Full wwPDB X-ray Structure Validation Report (i)

Apr 11, 2022 – 10:04 am BST

PDB ID	:	7PC5
Title	:	The third PDZ domain of PDZD7 complexed with the PDZ-binding motif of EXOC4
Authors	:	Cousido-Siah, A.; Trave, G.; Gogl, G.
Deposited on	:	2021-08-03
Resolution	:	1.70 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:

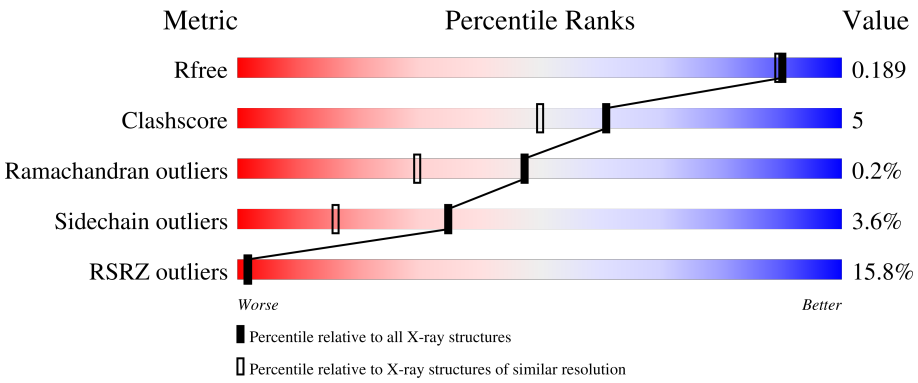
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.27
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

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 1.70 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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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 of 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	420	14%; 85%; 11%; .
2	B	10	50%; 30%; 20%; 10%; 40%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 3882 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 PDZ domain-containing protein 7,Annexin A2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	405	3313	2078	577	644	14	0	12	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	853	GLY	-	expression tag	UNP Q9H5P4
A	854	SER	-	expression tag	UNP Q9H5P4
A	855	HIS	-	expression tag	UNP Q9H5P4
A	856	MET	-	expression tag	UNP Q9H5P4
A	857	GLY	-	expression tag	UNP Q9H5P4
A	954	GLY	-	linker	UNP Q9H5P4
A	999	GLU	ALA	conflict	UNP P07355

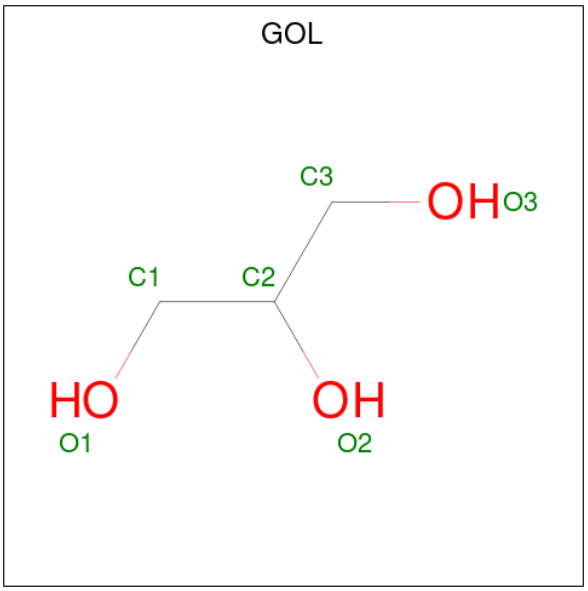
- Molecule 2 is a protein called Exocyst complex component 4.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	6	48	31	8	9	0	0	0

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
3	A	5	5	5	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	A	1	6	3	3	0	0
			Total	C	O		
4	A	1	6	3	3	0	0

- Molecule 5 is water.

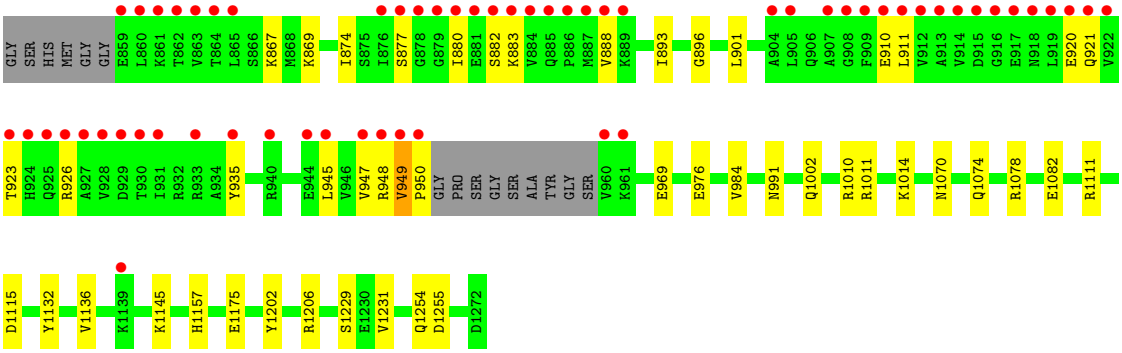
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	502	502	502	0	0
			Total	O		
5	B	2	2	2	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

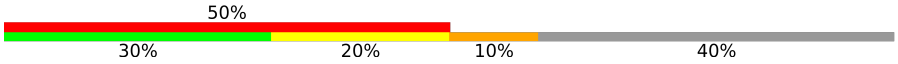
• Molecule 1: PDZ domain-containing protein 7,Annexin A2

Chain A:



• Molecule 2: Exocyst complex component 4

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, α, β, γ	54.11Å 60.35Å 181.68Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.75 – 1.70 42.75 – 1.70	Depositor EDS
% Data completeness (in resolution range)	97.6 (42.75-1.70) 97.6 (42.75-1.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.87 (at 1.70Å)	Xtriage
Refinement program	PHENIX 1.14_3260	Depositor
R, R_{free}	0.166 , 0.188 0.167 , 0.189	Depositor DCC
R_{free} test set	3246 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	24.3	Xtriage
Anisotropy	0.355	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	3882	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.61% 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 (i)

5.1 Standard geometry (i)

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

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	0.64	0/3357	0.70	0/4506
2	B	0.60	0/47	0.78	0/60
All	All	0.64	0/3404	0.70	0/4566

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 (i)

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	3313	0	3369	30	0
2	B	48	0	59	2	0
3	A	5	0	0	0	0
4	A	12	0	15	1	0
5	A	502	0	0	13	0
5	B	2	0	0	0	0
All	All	3882	0	3443	31	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 5.

All (31) 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:1231:VAL:HG23	5:A:1412:HOH:O	1.82	0.79
1:A:1078:ARG:NH1	1:A:1082:GLU:OE1	2.21	0.72
1:A:949:VAL:HG22	1:A:950:PRO:HD2	1.78	0.65
1:A:1115:ASP:HB3	5:A:1630:HOH:O	1.99	0.63
1:A:976:GLU:OE1	1:A:1011[B]:ARG:NH1	2.36	0.53
1:A:1175[B]:GLU:HG3	5:A:1646:HOH:O	2.09	0.52
1:A:869:LYS:HE3	1:A:901:LEU:HD11	1.92	0.51
1:A:910:GLU:HB3	1:A:948:ARG:HB3	1.92	0.51
1:A:1010:ARG:NH1	5:A:1417:HOH:O	2.45	0.50
1:A:1145:LYS:NZ	5:A:1410:HOH:O	2.32	0.50
1:A:888:VAL:HG22	1:A:911:LEU:HD12	1.94	0.50
1:A:1254:GLN:OE1	5:A:1401:HOH:O	2.20	0.50
1:A:1202:TYR:CZ	1:A:1206:ARG:HD2	2.48	0.48
1:A:1014:LYS:HE2	1:A:1014:LYS:HB3	1.64	0.48
1:A:896:GLY:N	5:A:1414:HOH:O	2.40	0.48
1:A:874:ILE:HG22	1:A:893:ILE:HD13	1.95	0.47
1:A:1070:ASN:O	1:A:1074[A]:GLN:HG2	2.14	0.47
1:A:991[A]:ASN:OD1	5:A:1402:HOH:O	2.21	0.46
1:A:888:VAL:HG13	1:A:911:LEU:HB2	1.98	0.45
1:A:877:SER:O	1:A:888:VAL:HA	2.17	0.45
2:B:32:LYS:HA	2:B:32:LYS:HD2	1.57	0.44
1:A:867:LYS:HD2	1:A:935:TYR:HA	2.00	0.43
1:A:969:GLU:OE1	5:A:1403:HOH:O	2.21	0.43
1:A:1132:TYR:CE1	1:A:1136[B]:VAL:HG23	2.54	0.43
1:A:1111:ARG:NH2	5:A:1421:HOH:O	2.52	0.42
1:A:1132:TYR:CD1	1:A:1136[B]:VAL:HG23	2.55	0.42
1:A:1111:ARG:HD2	5:A:1771:HOH:O	2.20	0.41
1:A:1229[B]:SER:OG	5:A:1404:HOH:O	2.22	0.40
1:A:877:SER:HB3	2:B:33:ILE:HD13	2.02	0.40
1:A:1002:GLN:OE1	5:A:1405:HOH:O	2.22	0.40
1:A:1157:HIS:CG	4:A:1307:GOL:H11	2.56	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	A	413/420 (98%)	408 (99%)	4 (1%)	1 (0%)	47	30
2	B	4/10 (40%)	4 (100%)	0	0	100	100
All	All	417/430 (97%)	412 (99%)	4 (1%)	1 (0%)	47	30

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	984	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	364/360 (101%)	353 (97%)	11 (3%)	41	22
2	B	6/9 (67%)	4 (67%)	2 (33%)	0	0
All	All	370/369 (100%)	357 (96%)	13 (4%)	35	17

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	880	ILE
1	A	882	SER
1	A	883	LYS
1	A	920	GLU
1	A	921	GLN
1	A	923	THR
1	A	926	ARG
1	A	945	LEU
1	A	947	VAL
1	A	949	VAL
1	A	1255	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	31	LYS
2	B	32	LYS

Sometimes 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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 5 are monoatomic - leaving 2 for Mogul analysis.

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	$\#\|Z\| > 2$
4	GOL	A	1306	-	5,5,5	1.08	0	5,5,5	1.26	0
4	GOL	A	1307	-	5,5,5	1.08	0	5,5,5	1.03	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	A	1306	-	-	0/4/4/4	-
4	GOL	A	1307	-	-	3/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1307	GOL	O1-C1-C2-C3
4	A	1307	GOL	O1-C1-C2-O2
4	A	1307	GOL	O2-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1307	GOL	1	0

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	405/420 (96%)	0.86	60 (14%)	2	2	17, 28, 100, 127	0
2	B	6/10 (60%)	3.90	5 (83%)	0	0	82, 90, 110, 120	0
All	All	411/430 (95%)	0.91	65 (15%)	2	2	17, 28, 101, 127	0

All (65) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	880	ILE	13.3
1	A	884	VAL	13.0
1	A	950	PRO	11.9
1	A	883	LYS	11.2
1	A	949	VAL	11.0
1	A	882	SER	8.6
1	A	916	GLY	8.6
2	B	33	ILE	8.1
1	A	887	MET	7.8
1	A	879	GLY	7.7
1	A	924	HIS	7.5
1	A	921	GLN	6.4
1	A	918	ASN	6.3
1	A	960	VAL	6.2
1	A	886	PRO	5.8
1	A	927	ALA	5.8
1	A	917	GLU	5.7
1	A	926	ARG	5.6
1	A	947	VAL	5.5
1	A	919	LEU	5.3
1	A	860	LEU	5.1
1	A	922	VAL	5.0
1	A	910	GLU	5.0
1	A	861	LYS	4.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	881	GLU	4.8
1	A	925	GLN	4.8
2	B	32	LYS	4.7
1	A	912	VAL	4.7
1	A	864	THR	4.6
1	A	859	GLU	4.6
1	A	876	ILE	4.6
2	B	31	LYS	4.6
1	A	931	ILE	4.5
1	A	904	ALA	4.4
1	A	923	THR	4.4
1	A	888	VAL	4.3
1	A	907	ALA	4.3
1	A	908	GLY	4.3
1	A	928	VAL	4.3
1	A	920	GLU	4.2
1	A	929	ASP	4.2
1	A	933	ARG	4.1
1	A	944	GLU	3.9
2	B	34	THR	3.7
1	A	885	GLN	3.6
1	A	961	LYS	3.6
1	A	914	VAL	3.6
1	A	862	THR	3.5
1	A	948	ARG	3.5
1	A	905	LEU	3.3
1	A	878	GLY	3.2
1	A	911	LEU	3.2
1	A	940	ARG	3.1
1	A	865	LEU	3.1
1	A	877	SER	3.1
1	A	913	ALA	2.9
1	A	930	THR	2.9
1	A	1139	LYS	2.9
1	A	945	LEU	2.9
1	A	915	ASP	2.6
1	A	909	PHE	2.5
1	A	889	LYS	2.3
2	B	36	VAL	2.2
1	A	935	TYR	2.0
1	A	863	VAL	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 monosaccharides 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(Å^2)	Q<0.9
3	CA	A	1302	1/1	0.91	0.25	64,64,64,64	0
4	GOL	A	1307	6/6	0.93	0.12	25,38,43,49	0
3	CA	A	1301	1/1	0.96	0.14	58,58,58,58	0
3	CA	A	1303	1/1	0.97	0.17	50,50,50,50	0
4	GOL	A	1306	6/6	0.98	0.09	20,24,29,33	0
3	CA	A	1304	1/1	0.98	0.06	32,32,32,32	0
3	CA	A	1305	1/1	0.99	0.11	20,20,20,20	0

6.5 Other polymers (i)

There are no such residues in this entry.