Full wwPDB X-ray Structure Validation Report (i)

May 13, 2020 – 10:48 am BST

PDB ID	:	2XV4
Title	:	Structure of Human RPC62 (partial)
Authors	:	Lefevre, S.; Legrand, P.; Fribourg, S.
Deposited on	:	2010-10-22
Resolution	:	2.95 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
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 (i)) were used in the production of this report:

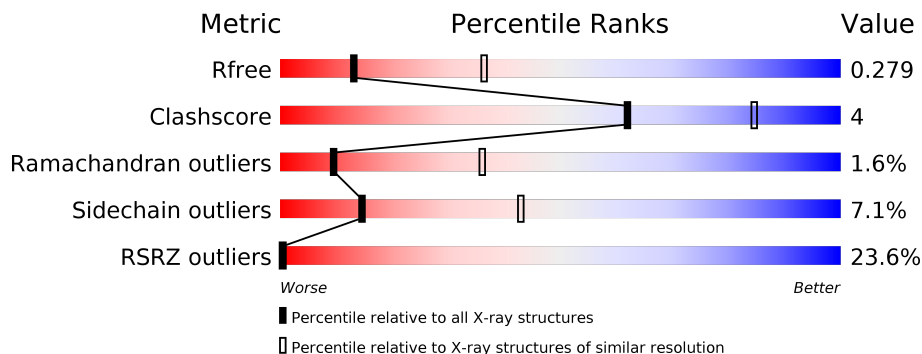|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.95 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	Percentile Ranks	Value
Rfree		0.279
Clashscore		4
Ramachandran outliers		1.6%
Sidechain outliers		7.1%
RSRZ outliers		23.6%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	3104 (3.00-2.92)
Clashscore	141614	3462 (3.00-2.92)
Ramachandran outliers	138981	3340 (3.00-2.92)
Sidechain outliers	138945	3343 (3.00-2.92)
RSRZ outliers	127900	2986 (3.00-2.92)

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	S	534	18% 61% 13% • 25%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 3217 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 DNA-DIRECTED RNA POLYMERASE III SUBUNIT RPC3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	S	399	Total	C	N	O	S	0	0	0
			3184	2014	555	596	19			

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	S	1	Total	O	P	0	0
			4	3	1		
2	S	1	Total	O	P	0	0
			5	4	1		
2	S	1	Total	O	P	0	0
			5	4	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	S	19	Total 19	O 19	0	0

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: DNA-DIRECTED RNA POLYMERASE III SUBUNIT RPC3
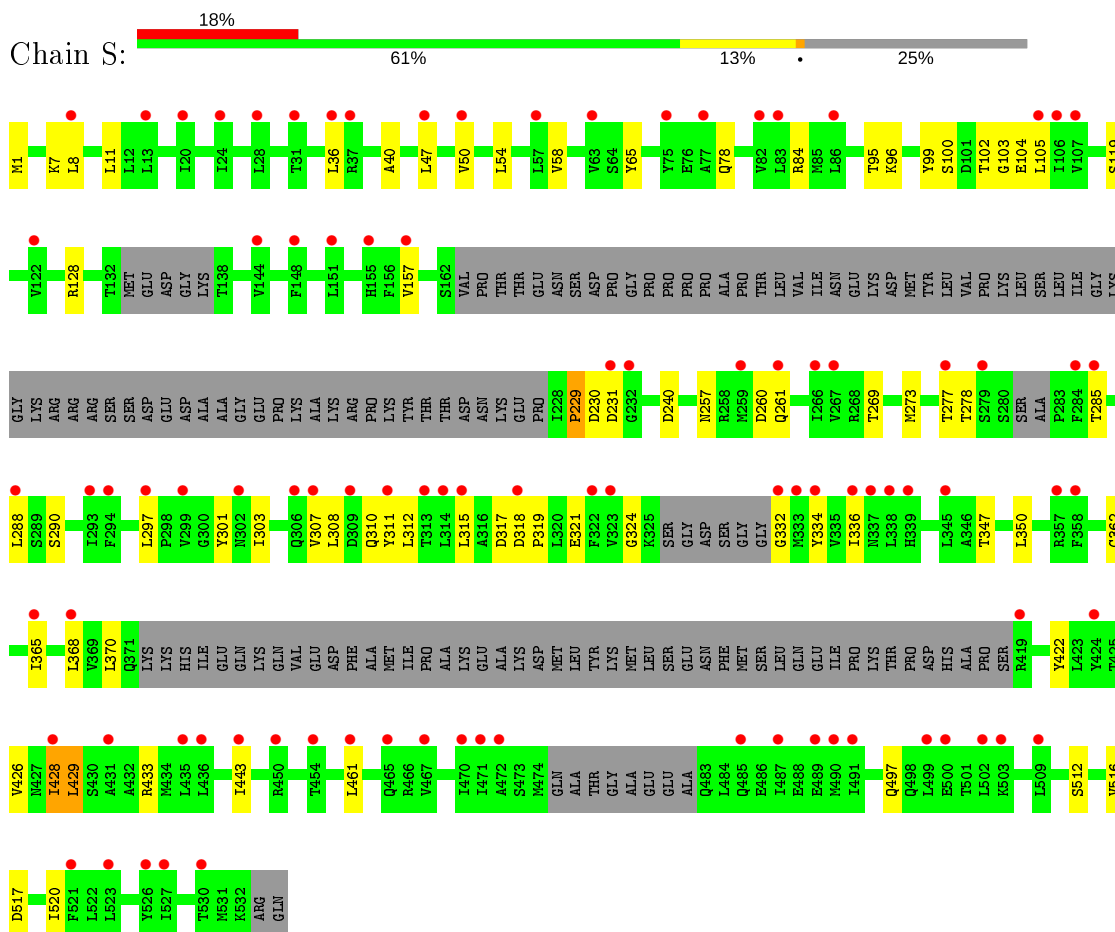
4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	162.54Å 69.92Å 75.97Å 90.00° 112.84° 90.00°	Depositor
Resolution (Å)	70.01 – 2.95 65.34 – 2.95	Depositor EDS
% Data completeness (in resolution range)	(Not available) (70.01-2.95) 99.8 (65.34-2.95)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.75 (at 2.96Å)	Xtriage
Refinement program	BUSTER 2.8.0	Depositor
R, R_{free}	0.228 , 0.256 0.237 , 0.279	Depositor DCC
R_{free} test set	846 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	84.2	Xtriage
Anisotropy	0.557	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 92.7	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	0.93	EDS
Total number of atoms	3217	wwPDB-VP
Average B, all atoms (Å^2)	104.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.31% 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: PO4

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	S	0.51	0/3225	0.71	1/4350 (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	S	229	PRO	N-CA-CB	7.21	111.95	103.30

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	S	3184	0	3222	28	0
2	S	14	0	0	0	0
3	S	19	0	0	1	0
All	All	3217	0	3222	28	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.

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:S:290:SER:HG	1:S:332:GLY:N	1.99	0.60
1:S:429:LEU:HD21	1:S:433:ARG:HH11	1.68	0.58
1:S:347:THR:HA	1:S:350:LEU:HD12	1.88	0.56
1:S:310:GLN:NE2	3:S:2016:HOH:O	2.39	0.52
1:S:54:LEU:O	1:S:58:VAL:HG23	2.09	0.52
1:S:324:GLY:HA2	1:S:334:TYR:HA	1.93	0.50
1:S:317:ASP:OD2	1:S:319:PRO:HD2	2.13	0.49
1:S:260:ASP:O	1:S:261:GLN:HB2	2.14	0.48
1:S:308:LEU:HA	1:S:311:TYR:CD2	2.49	0.48
1:S:285:THR:HB	1:S:336:ILE:H	1.80	0.47
1:S:273:MET:CG	1:S:288:LEU:HD21	2.46	0.46
1:S:362:CYS:HA	1:S:365:ILE:HD12	1.98	0.45
1:S:315:LEU:HD11	1:S:334:TYR:HE1	1.82	0.45
1:S:273:MET:HG3	1:S:288:LEU:HD21	1.99	0.44
1:S:96:LYS:HB2	1:S:104:GLU:HG3	2.00	0.44
1:S:8:LEU:HD11	1:S:443:ILE:HG22	2.00	0.44
1:S:7:LYS:HE3	1:S:11:LEU:HD11	1.99	0.44
1:S:312:LEU:HA	1:S:315:LEU:HD12	1.99	0.44
1:S:105:LEU:HD22	1:S:128:ARG:NH1	2.34	0.43
1:S:517:ASP:HA	1:S:520:ILE:HB	2.00	0.43
1:S:512:SER:O	1:S:516:VAL:HG23	2.18	0.43
1:S:269:THR:HG21	1:S:297:LEU:HD23	2.00	0.42
1:S:443:ILE:HD12	1:S:520:ILE:HD12	2.02	0.41
1:S:99:TYR:HB2	1:S:103:GLY:HA3	2.02	0.41
1:S:315:LEU:HD21	1:S:334:TYR:HD1	1.85	0.40
1:S:428:ILE:H	1:S:428:ILE:HD13	1.86	0.40
1:S:40:ALA:HA	1:S:50:VAL:HG21	2.02	0.40
1:S:96:LYS:O	1:S:99:TYR:O	2.39	0.40

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	S	385/534 (72%)	362 (94%)	17 (4%)	6 (2%)	9	36

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	S	229	PRO
1	S	231	ASP
1	S	422	TYR
1	S	100	SER
1	S	277	THR
1	S	230	ASP

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	S	353/476 (74%)	328 (93%)	25 (7%)	14	43

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	S	1	MET
1	S	36	LEU
1	S	47	LEU
1	S	65	TYR
1	S	78	GLN
1	S	84	ARG
1	S	95	THR
1	S	102	THR
1	S	119	SER
1	S	157	VAL
1	S	240	ASP
1	S	257	ASN
1	S	278	THR
1	S	301	TYR
1	S	303	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	S	307	VAL
1	S	318	ASP
1	S	321	GLU
1	S	368	LEU
1	S	370	LEU
1	S	426	VAL
1	S	428	ILE
1	S	429	LEU
1	S	461	LEU
1	S	497	GLN

Some 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 carbohydrates in this entry.

5.6   Ligand geometry (i)

3 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$\|Z\|>2$	Counts	RMSZ	#$\|Z\|>2$
2	PO4	S	1535	-	4,4,4	2.67	2 (50%)	6,6,6	1.18	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
2	PO4	S	1533	-	0,3,4	0.00	-	0,3,6	0.00	-
2	PO4	S	1534	-	4,4,4	2.13	1 (25%)	6,6,6	0.77	0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	S	1535	PO4	P-O1	4.33	1.61	1.50
2	S	1534	PO4	P-O1	3.40	1.58	1.50
2	S	1535	PO4	P-O2	2.28	1.61	1.54

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	S	399/534 (74%)	1.39	94 (23%)	0	0	62, 98, 153, 176	0

All (94) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	S	334	TYR	6.6
1	S	311	TYR	6.3
1	S	419	ARG	5.6
1	S	323	VAL	5.3
1	S	358	PHE	5.2
1	S	315	LEU	4.9
1	S	490	MET	4.5
1	S	470	ILE	4.4
1	S	500	GLU	4.3
1	S	322	PHE	4.3
1	S	318	ASP	4.3
1	S	461	LEU	4.2
1	S	526	TYR	4.2
1	S	37	ARG	4.0
1	S	313	THR	3.9
1	S	266	ILE	3.9
1	S	428	ILE	3.8
1	S	231	ASP	3.7
1	S	487	ILE	3.7
1	S	314	LEU	3.6
1	S	485	GLN	3.4
1	S	307	VAL	3.2
1	S	306	GLN	3.2
1	S	332	GLY	3.2
1	S	368	LEU	3.2
1	S	232	GLY	3.1
1	S	357	ARG	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	S	151	LEU	3.0
1	S	471	ILE	3.0
1	S	491	ILE	3.0
1	S	345	LEU	3.0
1	S	333	MET	3.0
1	S	499	LEU	2.9
1	S	472	ALA	2.9
1	S	489	GLU	2.9
1	S	293	ILE	2.9
1	S	302	ASN	2.8
1	S	36	LEU	2.8
1	S	338	LEU	2.8
1	S	467	VAL	2.8
1	S	259	MET	2.8
1	S	336	ILE	2.7
1	S	521	PHE	2.7
1	S	502	LEU	2.7
1	S	86	LEU	2.6
1	S	285	THR	2.6
1	S	523	LEU	2.6
1	S	435	LEU	2.6
1	S	106	ILE	2.6
1	S	530	THR	2.5
1	S	122	VAL	2.5
1	S	261	GLN	2.5
1	S	28	LEU	2.5
1	S	297	LEU	2.4
1	S	105	LEU	2.4
1	S	13	LEU	2.4
1	S	20	ILE	2.4
1	S	267	VAL	2.3
1	S	337	ASN	2.3
1	S	299	VAL	2.3
1	S	144	VAL	2.3
1	S	288	LEU	2.3
1	S	443	ILE	2.3
1	S	294	PHE	2.3
1	S	431	ALA	2.3
1	S	465	GLN	2.3
1	S	107	VAL	2.2
1	S	47	LEU	2.2
1	S	77	ALA	2.2

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	S	63	VAL	2.2
1	S	503	LYS	2.2
1	S	309	ASP	2.2
1	S	24	ILE	2.2
1	S	365	ILE	2.2
1	S	57	LEU	2.2
1	S	155	HIS	2.2
1	S	8	LEU	2.1
1	S	148	PHE	2.1
1	S	424	TYR	2.1
1	S	83	LEU	2.1
1	S	277	THR	2.1
1	S	284	PHE	2.1
1	S	454	THR	2.1
1	S	339	HIS	2.1
1	S	75	TYR	2.1
1	S	157	VAL	2.1
1	S	509	LEU	2.0
1	S	450	ARG	2.0
1	S	82	VAL	2.0
1	S	527	ILE	2.0
1	S	279	SER	2.0
1	S	31	THR	2.0
1	S	50	VAL	2.0
1	S	436	LEU	2.0

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)

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(Å²)	Q<0.9
2	PO4	S	1535	5/5	0.83	0.53	103,106,109,112	0
2	PO4	S	1533	4/5	0.85	0.33	119,123,125,126	0
2	PO4	S	1534	5/5	0.85	0.28	143,148,149,149	0

6.5 Other polymers (i)

There are no such residues in this entry.