wwPDB X-ray Structure Validation Summary Report (i)

Jan 13, 2024 – 11:51 am GMT

PDB ID	:	6T9D
Title	:	Crystal structure of a bispecific DutaFab in complex with human VEGF121
Authors	:	Kimbung, R.; Logan, D.T.; Beckmann, R.; Jensen, K.; Speck, J.; Fenn, S.; Kettenberger, H.
Deposited on	:	2019-10-28
Resolution	:	2.90 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:

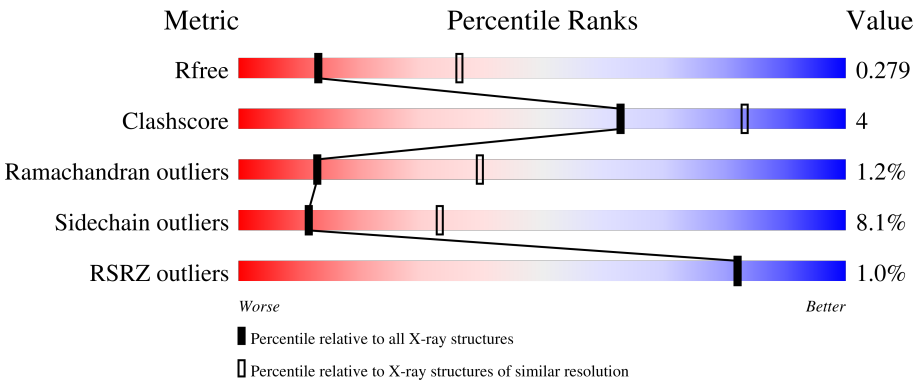
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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 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	AAA	220	% 81% 15% •
1	HHH	220	% 78% 16% • 5%
2	BBB	213	80% 18% •
2	LLL	213	81% 17% ••
3	CCC	121	66% 7% • 25%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	DDD	121	3% 69% 7% • 21%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 8284 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 VP mat DutaFab VH chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	212	1597	1013	264	314	6	0	0	0
1	HHH	210	1586	1007	262	311	6	0	0	0

- Molecule 2 is a protein called VP mat DutaFab VL chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	BBB	213	1678	1056	283	334	5	0	0	0
2	LLL	211	1677	1057	282	333	5	0	2	0

- Molecule 3 is a protein called Vascular endothelial growth factor A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	CCC	91	747	468	124	141	14	0	1	0
3	DDD	96	786	492	133	147	14	0	1	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
CCC	115	ASN	LYS	conflict	UNP P15692
DDD	115	ASN	LYS	conflict	UNP P15692

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	AAA	28	28	28	0	0

Continued on next page...

Continued from previous page...

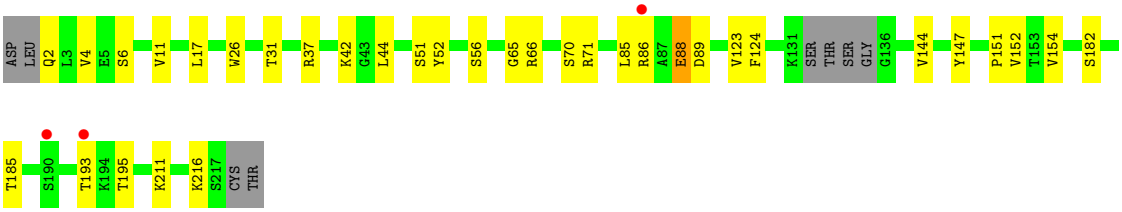
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	BBB	49	49	49	0	0
			Total	O		
4	CCC	25	25	25	0	0
			Total	O		
4	DDD	25	25	25	0	0
			Total	O		
4	HHH	40	40	40	0	0
			Total	O		
4	LLL	46	46	46	0	0

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.

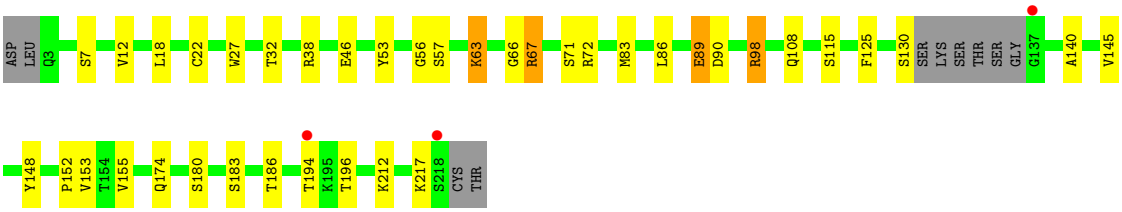
• Molecule 1: VP mat DutaFab VH chain

Chain AAA:



• Molecule 1: VP mat DutaFab VH chain

Chain HHH:

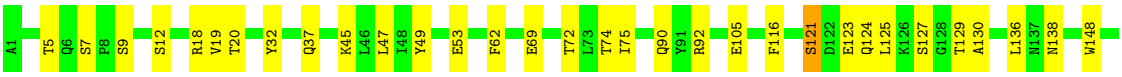


• Molecule 2: VP mat DutaFab VL chain

Chain BBB:

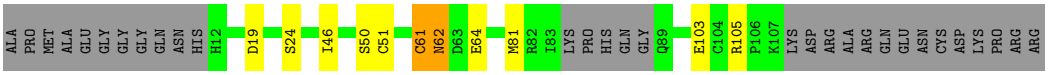
• Molecule 2: VP mat DutaFab VL chain

Chain LLL:



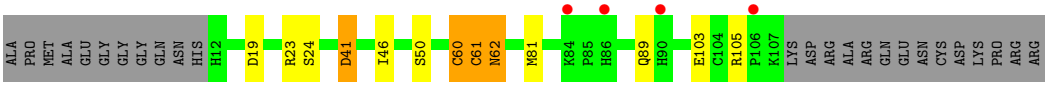
• Molecule 3: Vascular endothelial growth factor A

Chain CCC:



• Molecule 3: Vascular endothelial growth factor A

Chain DDD:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	70.08Å 107.88Å 87.37Å 90.00° 95.34° 90.00°	Depositor
Resolution (Å)	29.29 – 2.90 29.29 – 2.91	Depositor EDS
% Data completeness (in resolution range)	98.9 (29.29-2.90) 99.0 (29.29-2.91)	Depositor EDS
R_{merge}	0.21	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.47 (at 2.90Å)	Xtriage
Refinement program	REFMAC 5.8.0257	Depositor
R, R_{free}	0.212 , 0.280 0.217 , 0.279	Depositor DCC
R_{free} test set	1416 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	37.5	Xtriage
Anisotropy	0.048	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 42.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	8284	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.04% 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	AAA	0.68	0/1638	0.86	0/2231
1	HHH	0.69	0/1627	0.85	0/2216
2	BBB	0.67	0/1721	0.84	0/2336
2	LLL	0.67	0/1726	0.85	0/2343
3	CCC	0.68	0/763	0.81	0/1027
3	DDD	0.67	0/805	0.80	0/1085
All	All	0.68	0/8280	0.84	0/11238

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	AAA	0	1
1	HHH	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	AAA	65	GLY	Peptide
1	HHH	66	GLY	Peptide

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	1597	0	1541	17	0
1	HHH	1586	0	1534	17	0
2	BBB	1678	0	1609	17	0
2	LLL	1677	0	1614	17	0
3	CCC	747	0	704	4	0
3	DDD	786	0	742	7	0
4	AAA	28	0	0	2	0
4	BBB	49	0	0	2	0
4	CCC	25	0	0	0	0
4	DDD	25	0	0	0	0
4	HHH	40	0	0	1	0
4	LLL	46	0	0	0	0
All	All	8284	0	7744	68	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 4.

The worst 5 of 68 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:BBB:143:GLU:HB2	4:BBB:314:HOH:O	1.76	0.86
1:AAA:88:GLU:HG3	4:AAA:305:HOH:O	1.99	0.61
1:HHH:32:THR:OG1	1:HHH:98:ARG:HD3	2.03	0.59
2:LLL:123[A]:GLU:H	2:LLL:123[A]:GLU:CD	2.08	0.57
2:LLL:18:ARG:NH1	2:LLL:74:THR:HG21	2.19	0.57

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	208/220 (94%)	197 (95%)	8 (4%)	3 (1%)	11	36
1	HHH	206/220 (94%)	194 (94%)	9 (4%)	3 (2%)	10	34
2	BBB	211/213 (99%)	199 (94%)	11 (5%)	1 (0%)	29	61
2	LLL	211/213 (99%)	199 (94%)	11 (5%)	1 (0%)	29	61
3	CCC	88/121 (73%)	84 (96%)	3 (3%)	1 (1%)	14	42
3	DDD	95/121 (78%)	87 (92%)	5 (5%)	3 (3%)	4	16
All	All	1019/1108 (92%)	960 (94%)	47 (5%)	12 (1%)	13	40

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	DDD	41	ASP
1	HHH	27	TRP
1	AAA	26	TRP
1	AAA	88	GLU
3	CCC	62	ASN

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	176/184 (96%)	167 (95%)	9 (5%)	24	56
1	HHH	175/184 (95%)	160 (91%)	15 (9%)	10	30
2	BBB	188/188 (100%)	168 (89%)	20 (11%)	6	20
2	LLL	189/188 (100%)	176 (93%)	13 (7%)	15	41
3	CCC	88/110 (80%)	80 (91%)	8 (9%)	9	28
3	DDD	92/110 (84%)	83 (90%)	9 (10%)	8	24
All	All	908/964 (94%)	834 (92%)	74 (8%)	11	33

5 of 74 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	HHH	152	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	LLL	168	SER
1	HHH	194	THR
2	LLL	20	THR
2	BBB	181	LEU

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)

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	AAA	212/220 (96%)	-0.02	3 (1%)	75	75	22, 38, 63, 81	0
1	HHH	210/220 (95%)	-0.05	3 (1%)	75	75	20, 35, 60, 75	0
2	BBB	213/213 (100%)	-0.21	0	100	100	20, 36, 51, 72	0
2	LLL	211/213 (99%)	-0.28	0	100	100	19, 34, 50, 61	0
3	CCC	91/121 (75%)	0.07	0	100	100	22, 42, 68, 78	0
3	DDD	96/121 (79%)	0.30	4 (4%)	36	32	24, 42, 66, 75	0
All	All	1033/1108 (93%)	-0.08	10 (0%)	82	82	19, 36, 60, 81	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	DDD	86	HIS	2.7
1	AAA	190	SER	2.5
3	DDD	106	PRO	2.5
1	AAA	193	THR	2.4
3	DDD	90	HIS	2.4

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.