wwPDB X-ray Structure Validation Summary Report (i)

May 13, 2020 – 09:42 am BST

PDB ID	:	4GK0
Title	:	Crystal structure of human Rev3-Rev7-Rev1 complex
Authors	:	Tao, J.; Min, X.; Wei, X.
Deposited on	:	2012-08-10
Resolution	:	2.70 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 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	:	2.11
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 2.70 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.
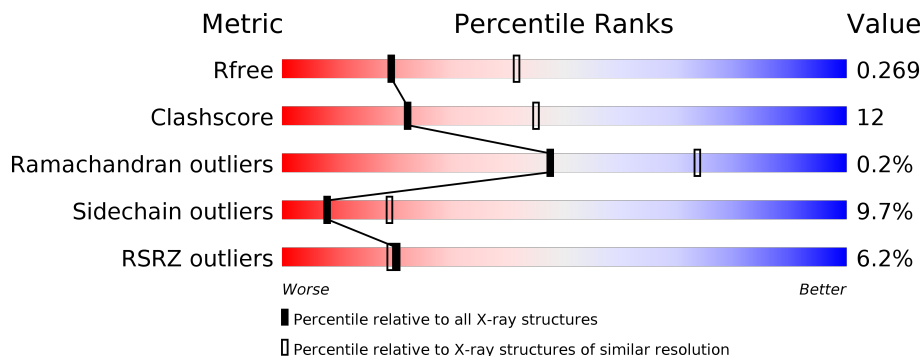
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 <=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	238	4% / 61% 18% · 17%
1	B	238	4% / 58% 18% · 20%
2	C	52	2% / 21% 15% · 62%
2	D	52	6% / 21% 17% 62%
3	E	136	% / 52% 17% · 29%
3	F	136	10% / 43% 22% · 32%

2   Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4987 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 Mitotic spindle assembly checkpoint protein MAD2B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	198	Total	C	N	O	S	0	0	0
			1605	1037	270	288	10			
1	B	190	Total	C	N	O	S	0	0	0
			1546	999	261	276	10			

There are 56 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-15	MET	-	EXPRESSION TAG	UNP Q9UI95
A	-14	GLY	-	EXPRESSION TAG	UNP Q9UI95
A	-13	SER	-	EXPRESSION TAG	UNP Q9UI95
A	-12	SER	-	EXPRESSION TAG	UNP Q9UI95
A	-11	HIS	-	EXPRESSION TAG	UNP Q9UI95
A	-10	HIS	-	EXPRESSION TAG	UNP Q9UI95
A	-9	HIS	-	EXPRESSION TAG	UNP Q9UI95
A	-8	HIS	-	EXPRESSION TAG	UNP Q9UI95
A	-7	HIS	-	EXPRESSION TAG	UNP Q9UI95
A	-6	HIS	-	EXPRESSION TAG	UNP Q9UI95
A	-5	SER	-	EXPRESSION TAG	UNP Q9UI95
A	-4	GLN	-	EXPRESSION TAG	UNP Q9UI95
A	-3	ASP	-	EXPRESSION TAG	UNP Q9UI95
A	-2	PRO	-	EXPRESSION TAG	UNP Q9UI95
A	-1	ASN	-	EXPRESSION TAG	UNP Q9UI95
A	0	SER	-	EXPRESSION TAG	UNP Q9UI95
A	124	ALA	ARG	ENGINEERED MUTATION	UNP Q9UI95
A	212	GLY	-	EXPRESSION TAG	UNP Q9UI95
A	213	SER	-	EXPRESSION TAG	UNP Q9UI95
A	214	GLY	-	EXPRESSION TAG	UNP Q9UI95
A	215	SER	-	EXPRESSION TAG	UNP Q9UI95
A	216	GLY	-	EXPRESSION TAG	UNP Q9UI95
A	217	SER	-	EXPRESSION TAG	UNP Q9UI95
A	218	GLY	-	EXPRESSION TAG	UNP Q9UI95
A	219	SER	-	EXPRESSION TAG	UNP Q9UI95

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	220	GLY	-	EXPRESSION TAG	UNP Q9UI95
A	221	SER	-	EXPRESSION TAG	UNP Q9UI95
A	222	HIS	-	EXPRESSION TAG	UNP Q9UI95
B	-15	MET	-	EXPRESSION TAG	UNP Q9UI95
B	-14	GLY	-	EXPRESSION TAG	UNP Q9UI95
B	-13	SER	-	EXPRESSION TAG	UNP Q9UI95
B	-12	SER	-	EXPRESSION TAG	UNP Q9UI95
B	-11	HIS	-	EXPRESSION TAG	UNP Q9UI95
B	-10	HIS	-	EXPRESSION TAG	UNP Q9UI95
B	-9	HIS	-	EXPRESSION TAG	UNP Q9UI95
B	-8	HIS	-	EXPRESSION TAG	UNP Q9UI95
B	-7	HIS	-	EXPRESSION TAG	UNP Q9UI95
B	-6	HIS	-	EXPRESSION TAG	UNP Q9UI95
B	-5	SER	-	EXPRESSION TAG	UNP Q9UI95
B	-4	GLN	-	EXPRESSION TAG	UNP Q9UI95
B	-3	ASP	-	EXPRESSION TAG	UNP Q9UI95
B	-2	PRO	-	EXPRESSION TAG	UNP Q9UI95
B	-1	ASN	-	EXPRESSION TAG	UNP Q9UI95
B	0	SER	-	EXPRESSION TAG	UNP Q9UI95
B	124	ALA	ARG	ENGINEERED MUTATION	UNP Q9UI95
B	212	GLY	-	EXPRESSION TAG	UNP Q9UI95
B	213	SER	-	EXPRESSION TAG	UNP Q9UI95
B	214	GLY	-	EXPRESSION TAG	UNP Q9UI95
B	215	SER	-	EXPRESSION TAG	UNP Q9UI95
B	216	GLY	-	EXPRESSION TAG	UNP Q9UI95
B	217	SER	-	EXPRESSION TAG	UNP Q9UI95
B	218	GLY	-	EXPRESSION TAG	UNP Q9UI95
B	219	SER	-	EXPRESSION TAG	UNP Q9UI95
B	220	GLY	-	EXPRESSION TAG	UNP Q9UI95
B	221	SER	-	EXPRESSION TAG	UNP Q9UI95
B	222	HIS	-	EXPRESSION TAG	UNP Q9UI95

- Molecule 2 is a protein called DNA polymerase zeta catalytic subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	20	Total	C	N	O	S	0	0	0
			151	95	25	29	2			
2	D	20	Total	C	N	O	S	0	0	0
			151	95	25	29	2			

- Molecule 3 is a protein called DNA repair protein REV1.

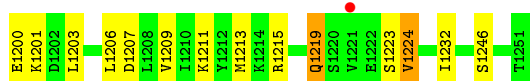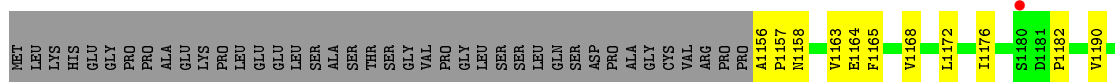| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	96	Total	C	N	O	S	0	0	0
			778	500	121	152	5			
3	F	93	Total	C	N	O	S	0	0	0
			756	487	118	146	5			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	1116	MET	-	EXPRESSION TAG	UNP Q9UBZ9
F	1116	MET	-	EXPRESSION TAG	UNP Q9UBZ9

3 Residue-property plots ⓘ

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: Mitotic spindle assembly checkpoint protein MAD2B

- Molecule 1: Mitotic spindle assembly checkpoint protein MAD2B

- Molecule 2: DNA polymerase zeta catalytic subunit

- Molecule 2: DNA polymerase zeta catalytic subunit

MET LEU THR PRO THR PRO ASP SER SER PRO ARG SER SER THR SER SER PRO SER GLN SER LYS ASN GLY SER PHE THR PRO ARG **T1874 A1875 N1876** R1887 E1888 E1889 I1890 M1891 A1892 T1893 LEU LEU ASP HIS ASP

- Molecule 3: DNA repair protein REV1

%

Chain E: 52% 17% · 29%

MET LEU LYS HIS GLU GLY PRO PRO ALA GLU LYS PRO LEU GLU GLU LEU SER ALA SER THR SER GLY VAL PRO GLY LEU SER SER LEU GLN SER ASP PRO ALA GLY CYS VAL ARG PRO PRO **A1156 P1157 N1158** V1163 E1164 F1165 V1168 L1172 I1176 S1180 D1181 P1182 V1190

E1200 K1201 D1202 L1203 L1206 D1207 L1208 V1209 I1210 K1211 Y1212 M1213 K1214 R1215 Q1219 S1220 V1221 E1222 S1223 V1224 I1232 S1246 T1251

- Molecule 3: DNA repair protein REV1

10%

Chain F: 43% 22% · 32%

MET LEU LYS HIS GLU GLY PRO PRO ALA GLU LYS PRO LEU GLU GLU LEU SER ALA SER THR SER GLY VAL PRO GLY LEU SER SER LEU GLN SER ASP PRO ALA GLY CYS VAL ARG PRO PRO **A1156 P1157 N1158 L1159 A1160** V1163 E1164 **T1170 L1171 L1172 R1173 E1174 W1175 I1176 T1177 T1178 I1179 S1180**

D1181 E1185 D1186 I1187 L1188 Q1189 V1190 Y1193 I1198 E1199 E1200 K1201 D1202 L1203 V1209 I1210 M1213 K1214 M1217 Q1218 Q1219 SER VAL GLU S1223 V1224 W1225 N1226 M1227 Q1237 L1240 Q1241 G1245 S1246 T1247 V1250 T1251

4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	121.09Å 71.89Å 106.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.70 35.19 – 2.70	Depositor EDS
% Data completeness (in resolution range)	95.6 (20.00-2.70) 95.6 (35.19-2.70)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	3.07 (at 2.68Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.218 , 0.273 0.215 , 0.269	Depositor DCC
R_{free} test set	1271 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å²)	53.1	Xtriage
Anisotropy	0.364	Xtriage
Bulk solvent k_{sol}(e/Å³), B_{sol}(Å²)	0.37 , 52.3	EDS
L-test for twinning[2]	$<|L|> = 0.49$, $<L^2> = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	4987	wwPDB-VP
Average B, all atoms (Å²)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.39% 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 ⓘ

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	0.61	0/1640	0.83	1/2228 (0.0%)
1	B	0.50	0/1579	0.70	0/2144
2	C	0.57	0/153	0.81	0/207
2	D	0.54	0/153	0.75	0/207
3	E	0.64	0/789	0.83	0/1070
3	F	0.49	0/766	0.72	0/1037
All	All	0.56	0/5080	0.77	1/6893 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	74	LEU	CA-CB-CG	5.08	126.98	115.30

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	1605	0	1640	35	0
1	B	1546	0	1577	42	0
2	C	151	0	160	7	0
2	D	151	0	160	8	0
3	E	778	0	789	20	0
3	F	756	0	768	22	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	4987	0	5094	118	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 12.

The worst 5 of 118 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:B:153:ARG:HG3	1:B:153:ARG:HH11	1.25	0.99
1:B:207:ALA:O	1:B:208:HIS:HB3	1.70	0.92
3:E:1172:LEU:HD13	3:E:1213:MET:CE	2.07	0.83
2:C:1874:THR:OG1	2:C:1875:ALA:N	2.12	0.82
3:F:1237:GLN:HG3	3:F:1247:THR:HG22	1.69	0.74

There are no symmetry-related clashes.

5.3　Torsion angles ⓘ

5.3.1　Protein backbone ⓘ

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	194/238 (82%)	186 (96%)	8 (4%)	0	100	100
1	B	186/238 (78%)	172 (92%)	14 (8%)	0	100	100
2	C	18/52 (35%)	14 (78%)	4 (22%)	0	100	100
2	D	18/52 (35%)	14 (78%)	4 (22%)	0	100	100
3	E	94/136 (69%)	89 (95%)	5 (5%)	0	100	100
3	F	89/136 (65%)	77 (86%)	11 (12%)	1 (1%)	14	34
All	All	599/852 (70%)	552 (92%)	46 (8%)	1 (0%)	47	73

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	F	1250	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	A	186/219 (85%)	167 (90%)	19 (10%)	7	17
1	B	178/219 (81%)	162 (91%)	16 (9%)	9	22
2	C	18/49 (37%)	14 (78%)	4 (22%)	1	2
2	D	18/49 (37%)	17 (94%)	1 (6%)	21	45
3	E	90/123 (73%)	85 (94%)	5 (6%)	21	45
3	F	87/123 (71%)	76 (87%)	11 (13%)	4	10
All	All	577/782 (74%)	521 (90%)	56 (10%)	8	19

5 of 56 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	104	GLN
1	B	164	GLN
3	F	1226	ASN
1	B	115	LEU
1	B	153	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	104	GLN
1	B	118	HIS
1	B	200	GLN
1	B	10	ASN
1	B	139	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 (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	198/238 (83%)	0.03	9 (4%)	33	31	26, 41, 63, 76	0
1	B	190/238 (79%)	0.12	9 (4%)	31	30	38, 59, 76, 80	0
2	C	20/52 (38%)	0.34	1 (5%)	28	27	32, 50, 81, 82	0
2	D	20/52 (38%)	0.50	3 (15%)	2	1	45, 63, 78, 79	0
3	E	96/136 (70%)	-0.05	2 (2%)	63	65	26, 44, 65, 71	0
3	F	93/136 (68%)	0.73	14 (15%)	2	1	47, 68, 80, 81	0
All	All	617/852 (72%)	0.18	38 (6%)	20	19	26, 53, 76, 82	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	1178	THR	5.7
3	F	1160	ALA	5.5
3	F	1179	ILE	4.6
3	E	1221	VAL	3.8
1	A	110	ILE	3.6

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.