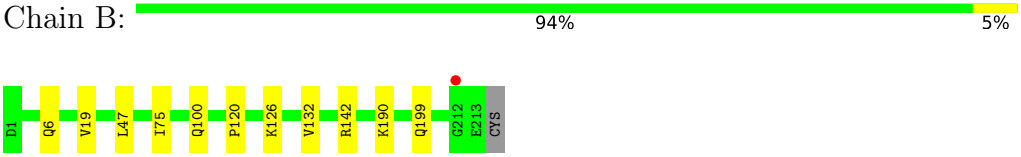
• Molecule 1: Fab heavy chain

Chain E:



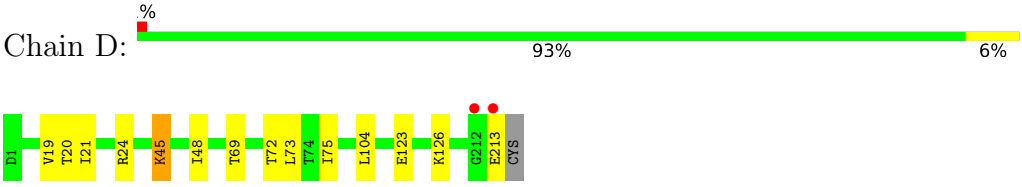
• Molecule 2: Fab light chain

Chain B:



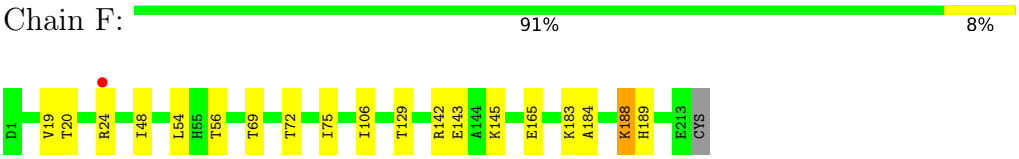
• Molecule 2: Fab light chain

Chain D:



• Molecule 2: Fab light chain

Chain F:


91%
8%
D1
V19
T20
R24
I48
L54
H55
T56
T69
T72
I75
I106
T129
R142
E143
A144
K145
E165
K183
A184
K188
H189
E213
CYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	60.60Å 127.93Å 93.81Å 90.00° 95.47° 90.00°	Depositor
Resolution (Å)	60.32 – 1.61 63.97 – 1.61	Depositor EDS
% Data completeness (in resolution range)	97.5 (60.32-1.61) 97.5 (63.97-1.61)	Depositor EDS
R_{merge}	0.03	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.16 (at 1.61Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.166 , 0.194 0.164 , 0.192	Depositor DCC
R_{free} test set	8952 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	20.4	Xtriage
Anisotropy	0.187	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 41.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	21212	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.05% 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	0.53	0/1726	0.70	0/2350
1	C	0.55	0/1726	0.70	0/2350
1	E	0.47	0/1726	0.66	0/2350
2	B	0.49	0/1677	0.68	1/2279 (0.0%)
2	D	0.46	0/1677	0.65	0/2279
2	F	0.49	0/1677	0.69	0/2279
All	All	0.50	0/10209	0.68	1/13887 (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	E	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})
2	B	47	LEU	CA-CB-CG	5.70	128.41	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	87	ARG	Sidechain

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	1681	1645	1647	8	0
1	C	1681	1645	1647	12	0
1	E	1681	1645	1647	17	0
2	B	1642	1593	1593	6	0
2	D	1642	1593	1593	9	0
2	F	1642	1593	1593	9	0
3	A	309	0	0	4	0
3	B	220	0	0	2	0
3	C	305	0	0	3	0
3	D	248	0	0	2	0
3	E	226	0	0	1	0
3	F	221	0	0	1	0
All	All	11498	9714	9720	58	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 3.

The worst 5 of 58 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:C:220:LYS:HD3	3:C:405:HOH:O	1.68	0.93
1:E:11:LEU:HD11	1:E:118:SER:HB3	1.65	0.78
1:E:83:MET:HB2	1:E:86:LEU:HD21	1.71	0.72
1:C:12:VAL:CG2	1:C:18:LEU:HD12	2.23	0.69
1:A:67:ARG:HD3	3:A:304:HOH:O	1.95	0.67

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	220/222 (99%)	217 (99%)	3 (1%)	0	100	100
1	C	220/222 (99%)	217 (99%)	3 (1%)	0	100	100
1	E	220/222 (99%)	217 (99%)	3 (1%)	0	100	100
2	B	211/214 (99%)	207 (98%)	4 (2%)	0	100	100
2	D	211/214 (99%)	207 (98%)	4 (2%)	0	100	100
2	F	211/214 (99%)	207 (98%)	4 (2%)	0	100	100
All	All	1293/1308 (99%)	1272 (98%)	21 (2%)	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	A	186/186 (100%)	185 (100%)	1 (0%)	88	80
1	C	186/186 (100%)	185 (100%)	1 (0%)	88	80
1	E	186/186 (100%)	181 (97%)	5 (3%)	44	18
2	B	189/190 (100%)	188 (100%)	1 (0%)	88	80
2	D	189/190 (100%)	186 (98%)	3 (2%)	62	40
2	F	189/190 (100%)	182 (96%)	7 (4%)	34	10
All	All	1125/1128 (100%)	1107 (98%)	18 (2%)	62	40

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	145	LYS
2	F	188	LYS
2	F	183	LYS
1	E	194	SER
2	F	142	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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

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, 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	222/222 (100%)	-0.30	0 100 100	11, 19, 34, 45	0
1	C	222/222 (100%)	-0.30	0 100 100	11, 18, 31, 46	0
1	E	222/222 (100%)	-0.04	12 (5%) 25 23	15, 25, 55, 66	0
2	B	213/214 (99%)	-0.29	1 (0%) 91 90	12, 20, 44, 58	0
2	D	213/214 (99%)	-0.37	2 (0%) 84 84	13, 21, 39, 63	0
2	F	213/214 (99%)	-0.32	1 (0%) 91 90	13, 21, 38, 50	0
All	All	1305/1308 (99%)	-0.27	16 (1%) 79 78	11, 20, 41, 66	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	222	CYS	7.0
1	E	196	GLY	5.6
1	E	141	THR	3.6
1	E	197	THR	3.5
1	E	221	SER	3.4

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 (i)

There are no such residues in this entry.