



wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 10:24 pm BST

PDB ID : 3S3K
Title : Crystal structure of the catalytic domain of PTP10D from *Drosophila melanogaster* with a small molecular inhibitor para-NitroCatechol Sulphate
Authors : Madan, L.L.; Gopal, B.
Deposited on : 2011-05-18
Resolution : 3.20 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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

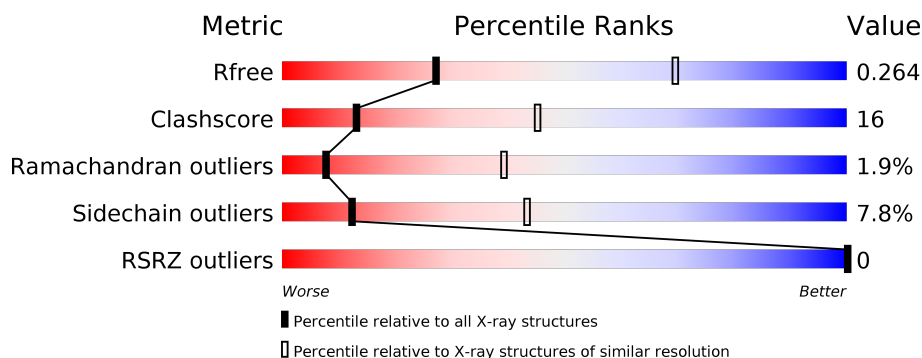
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.20 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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 $\leq 5\%$. 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	A	307	 64% 26% • 7%
1	B	307	 56% 29% 7% • 7%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4704 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 10D.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	287	Total	C	N	O	S	0	0	0
			2345	1490	420	418	17			
1	B	285	Total	C	N	O	S	0	0	0
			2318	1469	414	418	17			

There are 46 discrepancies between the modelled and reference sequences:

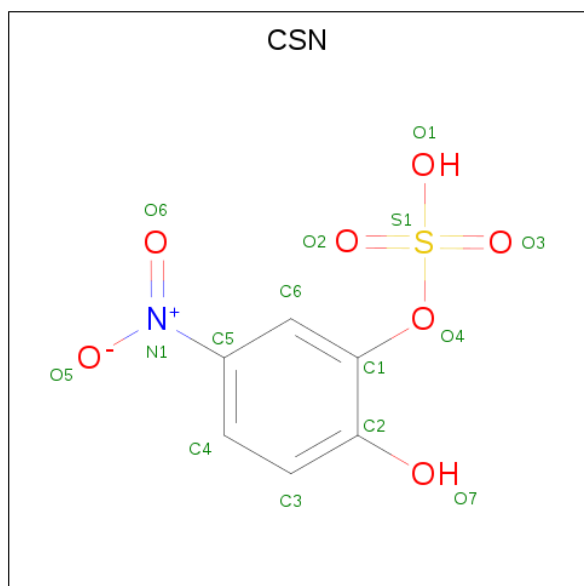
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	EXPRESSION TAG	UNP P35992
A	2	GLY	-	EXPRESSION TAG	UNP P35992
A	3	SER	-	EXPRESSION TAG	UNP P35992
A	4	SER	-	EXPRESSION TAG	UNP P35992
A	5	HIS	-	EXPRESSION TAG	UNP P35992
A	6	HIS	-	EXPRESSION TAG	UNP P35992
A	7	HIS	-	EXPRESSION TAG	UNP P35992
A	8	HIS	-	EXPRESSION TAG	UNP P35992
A	9	HIS	-	EXPRESSION TAG	UNP P35992
A	10	HIS	-	EXPRESSION TAG	UNP P35992
A	11	SER	-	EXPRESSION TAG	UNP P35992
A	12	SER	-	EXPRESSION TAG	UNP P35992
A	13	GLY	-	EXPRESSION TAG	UNP P35992
A	14	LEU	-	EXPRESSION TAG	UNP P35992
A	15	VAL	-	EXPRESSION TAG	UNP P35992
A	16	PRO	-	EXPRESSION TAG	UNP P35992
A	17	ARG	-	EXPRESSION TAG	UNP P35992
A	18	GLY	-	EXPRESSION TAG	UNP P35992
A	19	SER	-	EXPRESSION TAG	UNP P35992
A	20	HIS	-	EXPRESSION TAG	UNP P35992
A	21	MET	-	EXPRESSION TAG	UNP P35992
A	22	ALA	-	EXPRESSION TAG	UNP P35992
A	23	SER	-	EXPRESSION TAG	UNP P35992
B	1	MET	-	EXPRESSION TAG	UNP P35992
B	2	GLY	-	EXPRESSION TAG	UNP P35992

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	3	SER	-	EXPRESSION TAG	UNP P35992
B	4	SER	-	EXPRESSION TAG	UNP P35992
B	5	HIS	-	EXPRESSION TAG	UNP P35992
B	6	HIS	-	EXPRESSION TAG	UNP P35992
B	7	HIS	-	EXPRESSION TAG	UNP P35992
B	8	HIS	-	EXPRESSION TAG	UNP P35992
B	9	HIS	-	EXPRESSION TAG	UNP P35992
B	10	HIS	-	EXPRESSION TAG	UNP P35992
B	11	SER	-	EXPRESSION TAG	UNP P35992
B	12	SER	-	EXPRESSION TAG	UNP P35992
B	13	GLY	-	EXPRESSION TAG	UNP P35992
B	14	LEU	-	EXPRESSION TAG	UNP P35992
B	15	VAL	-	EXPRESSION TAG	UNP P35992
B	16	PRO	-	EXPRESSION TAG	UNP P35992
B	17	ARG	-	EXPRESSION TAG	UNP P35992
B	18	GLY	-	EXPRESSION TAG	UNP P35992
B	19	SER	-	EXPRESSION TAG	UNP P35992
B	20	HIS	-	EXPRESSION TAG	UNP P35992
B	21	MET	-	EXPRESSION TAG	UNP P35992
B	22	ALA	-	EXPRESSION TAG	UNP P35992
B	23	SER	-	EXPRESSION TAG	UNP P35992

- Molecule 2 is N,4-DIHYDROXY-N-OXO-3-(SULFOOXY)BENZENAMINIUM (three-letter code: CSN) (formula: C₆H₅NO₇S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	B	1	Total	C	N	O	S	0	0
			15	6	1	7	1		

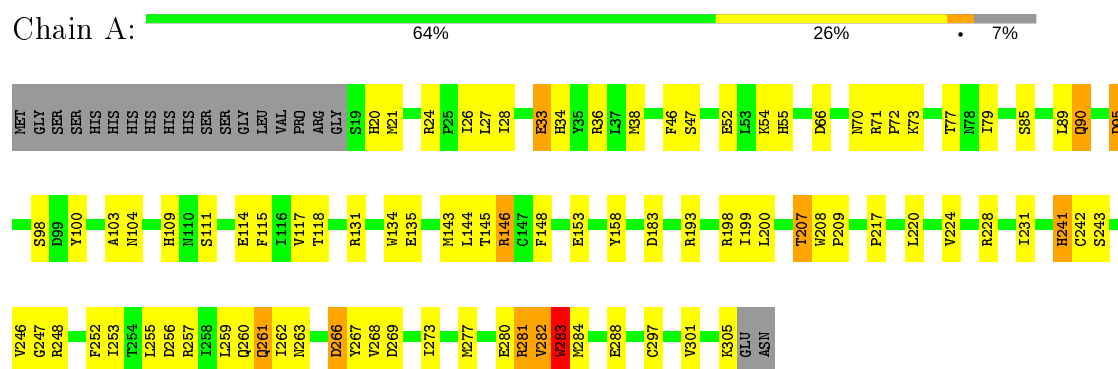
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	7	Total	O	0	0
			7	7		
3	B	19	Total	O	0	0
			19	19		

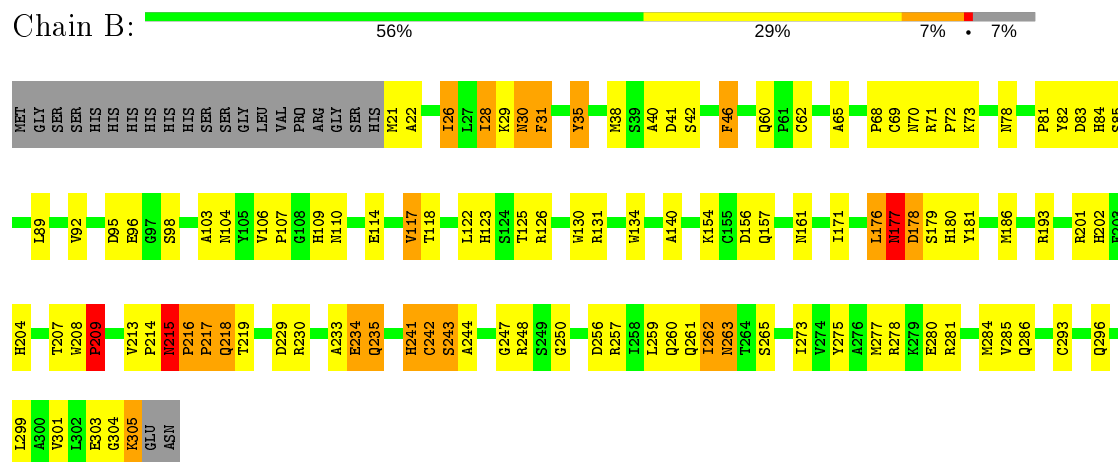
3 Residue-property plots

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 10D



• Molecule 1: Tyrosine-protein phosphatase 10D



4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	102.51Å 102.51Å 172.80Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.08 – 3.20 44.08 – 3.20	Depositor EDS
% Data completeness (in resolution range)	96.2 (44.08-3.20) 96.2 (44.08-3.20)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.37 (at 3.19Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.253 , 0.309 0.222 , 0.264	Depositor DCC
R_{free} test set	880 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å ²)	39.1	Xtriage
Anisotropy	0.027	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 56.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.047 for -h,-k,l	Xtriage
F_o, F_c correlation	0.86	EDS
Total number of atoms	4704	wwPDB-VP
Average B, all atoms (Å ²)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.56% 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 ⓘ

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

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	0.70	5/2412 (0.2%)	0.65	1/3272 (0.0%)
1	B	0.74	4/2383 (0.2%)	0.66	1/3236 (0.0%)
All	All	0.72	9/4795 (0.2%)	0.65	2/6508 (0.0%)

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	242	CYS	CB-SG	-11.89	1.62	1.82
1	A	158	TYR	CE1-CZ	-8.18	1.27	1.38
1	A	242	CYS	CB-SG	-6.27	1.71	1.82
1	A	158	TYR	CE2-CZ	-5.92	1.30	1.38
1	A	283	TRP	CB-CG	-5.44	1.40	1.50

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	198	ARG	NE-CZ-NH2	-6.66	116.97	120.30
1	B	176	LEU	CA-CB-CG	6.18	129.52	115.30

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	A	2345	0	2225	62	0
1	B	2318	0	2195	87	0
2	B	15	0	5	2	0
3	A	7	0	0	0	0
3	B	19	0	0	0	0
All	All	4704	0	4425	148	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 16.

The worst 5 of 148 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:B:29:LYS:C	1:B:30:ASN:HD22	1.46	1.18
1:B:304:GLY:O	1:B:305:LYS:HB2	1.41	1.12
1:B:176:LEU:O	1:B:177:ASN:HB2	1.51	1.04
1:A:95:ASP:HB3	1:A:98:SER:HB3	1.39	1.04
1:B:28:ILE:HD11	1:B:262:ILE:HG22	1.45	0.96

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	285/307 (93%)	258 (90%)	25 (9%)	2 (1%)	22	61
1	B	283/307 (92%)	248 (88%)	26 (9%)	9 (3%)	4	26
All	All	568/614 (92%)	506 (89%)	51 (9%)	11 (2%)	8	39

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	40	ALA
1	B	177	ASN
1	A	283	TRP
1	B	216	PRO
1	B	218	GLN

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	251/278 (90%)	235 (94%)	16 (6%)	17	52
1	B	250/278 (90%)	227 (91%)	23 (9%)	9	33
All	All	501/556 (90%)	462 (92%)	39 (8%)	12	43

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	28	ILE
1	B	85	SER
1	B	263	ASN
1	B	30	ASN
1	B	46	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	84	HIS
1	B	104	ASN
1	B	215	ASN
1	B	60	GLN
1	B	180	HIS

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](#)

1 ligand is 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	CSN	B	308	-	14,15,15	5.01	6 (42%)	19,22,22	2.00	4 (21%)

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	CSN	B	308	-	-	3/7/9/9	0/1/1/1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	308	CSN	O6-N1	11.75	1.42	1.22
2	B	308	CSN	C6-C5	8.28	1.53	1.39
2	B	308	CSN	C6-C1	8.16	1.53	1.38
2	B	308	CSN	O4-S1	-6.86	1.47	1.58
2	B	308	CSN	C5-N1	-4.06	1.35	1.45

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	308	CSN	C5-C6-C1	-6.66	111.55	119.12
2	B	308	CSN	C6-C1-C2	3.86	124.21	120.06
2	B	308	CSN	O6-N1-C5	2.07	121.73	118.80
2	B	308	CSN	C3-C4-C5	2.01	122.88	120.08

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	308	CSN	C1-O4-S1-O2
2	B	308	CSN	C1-O4-S1-O1
2	B	308	CSN	C1-O4-S1-O3

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	308	CSN	2	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	287/307 (93%)	-0.45	0 100 100	27, 53, 81, 104	0
1	B	285/307 (92%)	-0.39	0 100 100	27, 52, 87, 120	0
All	All	572/614 (93%)	-0.42	0 100 100	27, 52, 83, 120	0

There are no RSRZ outliers to report.

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(Å ²)	Q<0.9
2	CSN	B	308	15/15	0.91	0.34	38,38,38,38	0

6.5 Other polymers [i](#)

There are no such residues in this entry.