W O R L D W I D E

wwPDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

May 13, 2020 – 12:58 am BST

PDB ID	:	2V7N
Title	:	Unusual twinning in crystals of the CitS binding antibody Fab fragment f3p4
Authors	:	Frey, D.; Huber, T.; Plueckthun, A.; Gruetter, M.G.
Deposited on	:	2007-07-31
Resolution	:	1.92 Å(reported)

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

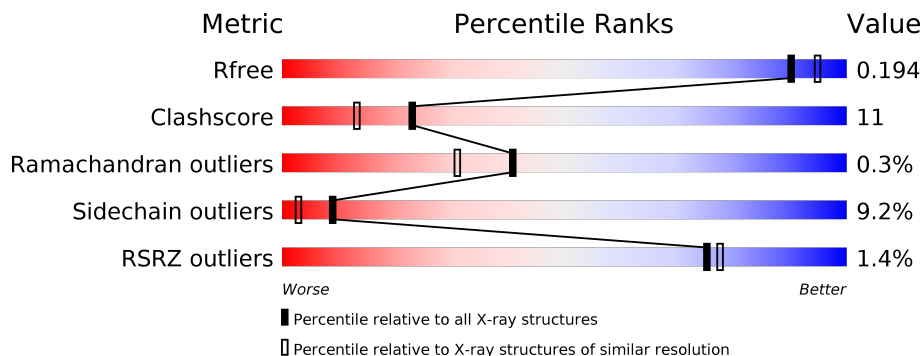|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 1.92 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.
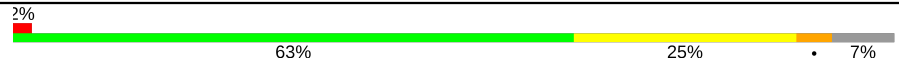
Metric	Percentile Ranks	Value
Rfree		0.194
Clashscore		11
Ramachandran outliers		0.3%
Sidechain outliers		9.2%
RSRZ outliers		1.4%

Worse   *Better*

∎ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	7937 (1.94-1.90)
Clashscore	141614	8644 (1.94-1.90)
Ramachandran outliers	138981	8530 (1.94-1.90)
Sidechain outliers	138945	8530 (1.94-1.90)
RSRZ outliers	127900	7793 (1.94-1.90)

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	215	71% 23% • ••
1	C	215	71% 24% • •
1	E	215	69% 25% 5% •
1	G	215	% 74% 23% ••
2	B	229	70% 21% • 7%
2	D	229	4% 68% 21% • 7%

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	229	2% 63% 25% • 7%
2	H	229	2% 66% 22% • 8%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 13417 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 IMMUNOGLOBULIN LIGHT CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	213	Total 1631	C 1019	N 275	O 333	S 4	0	0	0
1	C	212	Total 1627	C 1017	N 274	O 332	S 4	0	0	0
1	E	213	Total 1631	C 1019	N 275	O 333	S 4	0	0	0
1	G	212	Total 1627	C 1017	N 274	O 332	S 4	0	0	0

- Molecule 2 is a protein called IMMUNOGLOBULIN HEAVY CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	214	Total 1598	C 1011	N 272	O 309	S 6	0	0	0
2	D	213	Total 1595	C 1010	N 272	O 307	S 6	0	0	0
2	F	212	Total 1586	C 1005	N 270	O 305	S 6	0	0	0
2	H	211	Total 1582	C 1003	N 269	O 304	S 6	0	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	85	Total 85	O 85	0	0
3	B	58	Total 58	O 58	0	0
3	C	80	Total 80	O 80	0	0
3	D	61	Total 61	O 61	0	0

Continued on next page...

wwPDB X-ray Structure Validation Summary Report

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	64	Total 64	O 64	0	0
3	F	51	Total 51	O 51	0	0
3	G	80	Total 80	O 80	0	0
3	H	61	Total 61	O 61	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.
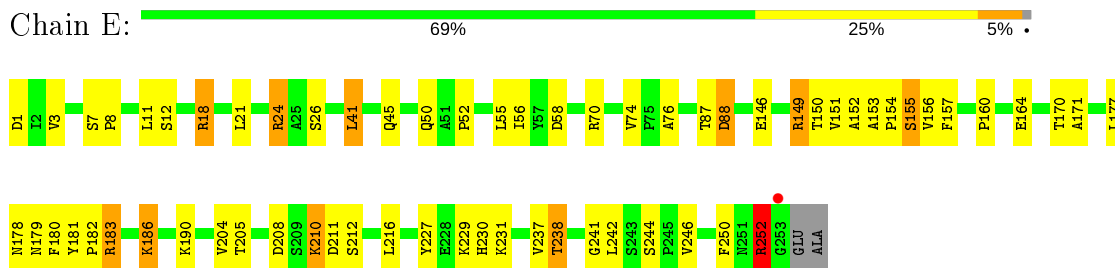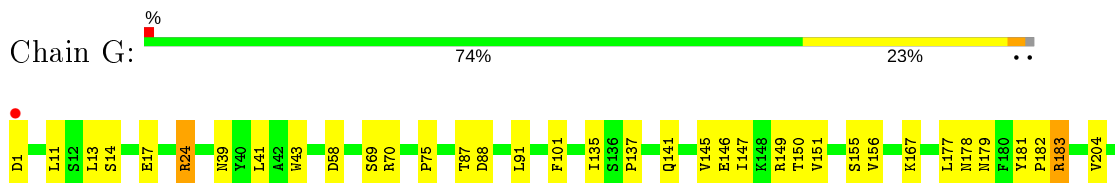
- Molecule 1: IMMUNOGLOBULIN LIGHT CHAIN

Chain A:

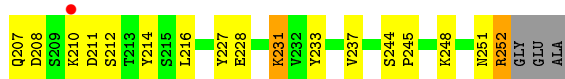- Molecule 1: IMMUNOGLOBULIN LIGHT CHAIN

Chain C:

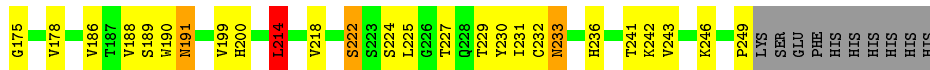- Molecule 1: IMMUNOGLOBULIN LIGHT CHAIN
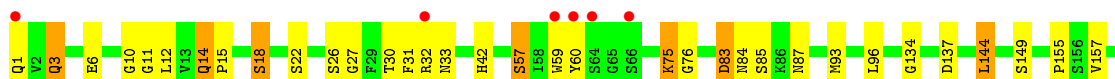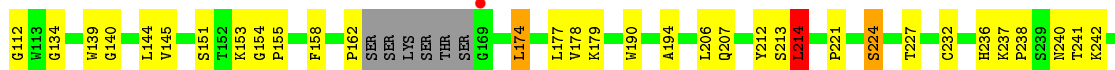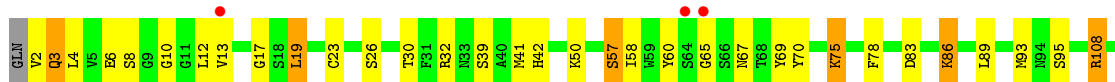
Chain E:

- Molecule 1: IMMUNOGLOBULIN LIGHT CHAIN

Chain G:

Q207 D208 S209 K210 D211 S212 T213 Y214 S215 L216 Y227 E228 K231 V232 Y233 V237 S244 P245 K248 N251 R252 GLY GLU ALA

● Molecule 2: IMMUNOGLOBULIN HEAVY CHAIN

Chain B:

70%　　　21%　　　·　7%

GLN V2 E6 L12 V13 Q14 L19 T30 F31 R32 R45 K50 W54 W59 Y60 S64 G65 S66 N67 Y70 A71 V74 K75 D83 N84 S85 K86 C106 G112 W139 S149 P155 S156 V157 P162 S163 S164 LYS SER THR SER G169

G175 V178 V186 T187 V188 S189 W190 N191 V199 H200 L214 V218 S222 S223 S224 L225 G226 T227 Q228 T229 Y230 I231 C232 N233 H236 T241 K242 V243 K246 P249 LYS SER GLU PHE HIS HIS HIS HIS HIS

● Molecule 2: IMMUNOGLOBULIN HEAVY CHAIN

Chain D:

4%

68%　　　21%　　·　7%

Q1 V2 Q3 E6 G10 G11 L12 V13 Q14 P15 S18 S22 S26 G27 F29 T30 F31 R32 N33 H42 S57 I58 W59 Y60 S64 G65 S66 K75 G76 D83 N84 S85 K86 N87 M93 L96 G134 D137 L144 S149 P155 S156 V157

P162 SER SER LYS SER THR SER G169 G170 D180 Y181 N191 H200 L211 Y212 S213 L214 P221 S222 S223 S224 L225 G226 T227 I231 C232 N233 H236 K237 T241 K242 K245 E248 P249 LYS SER GLU PHE HIS HIS HIS HIS

● Molecule 2: IMMUNOGLOBULIN HEAVY CHAIN

Chain F:

2%

63%　　　25%　·　7%

GLN V2 Q3 L4 V5 E6 S8 G9 G10 G11 L12 V13 G17 S18 L19 C23 S26 T30 F31 R32 N33 S39 A40 M41 H42 K50 S57 I58 W59 Y60 S64 G65 S66 N67 T68 Y69 Y70 K75 F78 D83 K86 L89 M93 N94 S95 R108

G112 W113 G134 W139 G140 L144 V145 S151 T152 K153 G154 P155 F158 P162 SER SER LYS SER THR SER G169 L174 L177 V178 K179 W190 A194 L206 Q207 S213 L214 P221 S224 T227 C232 H236 K237 P238 S239 N240 T241 K242

K245 P249 LYS SER GLU PHE HIS HIS HIS HIS HIS

● Molecule 2: IMMUNOGLOBULIN HEAVY CHAIN

Chain H:

2%

66%　　　22%　·　8%

GLN V2 E6 L12 V13 Q14 G17 R30 G27 F29 T30 F31 R32 M41 H42 W43 P48 G49 K50 W54 S57 Y60 N67 T68 Y69 D72 K75 R82 D83 N84 L89 Y90 L91 T101 G134 Q141 G142 T143 L144 V145 T146

P155 P159 P162 SER SER LYS SER THR SER GLY G170 T171 A172 D180 Y181 N191 V199 A204 L211 Y212 S213 L214 V218 L225 G226 T227 Q228 I231 C232 N233 H236 K237 P238 S239 N240 T241 K242 K245 K246 V247 E248 P249 LYS SER GLU PHE HIS HIS HIS

HIS HIS HIS HIS

4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	102.79Å 185.92Å 102.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 1.92 29.99 – 1.90	Depositor EDS
% Data completeness (in resolution range)	95.1 (15.00-1.92) 96.7 (29.99-1.90)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	3.99 (at 1.91Å)	Xtriage
Refinement program	SHELXL-97	Depositor
R, R_{free}	0.155 , 0.202 0.151 , 0.194	Depositor DCC
R_{free} test set	4240 reflections (2.80%)	wwPDB-VP
Wilson B-factor (Å^2)	18.9	Xtriage
Anisotropy	0.230	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 34.1	EDS
L-test for twinning[2]	$<|L|> = 0.37, <L^2> = 0.20$	Xtriage
Estimated twinning fraction	0.389 for l,-k,h	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	13417	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.24% 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 (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	0.39	0/1667	1.00	6/2267 (0.3%)
1	C	0.38	0/1663	1.03	5/2262 (0.2%)
1	E	0.39	0/1667	1.04	5/2267 (0.2%)
1	G	0.39	0/1663	1.02	4/2262 (0.2%)
2	B	0.37	0/1639	1.02	6/2234 (0.3%)
2	D	0.37	0/1636	0.98	5/2230 (0.2%)
2	F	0.37	0/1627	1.04	6/2218 (0.3%)
2	H	0.36	0/1623	0.97	3/2213 (0.1%)
All	All	0.38	0/13185	1.01	40/17953 (0.2%)

There are no bond length outliers.

The worst 5 of 40 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	183	ARG	CD-NE-CZ	14.47	143.87	123.60
1	E	252	ARG	NE-CZ-NH1	-10.07	115.27	120.30
2	F	214	LEU	CA-CB-CG	9.60	137.38	115.30
1	A	252	ARG	NE-CZ-NH1	-8.86	115.87	120.30
2	B	83	ASP	CB-CG-OD2	8.74	126.17	118.30

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	1631	0	1580	35	0
1	C	1627	0	1577	39	0
1	E	1631	0	1580	51	0
1	G	1627	0	1577	27	0
2	B	1598	0	1549	38	0
2	D	1595	0	1550	28	0
2	F	1586	0	1539	40	0
2	H	1582	0	1536	42	0
3	A	85	0	0	1	0
3	B	58	0	0	2	0
3	C	80	0	0	1	0
3	D	61	0	0	1	0
3	E	64	0	0	3	0
3	F	51	0	0	1	0
3	G	80	0	0	0	0
3	H	61	0	0	1	0
All	All	13417	0	12488	288	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 11.

The worst 5 of 288 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:B:229:THR:HG23	2:B:246:LYS:HE3	1.55	0.88
1:C:6:GLN:H	1:C:141:GLN:HE22	1.20	0.87
1:G:204:VAL:HG22	1:G:216:LEU:HD12	1.59	0.84
2:F:236:HIS:HB3	2:F:241:THR:HB	1.61	0.83
1:C:164:GLU:HA	1:C:167:LYS:HD3	1.60	0.81

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	211/215 (98%)	206 (98%)	5 (2%)	0	100	100
1	C	210/215 (98%)	203 (97%)	7 (3%)	0	100	100
1	E	211/215 (98%)	205 (97%)	6 (3%)	0	100	100
1	G	210/215 (98%)	206 (98%)	4 (2%)	0	100	100
2	B	210/229 (92%)	201 (96%)	8 (4%)	1 (0%)	29	18
2	D	209/229 (91%)	200 (96%)	7 (3%)	2 (1%)	15	6
2	F	208/229 (91%)	200 (96%)	7 (3%)	1 (0%)	29	18
2	H	207/229 (90%)	200 (97%)	6 (3%)	1 (0%)	29	18
All	All	1676/1776 (94%)	1621 (97%)	50 (3%)	5 (0%)	41	31

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	134	GLY
2	D	10	GLY
2	H	134	GLY
2	D	134	GLY
2	B	65	GLY

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	186/187 (100%)	166 (89%)	20 (11%)	6	2
1	C	186/187 (100%)	169 (91%)	17 (9%)	9	3
1	E	186/187 (100%)	171 (92%)	15 (8%)	11	4
1	G	186/187 (100%)	172 (92%)	14 (8%)	13	5
2	B	175/190 (92%)	164 (94%)	11 (6%)	18	8
2	D	174/190 (92%)	155 (89%)	19 (11%)	6	2
2	F	173/190 (91%)	153 (88%)	20 (12%)	5	1
2	H	173/190 (91%)	157 (91%)	16 (9%)	9	3

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1439/1508 (95%)	1307 (91%)	132 (9%)	9 3

5 of 132 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	222	SER
1	E	149	ARG
2	H	75	LYS
2	D	227	THR
1	E	7	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 39 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	230	HIS
2	D	84	ASN
2	H	84	ASN
1	C	251	ASN
2	D	14	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	213/215 (99%)	-0.15	0	100	100	13, 21, 34, 51	0
1	C	212/215 (98%)	-0.22	0	100	100	13, 21, 33, 51	0
1	E	213/215 (99%)	-0.15	1 (0%)	91	92	14, 22, 34, 51	0
1	G	212/215 (98%)	-0.21	2 (0%)	84	85	13, 21, 33, 51	0
2	B	214/229 (93%)	0.03	1 (0%)	91	92	14, 24, 45, 60	0
2	D	213/229 (93%)	0.14	9 (4%)	36	39	15, 24, 46, 59	0
2	F	212/229 (92%)	0.08	5 (2%)	59	62	16, 24, 44, 60	0
2	H	211/229 (92%)	-0.02	5 (2%)	59	62	15, 24, 43, 60	0
All	All	1700/1776 (95%)	-0.06	23 (1%)	75	77	13, 23, 40, 60	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	169	GLY	7.4
2	F	169	GLY	3.5
2	H	225	LEU	3.2
2	H	227	THR	3.1
2	F	65	GLY	3.0

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.