



wwPDB X-ray Structure Validation Summary Report ⓘ

May 21, 2020 – 08:05 am BST

PDB ID : 5D1Q
Title : IsdB NEAT2 bound by clone D2-06
Authors : Deng, X.
Deposited on : 2015-08-04
Resolution : 3.22 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 : 2.11
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

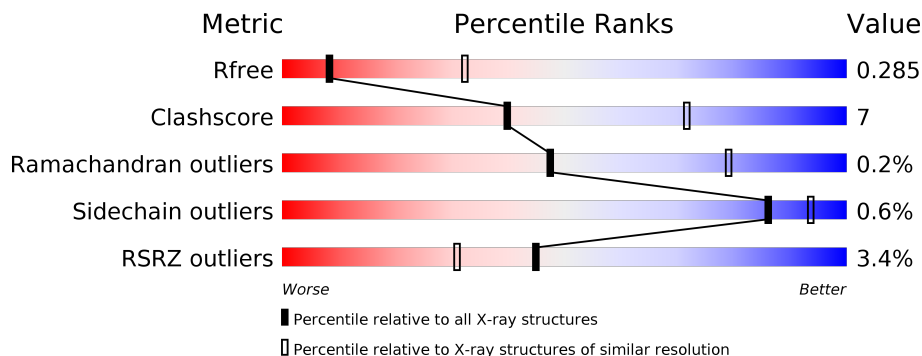
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.22 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	1335 (3.24-3.20)
Clashscore	141614	1460 (3.24-3.20)
Ramachandran outliers	138981	1437 (3.24-3.20)
Sidechain outliers	138945	1436 (3.24-3.20)
RSRZ outliers	127900	1291 (3.24-3.20)

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 $\leq 5\%$. 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	214	 91% 9%
2	B	261	 3% 70% 11% 17%
3	C	274	 4% 74% 10% 16%
4	D	216	 6% 81% 17% ..
5	E	120	 92% 6% ..

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 7583 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 D2-06 Light Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	214	1645	1030	276	334	5	0	0	0

- Molecule 2 is a protein called D2-06 Heavy Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	216	1596	1002	270	317	7	0	0	0

- Molecule 3 is a protein called P5 Heavy Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	231	1724	1083	288	347	6	0	0	0

- Molecule 4 is a protein called P5 Light Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	214	1641	1029	277	330	5	0	0	0

- Molecule 5 is a protein called Iron-regulated surface determinant protein B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	118	977	621	161	187	8	0	1	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	339	PRO	-	expression tag	UNP Q2YX96
E	340	GLY	-	expression tag	UNP Q2YX96

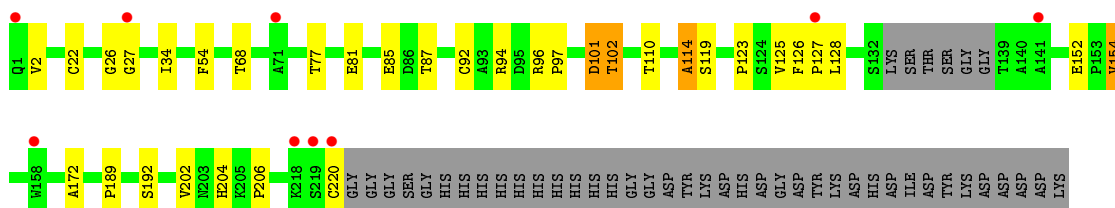
3 Residue-property plots [i](#)

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.

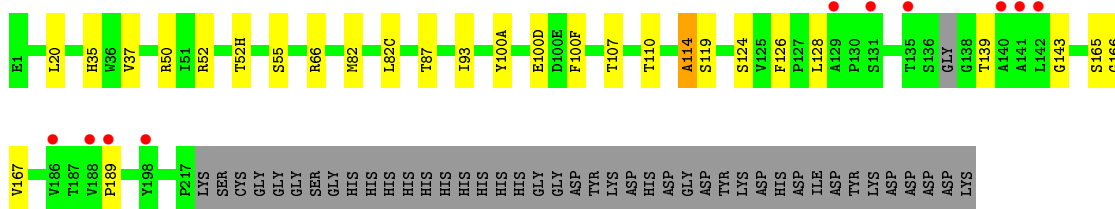
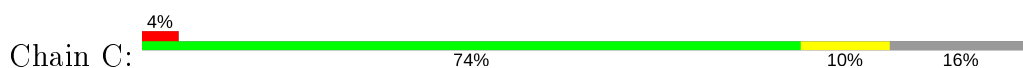
- Molecule 1: D2-06 Light Chain



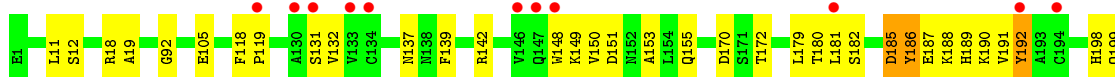
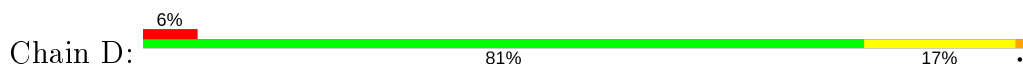
- Molecule 2: D2-06 Heavy Chain



- Molecule 3: P5 Heavy Chain



- Molecule 4: P5 Light Chain





- Molecule 5: Iron-regulated surface determinant protein B

Chain E: 92% 6% ..



4 Data and refinement statistics

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, α , β , γ	121.39Å 121.39Å 193.33Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.33 – 3.22 47.33 – 3.22	Depositor EDS
% Data completeness (in resolution range)	98.8 (47.33-3.22) 98.9 (47.33-3.22)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.27 (at 3.19Å)	Xtrriage
Refinement program	REFMAC 5.8.0123	Depositor
R, R_{free}	0.239 , 0.288 0.237 , 0.285	Depositor DCC
R_{free} test set	1242 reflections (5.19%)	wwPDB-VP
Wilson B-factor (Å ²)	88.1	Xtrriage
Anisotropy	0.046	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 16.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	7583	wwPDB-VP
Average B, all atoms (Å ²)	74.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.58% 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](#)

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	0/1682	0.66	0/2287
2	B	0.74	1/1632 (0.1%)	0.79	4/2225 (0.2%)
3	C	0.77	2/1762 (0.1%)	0.71	4/2399 (0.2%)
4	D	0.46	0/1678	0.73	1/2278 (0.0%)
5	E	0.40	0/996	0.70	0/1340
All	All	0.59	3/7750 (0.0%)	0.72	9/10529 (0.1%)

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
2	B	0	2
3	C	0	2
4	D	0	2
All	All	0	6

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	114	ALA	C-N	25.63	1.93	1.34
2	B	114	ALA	C-N	22.24	1.85	1.34
3	C	166	GLY	N-CA	6.88	1.56	1.46

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	154	VAL	CB-CA-C	7.04	124.77	111.40
3	C	166	GLY	N-CA-C	6.04	128.20	113.10
4	D	205	VAL	CB-CA-C	-6.02	99.96	111.40
2	B	102	THR	N-CA-CB	5.51	120.78	110.30
3	C	114	ALA	CA-C-N	-5.51	105.07	117.20

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	101	ASP	Peptide
2	B	154	VAL	Peptide
3	C	165	SER	Peptide
3	C	167	VAL	Peptide
4	D	142	ARG	Peptide

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	1645	0	1592	16	0
2	B	1596	0	1573	30	0
3	C	1724	0	1685	18	0
4	D	1641	0	1598	50	0
5	E	977	0	970	9	0
All	All	7583	0	7418	110	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 7.

The worst 5 of 110 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)
4:D:181:LEU:HD13	4:D:186:TYR:CB	1.56	1.34
2:B:114:ALA:C	2:B:119:SER:N	1.85	1.29
3:C:114:ALA:C	3:C:119:SER:N	1.93	1.21
1:A:214:CYS:HB3	2:B:220:CYS:SG	1.85	1.14
4:D:181:LEU:HD13	4:D:186:TYR:HB2	1.18	1.10

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	212/214 (99%)	201 (95%)	11 (5%)	0	100	100
2	B	212/261 (81%)	197 (93%)	14 (7%)	1 (0%)	29	66
3	C	227/274 (83%)	212 (93%)	15 (7%)	0	100	100
4	D	212/216 (98%)	190 (90%)	21 (10%)	1 (0%)	29	66
5	E	117/120 (98%)	111 (95%)	6 (5%)	0	100	100
All	All	980/1085 (90%)	911 (93%)	67 (7%)	2 (0%)	47	79

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	185	ASP
2	B	127	PRO

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	188/188 (100%)	187 (100%)	1 (0%)	88	94
2	B	180/216 (83%)	179 (99%)	1 (1%)	86	93
3	C	195/230 (85%)	195 (100%)	0	100	100
4	D	185/187 (99%)	183 (99%)	2 (1%)	73	88
5	E	110/110 (100%)	108 (98%)	2 (2%)	59	81
All	All	858/931 (92%)	852 (99%)	6 (1%)	86	93

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	186	TYR
5	E	362[B]	MET
4	D	192	TYR
2	B	85	GLU
5	E	362[A]	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	38	GLN
2	B	39	GLN
2	B	208	ASN
4	D	138	ASN
4	D	155	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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	B	1
3	C	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	114:ALA	C	119:SER	N	1.93
1	B	114:ALA	C	119:SER	N	1.85

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%)	0.05	3 (1%) 75 64	51, 65, 106, 118	0
2	B	216/261 (82%)	0.22	9 (4%) 36 24	55, 78, 107, 123	0
3	C	231/274 (84%)	0.13	10 (4%) 35 23	54, 68, 112, 129	0
4	D	214/216 (99%)	0.21	12 (5%) 24 14	49, 75, 127, 142	0
5	E	118/120 (98%)	-0.05	0 100 100	51, 63, 80, 99	0
All	All	993/1085 (91%)	0.13	34 (3%) 45 31	49, 69, 115, 142	0

The worst 5 of 34 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	146	VAL	3.9
2	B	71	ALA	3.4
3	C	141	ALA	3.2
3	C	186	VAL	3.2
3	C	140	ALA	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.