



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 13, 2017 – 01:26 am GMT

PDB ID : 2BKG
Title : Crystal structure of E3_19 a designed ankyrin repeat protein
Authors : Binz, H.K.; Kohl, A.; Pluckthun, A.; Grutter, M.G.
Deposited on : 2005-02-16
Resolution : 1.90 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/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) : 1.9-1692
EDS : trunk28620
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : recalc28949

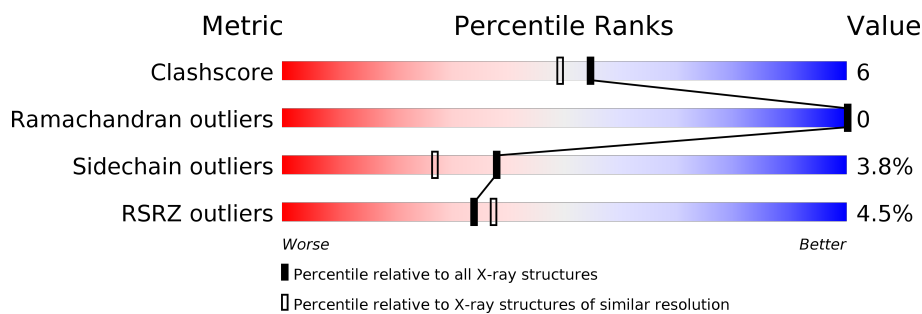
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.90 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))
Clashscore	112137	5731 (1.90-1.90)
Ramachandran outliers	110173	5669 (1.90-1.90)
Sidechain outliers	110143	5670 (1.90-1.90)
RSRZ outliers	101464	5100 (1.90-1.90)

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. 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	166	
1	B	166	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2660 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 SYNTHETIC CONSTRUCT ANKYRIN REPEAT PROTEIN E3_19.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	155	Total	C	N	O	S	0	5	1
			1169	730	199	238	2			
1	B	155	Total	C	N	O	S	0	3	1
			1161	724	197	239	1			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	158	Total	O	0	0
			158	158		
2	B	172	Total	O	0	0
			172	172		

● Molecule 1: SYNTHETIC CONSTRUCT ANKYRIN REPEAT PROTEIN E3 19



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	49.58Å 72.67Å 83.33Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.00 – 1.90 24.79 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.9 (25.00-1.90) 99.9 (24.79-1.90)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.18 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.176 , 0.227 0.189 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	17.6	Xtriage
Anisotropy	0.276	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 48.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	2660	wwPDB-VP
Average B, all atoms (Å ²)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.42% 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

5.1 Standard geometry

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	0.93	0/1208	1.04	7/1638 (0.4%)
1	B	0.92	0/1189	1.10	9/1615 (0.6%)
All	All	0.93	0/2397	1.07	16/3253 (0.5%)

There are no bond length outliers.

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	28	ASP	CB-CG-OD2	11.64	128.78	118.30
1	B	28[A]	ASP	CB-CG-OD2	10.03	127.33	118.30
1	B	28[B]	ASP	CB-CG-OD2	10.03	127.33	118.30
1	A	143	ASP	CB-CG-OD2	7.94	125.44	118.30
1	B	39	ASP	CB-CG-OD2	7.30	124.87	118.30
1	B	160	ASP	CB-CG-OD2	6.65	124.29	118.30
1	B	90	ARG	NE-CZ-NH2	-6.65	116.97	120.30
1	B	143	ASP	CB-CG-OD2	6.54	124.19	118.30
1	A	27	ASP	CB-CG-OD1	5.97	123.68	118.30
1	A	90	ARG	NE-CZ-NH1	5.84	123.22	120.30
1	A	110	ASP	CB-CG-OD1	5.72	123.45	118.30
1	A	90	ARG	NE-CZ-NH2	-5.43	117.59	120.30
1	B	110	ASP	CB-CG-OD1	5.20	122.98	118.30
1	B	122	ASP	CB-CG-OD2	5.17	122.96	118.30
1	B	105	ASP	CB-CG-OD1	5.17	122.95	118.30
1	A	39	ASP	CB-CG-OD1	5.01	122.81	118.30

There are no chirality outliers.

There are no planarity outliers.

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	1169	0	1152	19	0
1	B	1161	0	1140	9	0
2	A	158	0	0	7	1
2	B	172	0	0	5	1
All	All	2660	0	2292	28	1

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 6.

All (28) 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:A:150:PHE:CE2	1:A:154:ILE:HD11	2.08	0.89
1:A:150:PHE:CZ	1:A:154:ILE:HD11	2.21	0.76
1:A:23:ARG:NH2	2:A:2004:HOH:O	2.19	0.72
1:A:159[A]:GLU:OE2	2:A:2156:HOH:O	2.08	0.71
1:B:53:LEU:O	1:B:57[B]:VAL:HG22	1.98	0.63
1:A:21:ALA:HA	1:A:26[B]:GLN:HE21	1.63	0.62
1:A:129:VAL:HG21	1:A:161:LEU:HD21	1.82	0.62
1:B:89:LYS:NZ	2:B:2092:HOH:O	2.35	0.60
1:B:20:GLU:HG2	2:B:2007:HOH:O	2.04	0.56
1:A:121:ALA:HA	1:A:161:LEU:HD11	1.88	0.55
1:A:21:ALA:HA	1:A:26[B]:GLN:NE2	2.21	0.54
1:A:127:GLU:OE1	2:A:2124:HOH:O	2.18	0.54
1:B:125:HIS:HE1	2:B:2090:HOH:O	1.91	0.53
1:A:31:ARG:HD2	2:A:2019:HOH:O	2.08	0.53
1:A:147:LYS:NZ	1:A:155:ASP:OD1	2.37	0.51
1:A:68:LYS:HE2	2:A:2062:HOH:O	2.10	0.51
1:A:125:HIS:HE1	2:A:2088:HOH:O	1.94	0.51
1:B:12:SER:HB3	1:B:15:GLY:H	1.77	0.48
1:A:31:ARG:NH2	2:A:2018:HOH:O	2.46	0.47
1:A:42:ALA:O	1:A:49:THR:HA	2.14	0.47
1:B:31:ARG:NH1	2:B:2019:HOH:O	2.46	0.47
1:B:100:LEU:HD13	1:B:135:TYR:CD1	2.50	0.46
1:A:14:LEU:HB3	1:A:38:ALA:HB2	1.99	0.45
1:A:150:PHE:CE2	1:A:154:ILE:CD1	2.92	0.42
1:B:93:LEU:O	1:B:97:GLU:HG3	2.19	0.42
1:B:97:GLU:OE2	2:B:2100:HOH:O	2.22	0.42
1:A:42:ALA:O	1:A:50:PRO:HD3	2.20	0.41

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:2020:HOH:O	2:B:2134:HOH:O[4_556]	2.15	0.05

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	158/166 (95%)	157 (99%)	1 (1%)	0	100	100
1	B	156/166 (94%)	154 (99%)	2 (1%)	0	100	100
All	All	314/332 (95%)	311 (99%)	3 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	125/130 (96%)	120 (96%)	5 (4%)	36	25
1	B	123/130 (95%)	118 (96%)	5 (4%)	35	24
All	All	248/260 (95%)	238 (96%)	10 (4%)	38	25

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	12	SER
1	A	16	LYS
1	A	90	ARG
1	A	144	LYS
1	A	161	LEU
1	B	61[A]	GLU
1	B	61[B]	GLU
1	B	112	ILE
1	B	126	LEU
1	B	144	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	59	HIS
1	A	92	HIS
1	B	26	GLN
1	B	36	ASN
1	B	59	HIS
1	B	92	HIS

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

There are no such residues in this entry.

5.8 Polymer linkage issues

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	155/166 (93%)	0.21	8 (5%) 28 31	8, 17, 35, 50	0
1	B	155/166 (93%)	0.07	6 (3%) 40 44	8, 17, 34, 43	0
All	All	310/332 (93%)	0.14	14 (4%) 34 37	8, 17, 35, 50	0

All (14) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	12	SER	4.2
1	A	163	GLU	3.5
1	B	37	GLY	3.4
1	B	163	GLU	3.4
1	B	13	ASP	3.0
1	A	13	ASP	2.9
1	A	12	SER	2.8
1	A	14	LEU	2.7
1	B	160	ASP	2.7
1	B	35	ALA	2.4
1	A	160	ASP	2.4
1	A	166	GLN	2.4
1	A	159[A]	GLU	2.3
1	A	164	ILE	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 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.