W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

May 26, 2020 – 12:26 pm BST

PDB ID	:	2A6K
Title	:	Crystal Structure Analysis of the germline antibody 36-65 Fab in complex with the dodecapeptide SLGDNLTNHNLR
Authors	:	Sethi, D.K.; Agarwal, A.; Manivel, V.; Rao, K.V.; Salunke, D.M.
Deposited on	:	2005-07-03
Resolution	:	3.00 Å(reported)

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	:	2.11
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 3.00 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.
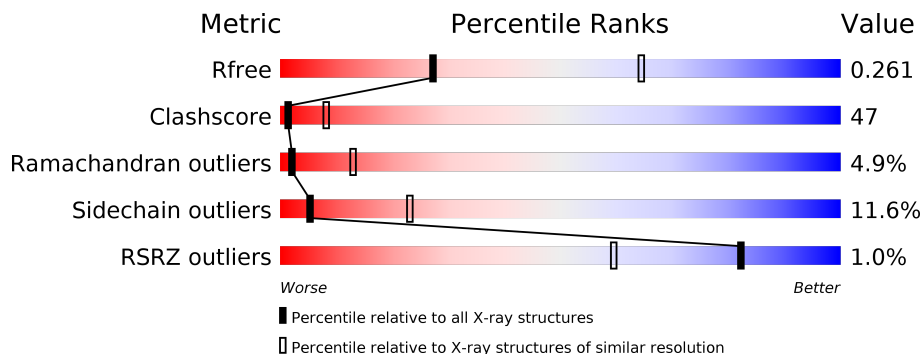
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 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	214	37% 53% 10%
1	L	214	41% 48% 10% •
2	B	222	% 41% 49% 8% •
2	H	222	39% 49% 9% • •
3	P	12	42% 8% 33% 25% 17% 17%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 6764 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 Germline antibody 36-65 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	214	Total 1651	C 1021	N 283	O 340	S 7	0	0	0
1	L	214	Total 1658	C 1024	N 283	O 344	S 7	0	0	0

- Molecule 2 is a protein called Germline antibody 36-65 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	218	Total 1634	C 1036	N 267	O 324	S 7	0	0	0
2	H	216	Total 1637	C 1039	N 267	O 324	S 7	0	0	0

- Molecule 3 is a protein called DODECAPEPTIDE: SLGDNLTNHNLR.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	P	10	Total 81	C 46	N 18	O 17	0	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	33	Total 33	O 33	0	0
4	B	21	Total 21	O 21	0	0
4	L	20	Total 20	O 20	0	0
4	H	26	Total 26	O 26	0	0
4	P	3	Total 3	O 3	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: Germline antibody 36-65 Fab light chain

Chain A:

- Molecule 1: Germline antibody 36-65 Fab light chain

Chain L:

- Molecule 2: Germline antibody 36-65 Fab heavy chain

Chain B:

G147 C148 L149 Y153 F154 P155 V158 T159 V160 T161 W162 S168 V171 H172 T173 F174 P175 A176 V177 L178 Q179 L182 L185 S186 P192 S193 S194 P195 R196 P197 S198 E199 T200 V201 T202 C203 N204 V205 A206 H207 S210 S211 T212 K213 V214 D215 K216 K217 I218 V219 P220 R221

D222

• **Molecule 2: Germline antibody 36-65 Fab heavy chain**

Chain H: 39% 49% 9%

E1 V2 Q3 L4 Q5 Q6 S7 E10 L11 V12 M20 S21 C22 K23 A24 S25 G26 Y27 T28 F29 T30 S31 Y32 G33 I34 N35 W36 V37 K38 Q39 R40 P41 G42 Q43 G44 L45 I48 G49 Y50 I51 N52 P53 G54 N55 G56 Y57 T58 K59 Y60 N61 E62 K63 G66 T69

V72 D73 K74 S75 S76 S77 Q82 L83 R84 S85 L86 T87 S88 E89 D90 S91 A92 F95 C96 A97 R98 S99 V100 Y101 Y102 G103 G104 S105 Y106 Y107 F108 Q113 L117 S120 S121 A122 K123 T124 T125 P126 P127 P131 L132 A133 P134 G135 SER ALA ALA GLN THR ASN S142 M143

V144 T145 L146 G147 V150 Y153 F154 P155 T159 W162 N163 S164 G165 S166 V171 F174 L178 Q179 S180 D181 L182 L185 S186 S187 S188 V189 T190 V191 P192 S193 S194 P195 R196 P197 S198 T202 C203 N204 V205 A206 H207 P208 A209 S210 S211 T212 K213 K216 V219

P220 R221 D222

• **Molecule 3: DODECAPEPTIDE: SLGDNLTNHNLR**

Chain P: 8% 42% 33% 25% 17% 17%

SER LEU G3 D4 N5 L6 T7 N8 H9 N10 L11 R12

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	53.13Å 144.92Å 70.99Å 90.00° 104.35° 90.00°	Depositor
Resolution (Å)	50.00 – 3.00 17.52 – 2.89	Depositor EDS
% Data completeness (in resolution range)	88.6 (50.00-3.00) 88.4 (17.52-2.89)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	0.43 (at 2.87Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.245 , 0.264 0.244 , 0.261	Depositor DCC
R_{free} test set	1813 reflections (8.24%)	wwPDB-VP
Wilson B-factor (Å^2)	35.3	Xtriage
Anisotropy	0.229	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 53.0	EDS
L-test for twinning[2]	$<|L|> = 0.51$, $<L^2> = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	6764	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.27% 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	0.46	0/1684	0.73	0/2284
1	L	0.52	1/1691 (0.1%)	0.80	2/2293 (0.1%)
2	B	0.60	4/1677 (0.2%)	0.81	2/2289 (0.1%)
2	H	0.49	0/1680	0.78	2/2290 (0.1%)
3	P	1.16	0/81	1.93	2/107 (1.9%)
All	All	0.53	5/6813 (0.1%)	0.80	8/9263 (0.1%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	102	TYR	CB-CG	7.40	1.62	1.51
1	L	154	GLU	CB-CG	6.79	1.65	1.52
2	B	101	TYR	CB-CG	5.70	1.60	1.51
2	B	101	TYR	CA-C	5.62	1.67	1.52
2	B	102	TYR	N-CA	5.55	1.57	1.46

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	101	TYR	N-CA-C	9.06	135.46	111.00
2	H	222	ASP	N-CA-C	7.86	132.23	111.00
1	L	154	GLU	N-CA-CB	-7.04	97.94	110.60
2	B	101	TYR	CA-CB-CG	6.79	126.30	113.40
3	P	5	ASN	CA-C-N	-6.59	102.69	117.20

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	1651	0	1569	164	1
1	L	1658	0	1581	151	2
2	B	1634	0	1565	141	1
2	H	1637	0	1589	165	0
3	P	81	0	73	41	0
4	A	33	0	0	13	0
4	B	21	0	0	6	0
4	H	26	0	0	1	0
4	L	20	0	0	5	0
4	P	3	0	0	2	0
All	All	6764	0	6377	618	2

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 47.

The worst 5 of 618 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)
3:P:4:ASP:OD1	3:P:5:ASN:N	1.76	1.16
1:L:31:ASN:HD22	1:L:51:THR:HG21	1.10	1.10
2:H:192:PRO:HD2	2:H:195:PRO:HG2	1.35	1.07
1:L:136:LEU:HD23	1:L:144:ILE:HD13	1.34	1.06
2:H:220:PRO:O	2:H:221:ARG:C	1.92	1.06

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:199:GLU:OE2	1:L:188:ARG:NH1[1_465]	2.09	0.11
1:A:60:SER:OG	1:L:79:GLU:OE1[2_555]	2.17	0.03

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	212/214 (99%)	185 (87%)	18 (8%)	9 (4%)	3	16
1	L	212/214 (99%)	182 (86%)	22 (10%)	8 (4%)	3	18
2	B	214/222 (96%)	184 (86%)	19 (9%)	11 (5%)	2	12
2	H	212/222 (96%)	184 (87%)	16 (8%)	12 (6%)	1	10
3	P	8/12 (67%)	3 (38%)	3 (38%)	2 (25%)	0	0
All	All	858/884 (97%)	738 (86%)	78 (9%)	42 (5%)	2	13

5 of 42 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	154	GLU
2	B	43	GLN
2	B	89	GLU
2	B	105	SER
1	L	8	THR

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	187/192 (97%)	168 (90%)	19 (10%)	7	28
1	L	190/192 (99%)	165 (87%)	25 (13%)	4	18
2	B	181/190 (95%)	165 (91%)	16 (9%)	10	36
2	H	184/190 (97%)	160 (87%)	24 (13%)	4	19
3	P	9/11 (82%)	6 (67%)	3 (33%)	0	1
All	All	751/775 (97%)	664 (88%)	87 (12%)	5	23

5 of 87 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	61	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	L	123	GLU
2	H	204	ASN
1	L	63	SER
1	L	96	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	82	GLN
1	L	31	ASN
3	P	9	HIS
2	B	113	GLN
1	L	6	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 (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	214/214 (100%)	-0.16	0	100	100	12, 32, 56, 63	0
1	L	214/214 (100%)	-0.20	1 (0%)	91	75	10, 30, 48, 60	0
2	B	218/222 (98%)	-0.06	2 (0%)	84	63	8, 38, 58, 68	0
2	H	216/222 (97%)	-0.22	1 (0%)	91	75	1, 31, 54, 65	0
3	P	10/12 (83%)	1.92	5 (50%)	0	0	35, 44, 51, 58	0
All	All	872/884 (98%)	-0.14	9 (1%)	82	59	1, 33, 56, 68	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	P	5	ASN	3.4
2	H	222	ASP	3.1
3	P	10	ASN	3.1
1	L	7	THR	2.9
3	P	4	ASP	2.8

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.