Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2018 – 04:43 pm GMT

PDB ID	:	2C0L
Title	:	TPR DOMAIN OF HUMAN PEX5P IN COMPLEX WITH HUMAN MSCP2
Authors	:	Stanley, W.A.; Kursula, P.; Wilmanns, M.
Deposited on	:	2005-09-05
Resolution	:	2.30 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 ⓘ symbol.

The following versions of software and data (see references ⓘ) were used in the production of this report:

|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	trunk30967
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk30967

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 2.30 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		0.288
Clashscore		6
Ramachandran outliers		0.7%
Sidechain outliers		2.4%
RSRZ outliers		4.8%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
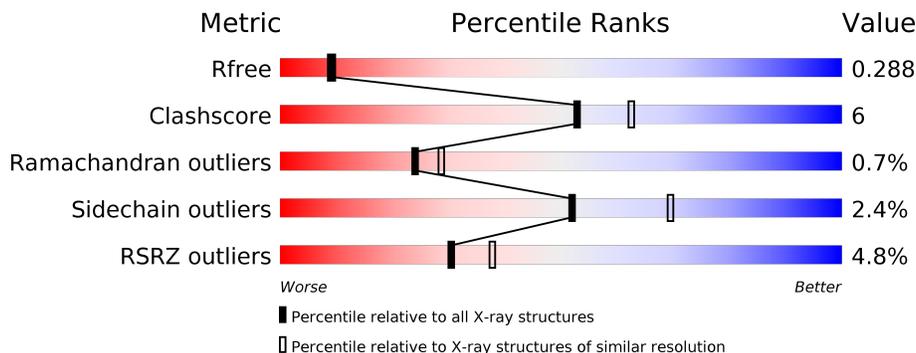
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	111664	4477 (2.30-2.30)
Clashscore	122126	5072 (2.30-2.30)
Ramachandran outliers	120053	5022 (2.30-2.30)
Sidechain outliers	120020	5021 (2.30-2.30)
RSRZ outliers	108989	4374 (2.30-2.30)

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 <=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	305	3% 80% 15% • •
2	B	122	10% 77% 20% •

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 3308 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 PEROXISOMAL TARGETING SIGNAL 1 RECEPTOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	292	Total 2287	C 1438	N 405	O 433	S 11	0	1	0

- Molecule 2 is a protein called NONSPECIFIC LIPID-TRANSFER PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	122	Total 922	C 587	N 154	O 175	S 6	0	0	0

- Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	81	Total 81	O 81	0	0
3	B	18	Total 18	O 18	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: PEROXISOMAL TARGETING SIGNAL 1 RECEPTOR
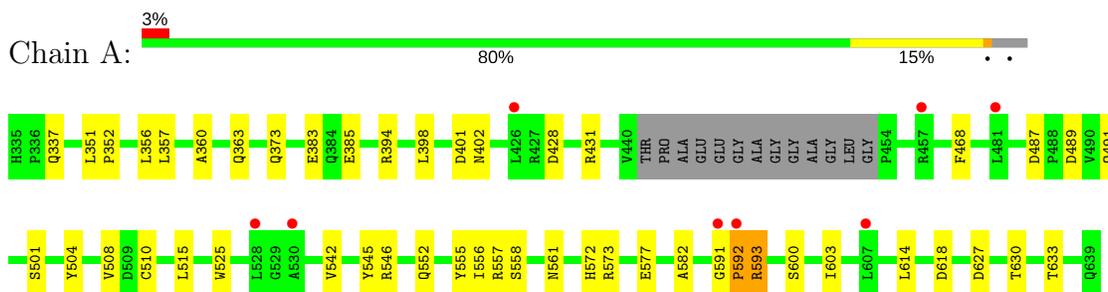
• Molecule 2: NONSPECIFIC LIPID-TRANSFER PROTEIN
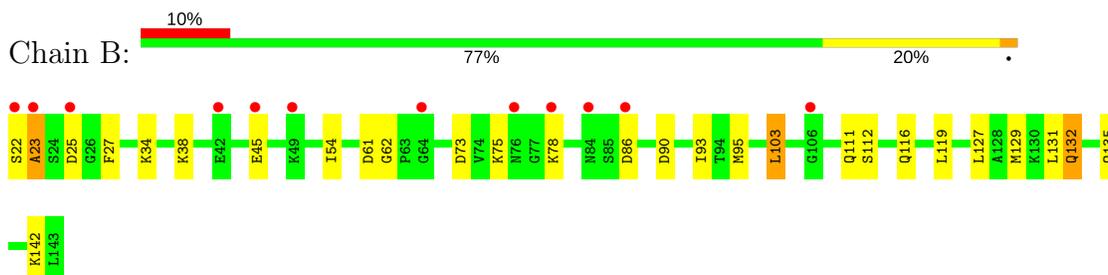
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	40.48Å 68.62Å 137.39Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.30 19.93 – 2.30	Depositor EDS
% Data completeness (in resolution range)	100.0 (20.00-2.30) 100.0 (19.93-2.30)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.43 (at 2.30Å)	Xtriage
Refinement program	REFMAC 5.1.9999	Depositor
R, R_{free}	0.202 , 0.256 0.222 , 0.288	Depositor DCC
R_{free} test set	885 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	33.8	Xtriage
Anisotropy	0.372	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 41.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.46, < L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	3308	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.00% 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)

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	0.33	0/2333	0.66	5/3159 (0.2%)
2	B	0.34	0/935	0.67	4/1246 (0.3%)
All	All	0.33	0/3268	0.66	9/4405 (0.2%)

There are no bond length outliers.

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	618	ASP	CB-CG-OD2	6.19	123.87	118.30
1	A	401	ASP	CB-CG-OD2	5.88	123.59	118.30
1	A	487	ASP	CB-CG-OD2	5.60	123.34	118.30
2	B	86	ASP	CB-CG-OD2	5.45	123.20	118.30
1	A	627	ASP	CB-CG-OD2	5.38	123.15	118.30
1	A	489	ASP	CB-CG-OD2	5.38	123.14	118.30
2	B	61	ASP	CB-CG-OD2	5.30	123.07	118.30
2	B	25	ASP	CB-CG-OD2	5.21	122.99	118.30
2	B	90	ASP	CB-CG-OD2	5.16	122.94	118.30

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	2287	0	2267	26	0
2	B	922	0	949	13	0

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	81	0	0	0	0
3	B	18	0	0	0	0
All	All	3308	0	3216	38	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 6.

All (38) 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:357:LEU:HG	1:A:592:PRO:HG3	1.45	0.97
1:A:356:LEU:HB2	1:A:592:PRO:HG2	1.55	0.88
1:A:373:GLN:HE21	1:A:402:ASN:HD21	1.36	0.71
2:B:132:GLN:HA	2:B:135:GLN:HG3	1.71	0.71
2:B:54:ILE:HD12	2:B:73:ASP:HA	1.74	0.70
1:A:572:HIS:HB2	1:A:614:LEU:HD13	1.79	0.65
1:A:428:ASP:HA	1:A:431:ARG:HG2	1.86	0.56
2:B:54:ILE:HD11	2:B:75:LYS:HB3	1.87	0.56
1:A:351:LEU:HB2	1:A:352:PRO:HD3	1.87	0.55
2:B:93:ILE:HD13	2:B:131:LEU:HD22	1.88	0.55
1:A:373:GLN:NE2	1:A:402:ASN:HD21	2.04	0.55
1:A:357:LEU:HG	1:A:592:PRO:CG	2.29	0.54
2:B:93:ILE:HG23	2:B:119:LEU:CD1	2.38	0.54
1:A:542:VAL:O	1:A:546:ARG:HG2	2.07	0.54
1:A:356:LEU:CB	1:A:592:PRO:HG2	2.33	0.53
1:A:468:PHE:CZ	1:A:501:SER:HB3	2.44	0.52
1:A:360:ALA:HA	1:A:363:GLN:HE21	1.73	0.52
1:A:557:ARG:HD2	2:B:142:LYS:HB2	1.92	0.51
1:A:360:ALA:HB1	1:A:593:ARG:HD3	1.91	0.51
2:B:112:SER:O	2:B:116:GLN:HG3	2.11	0.51
2:B:95:MET:SD	2:B:103:LEU:HD12	2.52	0.49
1:A:545:TYR:CZ	1:A:561:ASN:HB3	2.47	0.49
1:A:383:GLU:OE1	1:A:557:ARG:HD3	2.13	0.49
2:B:45:GLU:HB2	2:B:78:LYS:CE	2.45	0.47
1:A:394[A]:ARG:HH22	1:A:398:LEU:HD21	1.80	0.46
1:A:573:ARG:O	1:A:577:GLU:HG3	2.14	0.46
2:B:27:PHE:CE2	2:B:62:GLY:HA3	2.51	0.46
1:A:630:THR:O	1:A:633:THR:HB	2.16	0.45
2:B:34:LYS:O	2:B:38:LYS:HG3	2.17	0.45
1:A:556:ILE:HG23	1:A:582:ALA:HB1	1.98	0.45
1:A:504:TYR:O	1:A:508:VAL:HG23	2.17	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:515:LEU:HD22	1:A:525:TRP:CE2	2.52	0.44
1:A:572:HIS:CB	1:A:614:LEU:HD13	2.48	0.43
2:B:22:SER:O	2:B:23:ALA:C	2.57	0.43
2:B:111:GLN:HG3	2:B:129:MET:SD	2.60	0.41
1:A:600:SER:HB2	1:A:603:ILE:HD12	2.02	0.41
1:A:385:GLU:CD	1:A:557:ARG:HH22	2.24	0.40
1:A:491:GLN:O	1:A:510:CYS:HB3	2.21	0.40

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	289/305 (95%)	274 (95%)	13 (4%)	2 (1%)	24	29
2	B	120/122 (98%)	109 (91%)	10 (8%)	1 (1%)	21	25
All	All	409/427 (96%)	383 (94%)	23 (6%)	3 (1%)	24	29

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	591	GLY
1	A	592	PRO
2	B	23	ALA

5.3.2 Protein sidechains ⓘ

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	241/245 (98%)	236 (98%)	5 (2%)	56	73
2	B	98/98 (100%)	95 (97%)	3 (3%)	43	59
All	All	339/343 (99%)	331 (98%)	8 (2%)	52	69

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	337	GLN
1	A	552	GLN
1	A	555	TYR
1	A	558	SER
1	A	593	ARG
2	B	103	LEU
2	B	127	LEU
2	B	132	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	347	GLN
1	A	363	GLN
1	A	373	GLN
2	B	109	ASN
2	B	133	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

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 ⓘ

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	A	292/305 (95%)	0.33	8 (2%)	54	62	24, 33, 47, 61	0
2	B	122/122 (100%)	0.40	12 (9%)	7	10	19, 37, 52, 61	0
All	All	414/427 (96%)	0.35	20 (4%)	30	38	19, 34, 51, 61	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	591	GLY	7.8
1	A	592	PRO	5.1
2	B	25	ASP	3.3
2	B	86	ASP	3.1
2	B	23	ALA	3.1
1	A	457	ARG	3.0
2	B	49	LYS	2.9
2	B	84	ASN	2.9
1	A	426	LEU	2.9
1	A	528	LEU	2.8
2	B	22	SER	2.7
1	A	481	LEU	2.5
2	B	76	ASN	2.4
1	A	530	ALA	2.3
2	B	42	GLU	2.3
2	B	45	GLU	2.2
2	B	78	LYS	2.1
2	B	106	GLY	2.2
2	B	64	GLY	2.1
1	A	607	LEU	2.0

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.