



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 13, 2017 – 02:00 pm GMT

PDB ID : 4JHP
Title : The crystal structure of the RPGR RCC1-like domain in complex with PDE6D
Authors : Waetzlich, D.; Vetter, I.; Wittinghofer, A.; Ismail, S.
Deposited on : 2013-03-05
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

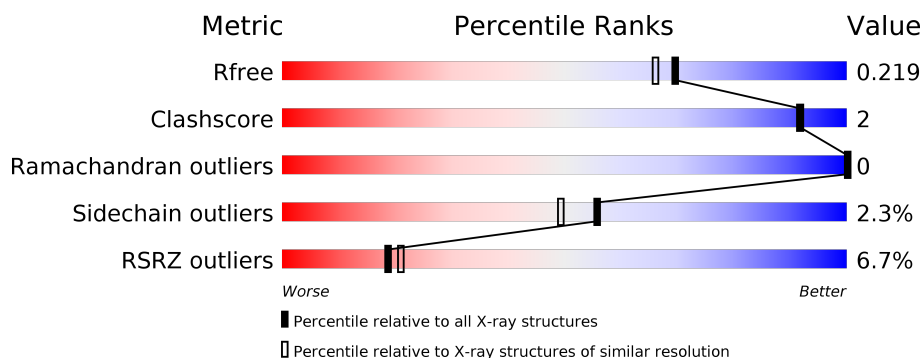
1 Overall quality at a glance

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))
R_{free}	100719	5047 (1.90-1.90)
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	B	152	<div> <div>7%</div> <div> <div></div> <div>89%</div> <div>• • 7%</div> </div> </div>
2	C	391	<div> <div>6%</div> <div> <div></div> <div>87%</div> <div>5% • 7%</div> </div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4336 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 Retinal rod rhodopsin-sensitive cGMP 3',5'-cyclic phosphodiesterase subunit delta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	142	Total	C	N	O	S	0	1	0
			1171	753	197	215	6			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-1	GLY	-	EXPRESSION TAG	UNP O43924
B	0	SER	-	EXPRESSION TAG	UNP O43924

- Molecule 2 is a protein called X-linked retinitis pigmentosa GTPase regulator.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	362	Total	C	N	O	S	3	1	0
			2752	1745	470	521	16			

There are 29 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	369	PRO	-	EXPRESSION TAG	UNP Q92834
C	370	ARG	-	EXPRESSION TAG	UNP Q92834
C	371	ARG	-	EXPRESSION TAG	UNP Q92834
C	372	PRO	-	EXPRESSION TAG	UNP Q92834
C	373	PRO	-	EXPRESSION TAG	UNP Q92834
C	374	ALA	-	EXPRESSION TAG	UNP Q92834
C	375	TYR	-	EXPRESSION TAG	UNP Q92834
C	376	VAL	-	EXPRESSION TAG	UNP Q92834
C	377	GLU	-	EXPRESSION TAG	UNP Q92834
C	378	GLN	-	EXPRESSION TAG	UNP Q92834
C	379	LYS	-	EXPRESSION TAG	UNP Q92834
C	380	LEU	-	EXPRESSION TAG	UNP Q92834
C	381	ILE	-	EXPRESSION TAG	UNP Q92834

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	382	SER	-	EXPRESSION TAG	UNP Q92834
C	383	GLU	-	EXPRESSION TAG	UNP Q92834
C	384	GLU	-	EXPRESSION TAG	UNP Q92834
C	385	ASP	-	EXPRESSION TAG	UNP Q92834
C	386	LEU	-	EXPRESSION TAG	UNP Q92834
C	387	ASN	-	EXPRESSION TAG	UNP Q92834
C	388	SER	-	EXPRESSION TAG	UNP Q92834
C	389	ALA	-	EXPRESSION TAG	UNP Q92834
C	390	VAL	-	EXPRESSION TAG	UNP Q92834
C	391	ASP	-	EXPRESSION TAG	UNP Q92834
C	392	HIS	-	EXPRESSION TAG	UNP Q92834
C	393	HIS	-	EXPRESSION TAG	UNP Q92834
C	394	HIS	-	EXPRESSION TAG	UNP Q92834
C	395	HIS	-	EXPRESSION TAG	UNP Q92834
C	396	HIS	-	EXPRESSION TAG	UNP Q92834
C	397	HIS	-	EXPRESSION TAG	UNP Q92834

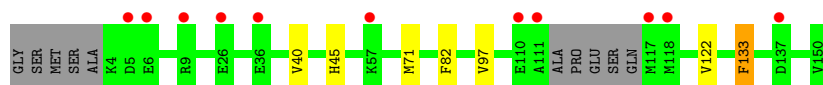
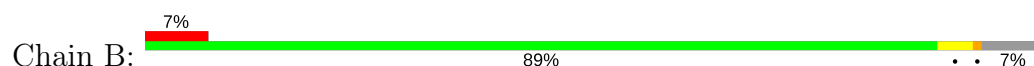
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	102	Total O 102 102	0	0
3	C	311	Total O 311 311	0	0

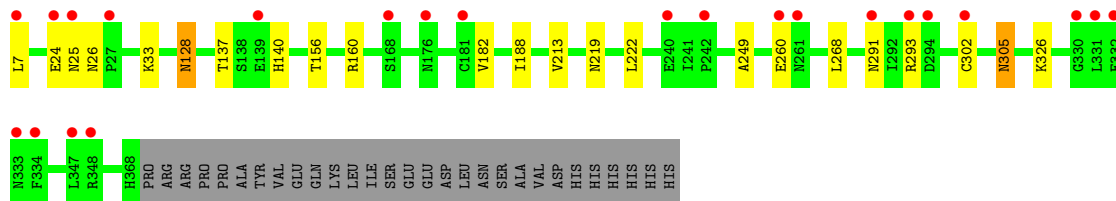
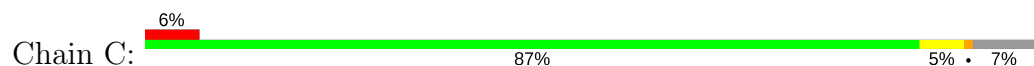
3 Residue-property plots [i](#)

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: Retinal rod rhodopsin-sensitive cGMP 3',5'-cyclic phosphodiesterase subunit delta



- Molecule 2: X-linked retinitis pigmentosa GTPase regulator



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	133.83Å 58.78Å 92.11Å 90.00° 123.32° 90.00°	Depositor
Resolution (Å)	29.39 – 1.90 29.39 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.6 (29.39-1.90) 99.6 (29.39-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.21 (at 1.91Å)	Xtriage
Refinement program	REFMAC 5.6.0093	Depositor
R, R_{free}	0.200 , 0.221 0.199 , 0.219	Depositor DCC
R_{free} test set	2359 reflections (5.27%)	DCC
Wilson B-factor (Å ²)	22.1	Xtriage
Anisotropy	0.279	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 43.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	4336	wwPDB-VP
Average B, all atoms (Å ²)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.74% 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	B	0.33	0/1198	0.46	0/1613
2	C	0.36	0/2819	0.49	0/3819
All	All	0.35	0/4017	0.48	0/5432

There are no bond length outliers.

There are no bond angle outliers.

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	B	1171	0	1163	5	0
2	C	2752	0	2685	10	0
3	B	102	0	0	0	0
3	C	311	0	0	0	0
All	All	4336	0	3848	15	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 2.

All (15) 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:C:25:ASN:HA	2:C:26:ASN:HB2	1.70	0.71

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:137:THR:H	2:C:140:HIS:HD2	1.45	0.65
2:C:249:ALA:HB1	2:C:302[B]:CYS:HB2	1.83	0.60
2:C:305:ASN:HD21	2:C:326:LYS:NZ	2.00	0.59
2:C:156:THR:HG22	2:C:160:ARG:H	1.71	0.55
2:C:219:ASN:HB3	2:C:268:LEU:HG	1.89	0.54
1:B:40:VAL:O	1:B:45:HIS:HE1	1.91	0.54
2:C:305:ASN:HD21	2:C:326:LYS:HZ1	1.59	0.50
2:C:213:VAL:HG21	2:C:222:LEU:O	2.14	0.47
1:B:133:PHE:C	1:B:133:PHE:HD2	2.20	0.46
1:B:133:PHE:C	1:B:133:PHE:CD2	2.90	0.45
2:C:128:ASN:C	2:C:128:ASN:HD22	2.21	0.43
1:B:71:MET:HB2	1:B:97:VAL:HB	2.01	0.42
2:C:24:GLU:HA	2:C:25:ASN:HA	1.80	0.42
1:B:82:PHE:HZ	1:B:122:VAL:HG12	1.85	0.41

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	B	139/152 (91%)	137 (99%)	2 (1%)	0	100	100
2	C	361/391 (92%)	352 (98%)	9 (2%)	0	100	100
All	All	500/543 (92%)	489 (98%)	11 (2%)	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	B	129/136 (95%)	128 (99%)	1 (1%)	85	85
2	C	300/326 (92%)	291 (97%)	9 (3%)	46	37
All	All	429/462 (93%)	419 (98%)	10 (2%)	56	49

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

Mol	Chain	Res	Type
1	B	133	PHE
2	C	7	LEU
2	C	33	LYS
2	C	128	ASN
2	C	182	VAL
2	C	188	ILE
2	C	260	GLU
2	C	291	ASN
2	C	293	ARG
2	C	305	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	18	ASN
1	B	45	HIS
1	B	85	GLN
1	B	126	ASN
2	C	54	ASN
2	C	128	ASN
2	C	140	HIS
2	C	144	GLN
2	C	226	ASN
2	C	247	GLN
2	C	295	GLN
2	C	305	ASN
2	C	333	ASN

5.3.3 RNA ⓘ

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

6.1 Protein, DNA and RNA chains

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	B	142/152 (93%)	0.17	11 (7%) 14 16	19, 26, 45, 57	0
2	C	362/391 (92%)	0.27	23 (6%) 20 23	15, 21, 37, 60	3 (0%)
All	All	504/543 (92%)	0.25	34 (6%) 19 21	15, 23, 40, 60	3 (0%)

All (34) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	331	LEU	9.4
2	C	7	LEU	6.0
2	C	332	GLU	5.1
2	C	302[A]	CYS	5.0
2	C	25	ASN	4.9
2	C	27	PRO	4.8
2	C	334	PHE	4.5
1	B	117	MET	4.3
1	B	111	ALA	4.2
2	C	168	SER	3.7
2	C	24	GLU	3.5
2	C	348	ARG	3.5
2	C	333	ASN	3.3
1	B	57	LYS	3.2
2	C	260	GLU	3.0
2	C	330	GLY	2.8
2	C	294	ASP	2.7
1	B	5	ASP	2.6
2	C	176	ASN	2.6
2	C	293	ARG	2.5
1	B	137	ASP	2.5
2	C	139	GLU	2.5
1	B	118	MET	2.4
2	C	240	GLU	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	6	GLU	2.4
2	C	291	ASN	2.4
1	B	36	GLU	2.4
1	B	9	ARG	2.3
1	B	26	GLU	2.2
2	C	242	PRO	2.2
1	B	110	GLU	2.1
2	C	181	CYS	2.1
2	C	261	ASN	2.0
2	C	347	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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.