



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 17, 2021 – 08:06 AM EDT

PDB ID : 1MHP
Title : Crystal structure of a chimeric alpha1 integrin I-domain in complex with the Fab fragment of a humanized neutralizing antibody
Authors : Karpusas, M.; Taylor, F.; Ferrant, J.; Weinreb, P.; Garber, E.
Deposited on : 2002-08-20
Resolution : 2.80 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 : 2.23.2
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) : 2.23.2

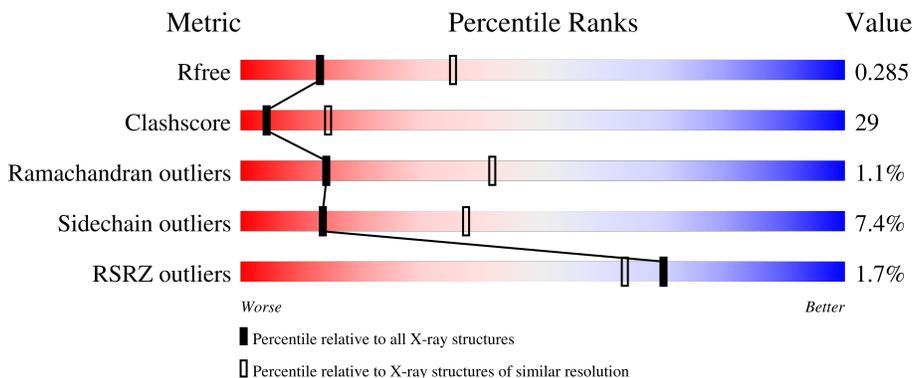
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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.



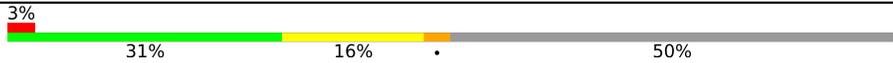
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 $\leq 5\%$. 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	192	 53% 39% 5% 3% 1%
1	B	192	 51% 44% 5% 2% 1%
2	H	219	 52% 44% 5% 1% 1%
2	X	219	 25% 24% 5% 5% 46%
3	L	212	 60% 36% 5% 1% 1%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Y	212	 <p>A horizontal bar chart representing the quality of chain. The bar is divided into four segments: a small red segment (3%), a green segment (31%), a yellow segment (16%), and a grey segment (50%). The percentages are labeled below the bar.</p>

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 7943 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 integrin alpha 1, (RESIDUES 169-360).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	184	1454	916	251	283	4	0	0	0
1	B	192	1521	956	266	295	4	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	217	VAL	GLY	engineered mutation	UNP P18614
A	218	GLN	ARG	engineered mutation	UNP P18614
A	219	ARG	GLN	engineered mutation	UNP P18614
A	222	ARG	LEU	engineered mutation	UNP P18614
B	217	VAL	GLY	engineered mutation	UNP P18614
B	218	GLN	ARG	engineered mutation	UNP P18614
B	219	ARG	GLN	engineered mutation	UNP P18614
B	222	ARG	LEU	engineered mutation	UNP P18614

- Molecule 2 is a protein called Fab fragment, heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	H	219	1629	1027	275	321	6	15	0	0
2	X	118	898	565	155	174	4	0	0	0

- Molecule 3 is a protein called FAB FRAGMENT, light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	L	212	1633	1024	273	330	6	0	0	0
3	Y	105	806	511	133	159	3	0	0	0

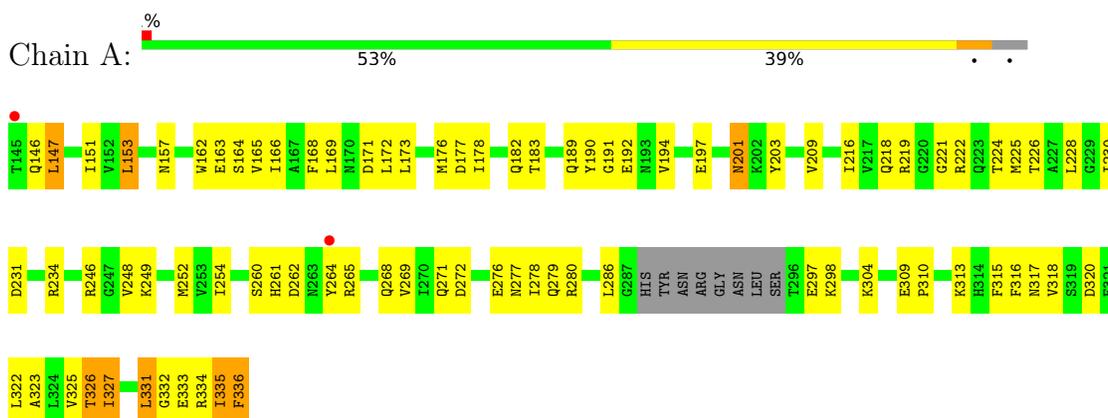
- Molecule 4 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total 1	Mn 1	0	0
4	B	1	Total 1	Mn 1	0	0

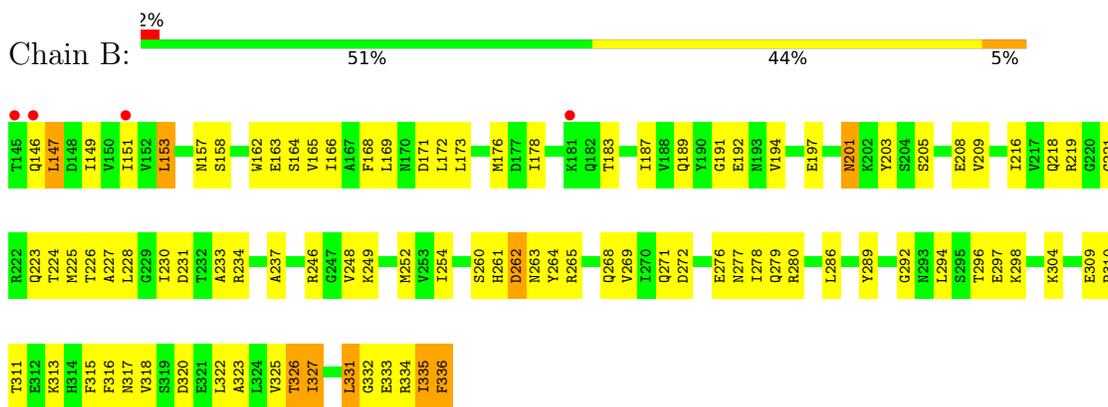
3 Residue-property plots [i](#)

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.

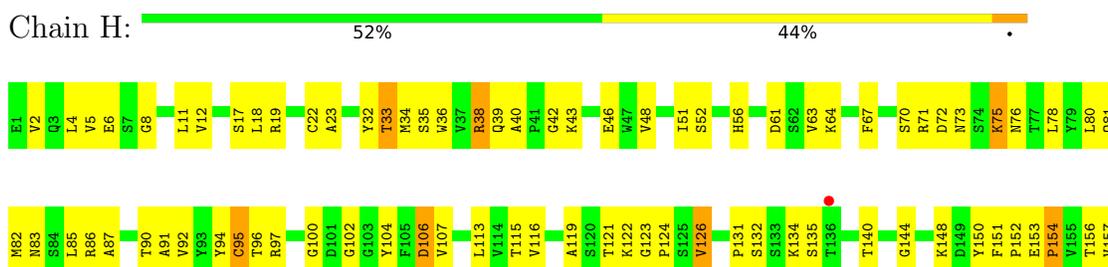
- Molecule 1: integrin alpha 1, (RESIDUES 169-360)



- Molecule 1: integrin alpha 1, (RESIDUES 169-360)



- Molecule 2: Fab fragment, heavy chain



4 Data and refinement statistics i

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α , β , γ	255.09Å 255.09Å 38.64Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	35.00 – 2.80 33.05 – 2.80	Depositor EDS
% Data completeness (in resolution range)	(Not available) (35.00-2.80) 94.0 (33.05-2.80)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.96 (at 2.81Å)	Xtrriage
Refinement program	CNX	Depositor
R, R_{free}	0.213 , 0.272 0.232 , 0.285	Depositor DCC
R_{free} test set	2830 reflections (8.03%)	wwPDB-VP
Wilson B-factor (Å ²)	38.7	Xtrriage
Anisotropy	0.465	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 59.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.016 for h,-h-k,-l	Xtrriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	7943	wwPDB-VP
Average B, all atoms (Å ²)	37.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.65% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MN

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	0.38	0/1472	0.61	0/1983
1	B	0.37	0/1542	0.61	0/2079
2	H	0.45	0/1668	0.74	0/2268
2	X	0.41	0/918	0.68	0/1242
3	L	0.43	0/1674	0.69	0/2277
3	Y	0.36	0/830	0.62	0/1132
All	All	0.40	0/8104	0.66	0/10981

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	1454	0	1463	75	0
1	B	1521	0	1524	84	0
2	H	1629	0	1594	95	0
2	X	898	0	860	78	0
3	L	1633	0	1567	102	0
3	Y	806	0	768	48	0
4	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	1	0	0	0	0
All	All	7943	0	7776	458	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 29.

The worst 5 of 458 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:Y:6:GLN:H	3:Y:99:GLN:NE2	1.48	1.11
3:L:6:GLN:H	3:L:99:GLN:NE2	1.48	1.10
2:H:71:ARG:HD3	2:H:73:ASN:HD21	1.18	1.09
2:X:33:THR:HB	2:X:52:SER:HA	1.45	0.98
1:B:249:LYS:HE2	1:B:279:GLN:HG2	1.48	0.95

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	180/192 (94%)	165 (92%)	14 (8%)	1 (1%)	25	56
1	B	190/192 (99%)	172 (90%)	17 (9%)	1 (0%)	29	61
2	H	217/219 (99%)	200 (92%)	16 (7%)	1 (0%)	29	61
2	X	116/219 (53%)	104 (90%)	8 (7%)	4 (3%)	3	13
3	L	210/212 (99%)	190 (90%)	17 (8%)	3 (1%)	11	34
3	Y	103/212 (49%)	94 (91%)	8 (8%)	1 (1%)	15	44
All	All	1016/1246 (82%)	925 (91%)	80 (8%)	11 (1%)	14	41

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	326	THR
3	L	54	ALA
1	B	326	THR
2	X	106	ASP
3	Y	54	ALA

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	160/167 (96%)	150 (94%)	10 (6%)	18	46
1	B	167/167 (100%)	156 (93%)	11 (7%)	16	44
2	H	181/181 (100%)	167 (92%)	14 (8%)	13	35
2	X	94/181 (52%)	87 (93%)	7 (7%)	13	37
3	L	187/187 (100%)	174 (93%)	13 (7%)	15	40
3	Y	91/187 (49%)	81 (89%)	10 (11%)	6	19
All	All	880/1070 (82%)	815 (93%)	65 (7%)	13	37

5 of 65 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	Y	15	VAL
3	Y	72	LEU
3	L	4	LEU
2	H	214	LYS
3	Y	77	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	268	GLN
1	B	279	GLN
3	Y	3	GLN
2	H	205	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	H	76	ASN

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	184/192 (95%)	-0.18	2 (1%) 80 75	6, 28, 81, 137	0
1	B	192/192 (100%)	-0.09	4 (2%) 63 54	13, 34, 79, 96	0
2	H	219/219 (100%)	-0.20	1 (0%) 91 88	8, 31, 67, 99	3 (1%)
2	X	118/219 (53%)	-0.20	0 100 100	15, 40, 71, 113	0
3	L	212/212 (100%)	-0.33	3 (1%) 75 70	9, 29, 62, 138	0
3	Y	105/212 (49%)	0.29	7 (6%) 17 10	17, 42, 95, 130	0
All	All	1030/1246 (82%)	-0.15	17 (1%) 70 63	6, 33, 76, 138	3 (0%)

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	Y	105	ILE	5.0
1	A	145	THR	4.2
3	Y	82	PHE	4.0
3	Y	14	SER	3.7
1	B	181	LYS	3.4

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 monosaccharides in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
4	MN	B	400	1/1	0.95	0.06	34,34,34,34	0
4	MN	A	400	1/1	0.97	0.13	34,34,34,34	0

6.5 Other polymers [i](#)

There are no such residues in this entry.