



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 4, 2021 – 02:02 pm GMT

PDB ID : 7NMU
Title : Crystal structure of human platelet glycoprotein VI in complex with an inhibitory nanobody.
Authors : Slater, A.; Jonas, E.; Watson, S.P.
Deposited on : 2021-02-23
Resolution : 2.50 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	:	2.17.1
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)	:	2.17.1

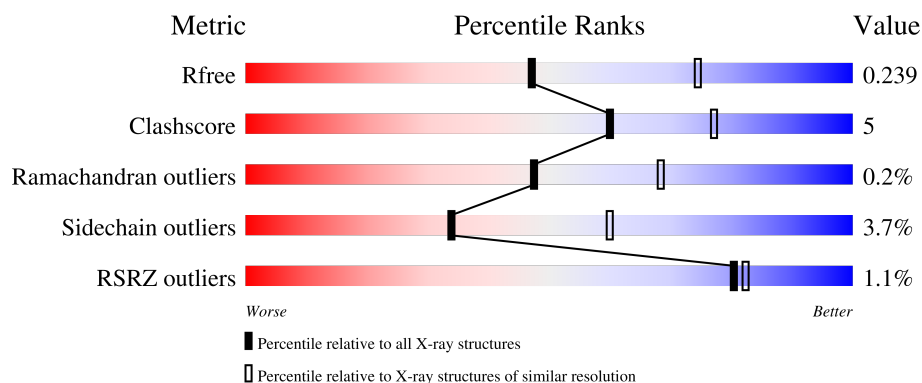
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION



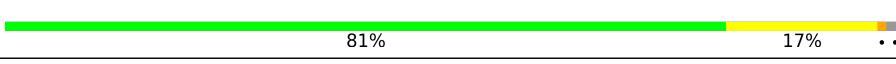

The reported resolution of this entry is 2.50 Å.

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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	AAA	194	 3% 78% 11% • 10%
1	BBB	194	 % 78% 12% • 8%
2	CCC	130	 81% 17% ••
2	DDD	130	 90% 8% ••

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 9445 atoms, of which 4511 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 Platelet glycoprotein VI.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	174	Total	C	H	N	O	S	64	0	0
			2647	853	1314	222	253	5			
1	BBB	178	Total	C	H	N	O	S	65	0	0
			2698	869	1340	226	258	5			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	1	ASP	-	expression tag	UNP Q9HCN6
AAA	2	LYS	-	expression tag	UNP Q9HCN6
AAA	3	LEU	-	expression tag	UNP Q9HCN6
AAA	4	ALA	-	expression tag	UNP Q9HCN6
AAA	5	SER	-	expression tag	UNP Q9HCN6
AAA	76	GLN	ASN	engineered mutation	UNP Q9HCN6
AAA	189	ASP	-	expression tag	UNP Q9HCN6
AAA	190	PRO	-	expression tag	UNP Q9HCN6
AAA	191	ILE	-	expression tag	UNP Q9HCN6
AAA	192	GLU	-	expression tag	UNP Q9HCN6
AAA	193	GLY	-	expression tag	UNP Q9HCN6
AAA	194	ARG	-	expression tag	UNP Q9HCN6
BBB	1	ASP	-	expression tag	UNP Q9HCN6
BBB	2	LYS	-	expression tag	UNP Q9HCN6
BBB	3	LEU	-	expression tag	UNP Q9HCN6
BBB	4	ALA	-	expression tag	UNP Q9HCN6
BBB	5	SER	-	expression tag	UNP Q9HCN6
BBB	76	GLN	ASN	engineered mutation	UNP Q9HCN6
BBB	189	ASP	-	expression tag	UNP Q9HCN6
BBB	190	PRO	-	expression tag	UNP Q9HCN6
BBB	191	ILE	-	expression tag	UNP Q9HCN6
BBB	192	GLU	-	expression tag	UNP Q9HCN6
BBB	193	GLY	-	expression tag	UNP Q9HCN6
BBB	194	ARG	-	expression tag	UNP Q9HCN6

- Molecule 2 is a protein called Nanobody 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	DDD	129	Total	C	H	N	O	S	59	0	0
			1922	619	932	163	202	6			
2	CCC	128	Total	C	H	N	O	S	59	0	0
			1908	614	925	162	201	6			

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	CCC	1	Total	Ca	0	0
			1	1		

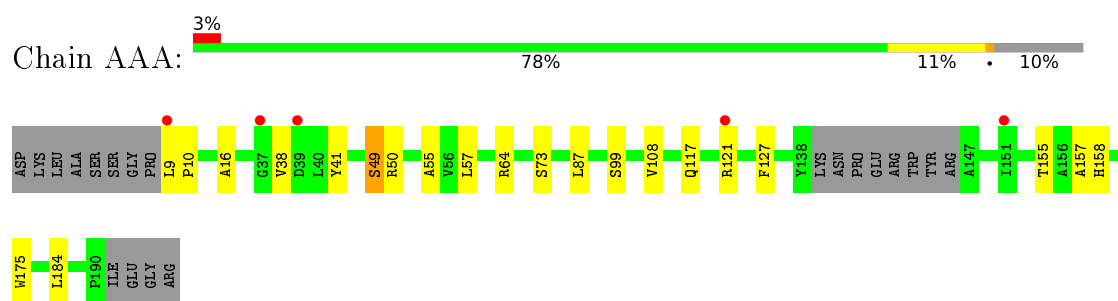
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	AAA	57	Total	O	0	0
			57	57		
4	BBB	84	Total	O	0	0
			84	84		
4	DDD	64	Total	O	0	0
			64	64		
4	CCC	64	Total	O	0	0
			64	64		

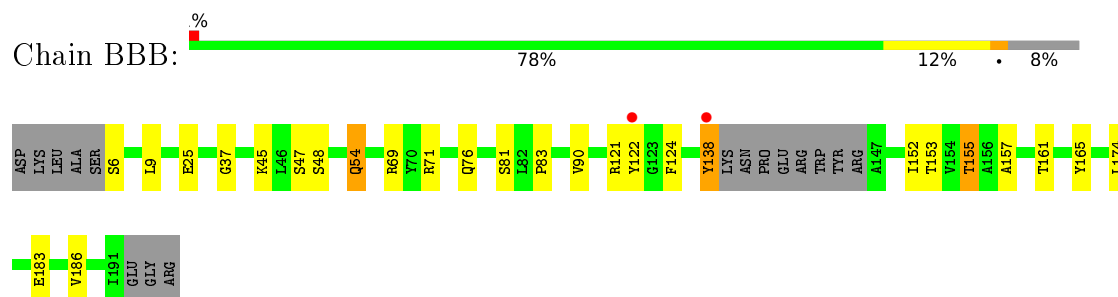
3 Residue-property plots [i](#)

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.

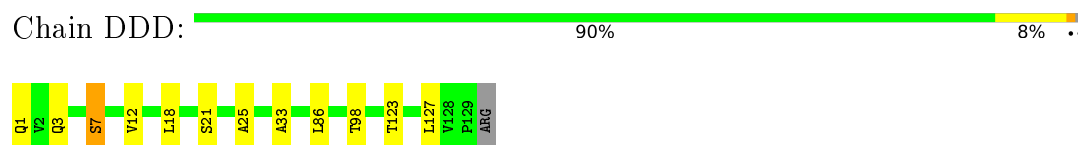
- Molecule 1: Platelet glycoprotein VI



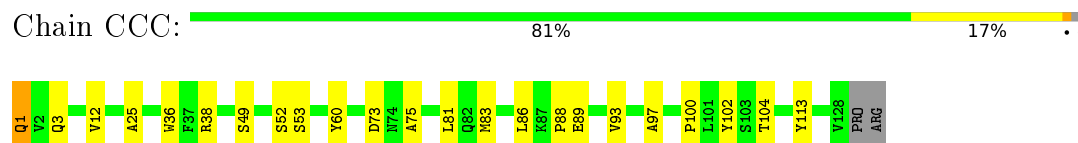
- Molecule 1: Platelet glycoprotein VI



- Molecule 2: Nanobody 2



- Molecule 2: Nanobody 2



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	69.91Å 84.75Å 124.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	70.08 – 2.50 69.98 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.9 (70.08-2.50) 99.9 (69.98-2.50)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.70 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.8.0257	Depositor
R, R_{free}	0.174 , 0.235 0.180 , 0.239	Depositor DCC
R_{free} test set	1329 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	40.4	Xtriage
Anisotropy	0.056	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 37.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	9445	wwPDB-VP
Average B, all atoms (Å ²)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.60% 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: CA

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	AAA	0.65	0/1371	0.82	0/1872
1	BBB	0.68	0/1397	0.86	0/1908
2	CCC	0.69	0/1004	0.89	0/1362
2	DDD	0.67	0/1012	0.86	0/1374
All	All	0.67	0/4784	0.86	0/6516

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	AAA	1333	1314	1311	12	0
1	BBB	1358	1340	1337	19	0
2	CCC	983	925	925	14	0
2	DDD	990	932	932	4	0
3	CCC	1	0	0	0	0
4	AAA	57	0	0	0	0
4	BBB	84	0	0	0	0
4	CCC	64	0	0	0	0
4	DDD	64	0	0	0	0
All	All	4934	4511	4505	42	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.

The worst 5 of 42 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:BBB:161:THR:OG1	1:BBB:183:GLU:OE1	2.16	0.58
1:BBB:155:THR:HG23	1:BBB:157:ALA:H	1.67	0.57
1:AAA:41:TYR:CD2	1:AAA:55:ALA:HB2	2.39	0.57
1:BBB:37:GLY:O	1:BBB:76:GLN:NE2	2.38	0.57
1:BBB:47:SER:OG	1:BBB:69:ARG:NH2	2.43	0.52

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	AAA	170/194 (88%)	161 (95%)	8 (5%)	1 (1%)	25	43
1	BBB	174/194 (90%)	169 (97%)	5 (3%)	0	100	100
2	CCC	126/130 (97%)	124 (98%)	2 (2%)	0	100	100
2	DDD	127/130 (98%)	126 (99%)	1 (1%)	0	100	100
All	All	597/648 (92%)	580 (97%)	16 (3%)	1 (0%)	47	68

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	38	VAL

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	AAA	148/165 (90%)	143 (97%)	5 (3%)	37	63
1	BBB	151/165 (92%)	146 (97%)	5 (3%)	38	64
2	CCC	106/108 (98%)	102 (96%)	4 (4%)	33	58
2	DDD	107/108 (99%)	102 (95%)	5 (5%)	26	49
All	All	512/546 (94%)	493 (96%)	19 (4%)	34	60

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	DDD	127	LEU
2	CCC	88	PRO
2	CCC	89	GLU
2	CCC	52	SER
1	BBB	138	TYR

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

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 1 ligands modelled in this entry, 1 is 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	AAA	174/194 (89%)	0.03	5 (2%) 51 55	25, 45, 84, 98	0
1	BBB	178/194 (91%)	-0.07	2 (1%) 80 82	23, 39, 70, 103	0
2	CCC	128/130 (98%)	-0.17	0 100 100	23, 36, 64, 82	0
2	DDD	129/130 (99%)	-0.20	0 100 100	20, 37, 72, 110	0
All	All	609/648 (93%)	-0.09	7 (1%) 80 82	20, 40, 74, 110	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	121	ARG	2.9
1	AAA	151	ILE	2.6
1	AAA	39	ASP	2.4
1	AAA	37	GLY	2.3
1	BBB	138	TYR	2.3

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(\AA^2)	Q<0.9
3	CA	CCC	201	1/1	1.00	0.18	23,23,23,23	0

6.5 Other polymers [i](#)

There are no such residues in this entry.