Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2020 – 11:45 am BST

PDB ID	:	4QUM
Title	:	Crystal structure of PTPN3 (PTPH1) in complex with a dually phosphorylated MAPK12 peptide
Authors	:	Chen, K.E.; Meng, T.C.; Wang, A.H.J.
Deposited on	:	2014-07-10
Resolution	:	2.52 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:

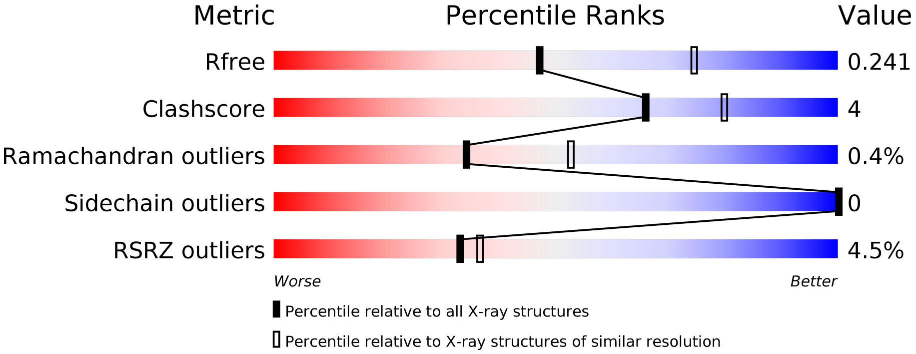
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA,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

1 Overall quality at a glance

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

The reported resolution of this entry is 2.52 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	5743 (2.54-2.50)
Clashscore	141614	6463 (2.54-2.50)
Ramachandran outliers	138981	6335 (2.54-2.50)
Sidechain outliers	138945	6337 (2.54-2.50)
RSRZ outliers	127900	5630 (2.54-2.50)

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	306	
2	B	9	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2339 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 Tyrosine-protein phosphatase non-receptor type 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	280	2241	1419	382	424	16	44	3	0

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	604	MET	-	EXPRESSION TAG	UNP P26045
A	605	HIS	-	EXPRESSION TAG	UNP P26045
A	606	HIS	-	EXPRESSION TAG	UNP P26045
A	607	HIS	-	EXPRESSION TAG	UNP P26045
A	608	HIS	-	EXPRESSION TAG	UNP P26045
A	609	HIS	-	EXPRESSION TAG	UNP P26045
A	610	HIS	-	EXPRESSION TAG	UNP P26045
A	611	SER	-	EXPRESSION TAG	UNP P26045
A	612	SER	-	EXPRESSION TAG	UNP P26045
A	613	GLY	-	EXPRESSION TAG	UNP P26045
A	614	VAL	-	EXPRESSION TAG	UNP P26045
A	615	ASP	-	EXPRESSION TAG	UNP P26045
A	616	LEU	-	EXPRESSION TAG	UNP P26045
A	617	GLY	-	EXPRESSION TAG	UNP P26045
A	618	THR	-	EXPRESSION TAG	UNP P26045
A	619	GLU	-	EXPRESSION TAG	UNP P26045
A	620	ASN	-	EXPRESSION TAG	UNP P26045
A	621	LEU	-	EXPRESSION TAG	UNP P26045
A	622	TYR	-	EXPRESSION TAG	UNP P26045
A	623	PHE	-	EXPRESSION TAG	UNP P26045
A	624	GLN	-	EXPRESSION TAG	UNP P26045
A	625	SER	-	EXPRESSION TAG	UNP P26045
A	626	ASN	-	EXPRESSION TAG	UNP P26045
A	627	ALA	-	EXPRESSION TAG	UNP P26045
A	811	ALA	ASP	ENGINEERED MUTATION	UNP P26045
A	842	SER	CYS	ENGINEERED MUTATION	UNP P26045

- Molecule 2 is a protein called Mitogen-activated protein kinase 12.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
2	B	8	71	40	11	17	2	1	6	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	190	ARG	TRP	ENGINEERED MUTATION	UNP P53778

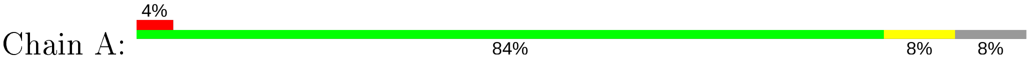
- Molecule 3 is water.

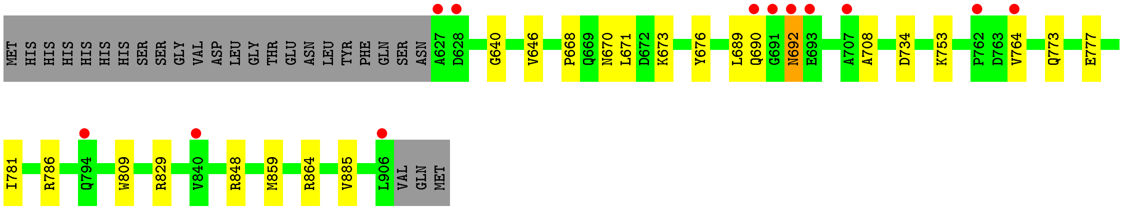
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	26	26	26	0	0
			Total	O		
3	B	1	1	1	0	0

3 Residue-property plots

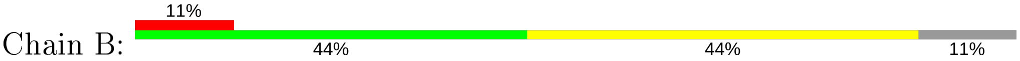
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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.

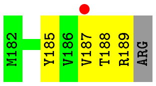
- Molecule 1: Tyrosine-protein phosphatase non-receptor type 3





- Molecule 2: Mitogen-activated protein kinase 12





4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	75.49Å 75.49Å 109.61Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	31.32 – 2.52 31.33 – 2.52	Depositor EDS
% Data completeness (in resolution range)	98.1 (31.32-2.52) 98.1 (31.33-2.52)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.86 (at 2.51Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, R_{free}	0.197 , 0.245 0.206 , 0.241	Depositor DCC
R_{free} test set	611 reflections (4.90%)	wwPDB-VP
Wilson B-factor (Å^2)	41.1	Xtriage
Anisotropy	0.565	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 39.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.030 for -h,-k,l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	2339	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.79% 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: TPO, PTR

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	0.45	0/2299	0.63	0/3126
2	B	0.40	0/41	0.75	0/50
All	All	0.45	0/2340	0.63	0/3176

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	A	2241	0	2224	16	0
2	B	71	0	62	3	0
3	A	26	0	0	0	0
3	B	1	0	0	0	0
All	All	2339	0	2286	18	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 (18) 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:670:ASN:HA	1:A:673:LYS:HE2	1.77	0.65
1:A:777:GLU:HG2	1:A:786:ARG:HG2	1.85	0.59
2:B:188:THR:OG1	2:B:189:ARG:N	2.38	0.56
1:A:829:ARG:HD3	1:A:859:MET:HE2	1.89	0.55
1:A:764:VAL:HG23	1:A:773:GLN:HB3	1.88	0.54
2:B:187:VAL:HG22	2:B:188:THR:H	1.74	0.52
1:A:640:GLY:HA3	1:A:646:VAL:HG13	1.92	0.51
1:A:668:PRO:HA	1:A:671:LEU:HD12	1.90	0.51
1:A:829:ARG:CD	1:A:859:MET:HE2	2.42	0.50
1:A:885:VAL:O	1:A:885:VAL:HG12	2.14	0.48
1:A:689:LEU:O	1:A:692:ASN:HA	2.15	0.46
1:A:753:LYS:HD2	1:A:753:LYS:HA	1.69	0.46
1:A:708:ALA:HB2	1:A:864:ARG:HD3	1.98	0.45
1:A:809:TRP:CE2	1:A:848:ARG:HG2	2.52	0.44
1:A:690:GLN:HB2	1:A:734:ASP:OD2	2.18	0.44
1:A:781:ILE:HA	1:A:781:ILE:HD12	1.93	0.41
1:A:689:LEU:HD23	1:A:689:LEU:HA	1.86	0.41
1:A:676:TYR:CG	2:B:185:PTR:HB3	2.56	0.41

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	281/306 (92%)	264 (94%)	16 (6%)	1 (0%)	34	53
2	B	4/9 (44%)	2 (50%)	2 (50%)	0	100	100
All	All	285/315 (90%)	266 (93%)	18 (6%)	1 (0%)	34	53

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	692	ASN

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	252/273 (92%)	252 (100%)	0	100	100
2	B	5/6 (83%)	5 (100%)	0	100	100
All	All	257/279 (92%)	257 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues 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	PTR	B	185	2	15,16,17	1.19	2 (13%)	19,22,24	0.58	0
2	TPO	B	183	2	8,10,11	1.30	1 (12%)	10,14,16	1.81	1 (10%)

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
2	PTR	B	185	2	-	0/10/11/13	0/1/1/1
2	TPO	B	183	2	-	5/9/11/13	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	185	PTR	OH-CZ	-3.67	1.32	1.40
2	B	185	PTR	P-OH	2.22	1.62	1.59
2	B	183	TPO	P-OG1	2.02	1.63	1.59

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	183	TPO	P-OG1-CB	-5.11	107.78	123.21

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	183	TPO	N-CA-CB-OG1
2	B	183	TPO	C-CA-CB-CG2
2	B	183	TPO	CB-OG1-P-O2P
2	B	183	TPO	N-CA-CB-CG2
2	B	183	TPO	O-C-CA-CB

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	185	PTR	1	0

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 ⓘ

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	280/306 (91%)	0.01	12 (4%)	35	39	27, 40, 65, 91	13 (4%)
2	B	6/9 (66%)	0.92	1 (16%)	1	1	60, 65, 82, 83	1 (16%)
All	All	286/315 (90%)	0.03	13 (4%)	33	36	27, 40, 65, 91	14 (4%)

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	693	GLU	8.6
1	A	627	ALA	7.4
1	A	906	LEU	4.4
1	A	690	GLN	4.3
1	A	692	ASN	4.3
1	A	762	PRO	3.0
1	A	691	GLY	2.7
1	A	764	VAL	2.7
1	A	628	ASP	2.4
2	B	187	VAL	2.4
1	A	794	GLN	2.4
1	A	707	ALA	2.2
1	A	840	VAL	2.1

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

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
2	TPO	B	183	11/12	0.85	0.15	65,74,98,117	0

Continued on next page...

Continued from previous page...

Mol	**Type**	**Chain**	**Res**	**Atoms**	**RSCC**	**RSR**	**B-factors($Å^2$)**	**Q<0.9**
2	PTR	B	185	16/17	0.97	0.18	34,41,57,57	0

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.