Full wwPDB X-ray Structure Validation Report (i)

Feb 5, 2024 – 09:19 PM EST

PDB ID	:	1YTS
Title	:	A LIGAND-INDUCED CONFORMATIONAL CHANGE IN THE YERSINIA PROTEIN TYROSINE PHOSPHATASE
Authors	:	Schubert, H.L.; Stuckey, J.A.; Fauman, E.B.; Dixon, J.E.; Saper, M.A.
Deposited on	:	1995-04-07
Resolution	:	2.50 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

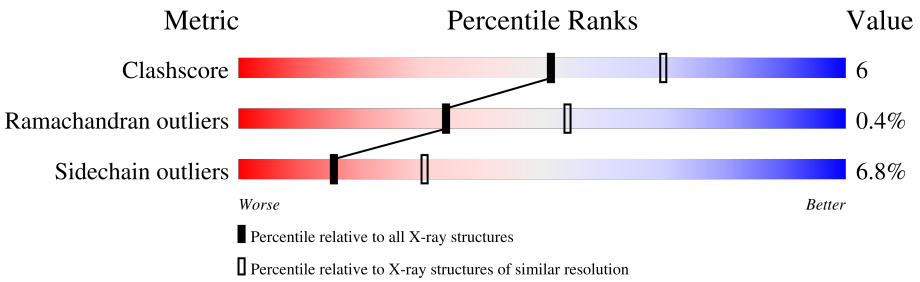
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.50 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))
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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 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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	278	70% 25% • •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2207 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.

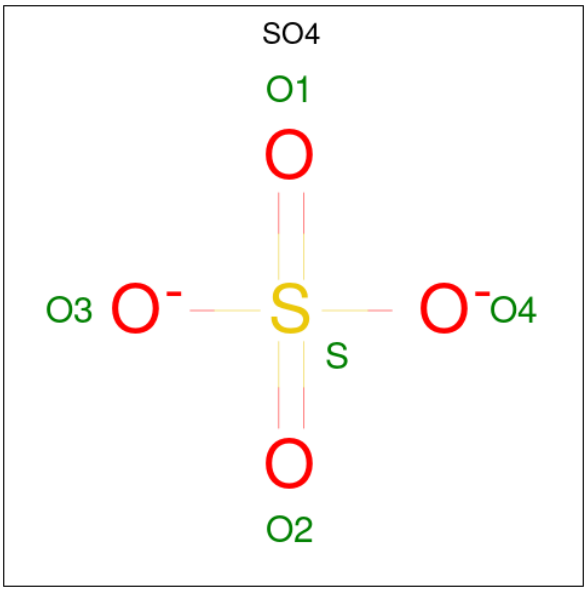
- Molecule 1 is a protein called YERSINIA PROTEIN TYROSINE PHOSPHATASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	278	2137	1304	398	420	15	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	235	ARG	CYS	conflict	UNP P15273
A	403	SER	CYS	conflict	UNP P15273

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
2	A	1	5	4	1	0	0
2	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	60	60	60	0	0

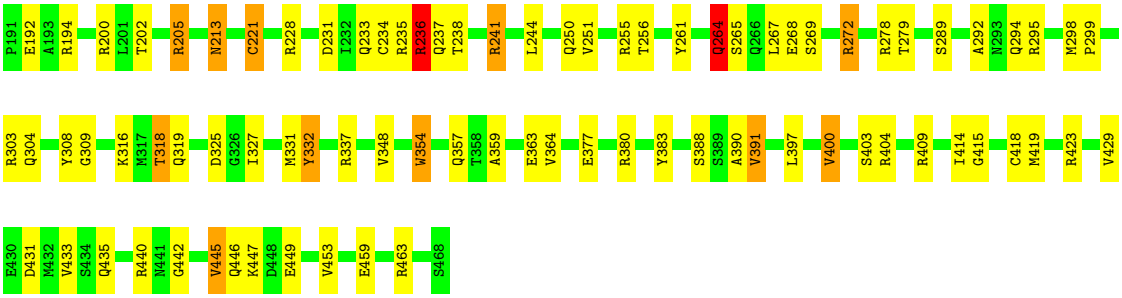
3 Residue-property plots (i)

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. 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. 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.

Note EDS was not executed.

• Molecule 1: YERSINIA PROTEIN TYROSINE PHOSPHATASE

Chain A:



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	56.40Å 49.80Å 100.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 2.50	Depositor
% Data completeness (in resolution range)	(Not available) (10.00-2.50)	Depositor
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, R_{free}	0.174 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2207	wwPDB-VP
Average B, all atoms (Å^2)	18.0	wwPDB-VP

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: SO4

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	1.05	1/2162 (0.0%)	1.93	50/2918 (1.7%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	264	GLN	CA-CB	-5.21	1.42	1.53

All (50) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	255	ARG	NE-CZ-NH2	12.80	126.70	120.30
1	A	331	MET	CG-SD-CE	-12.35	80.44	100.20
1	A	235	ARG	NE-CZ-NH1	11.59	126.09	120.30
1	A	331	MET	CA-CB-CG	-9.72	96.77	113.30
1	A	241	ARG	NE-CZ-NH2	-9.55	115.52	120.30
1	A	205	ARG	NE-CZ-NH1	9.49	125.04	120.30
1	A	332	TYR	CB-CG-CD2	-9.14	115.51	121.00
1	A	380	ARG	NE-CZ-NH1	8.85	124.73	120.30
1	A	200	ARG	NE-CZ-NH2	8.49	124.55	120.30
1	A	463	ARG	NE-CZ-NH2	-8.12	116.24	120.30
1	A	440	ARG	NE-CZ-NH2	7.63	124.12	120.30
1	A	380	ARG	NE-CZ-NH2	-7.42	116.59	120.30
1	A	354	TRP	CG-CD2-CE3	7.34	140.51	133.90
1	A	221	CYS	CA-C-N	-7.12	101.96	116.20
1	A	213	ASN	CA-CB-CG	-6.93	98.16	113.40
1	A	228	ARG	NE-CZ-NH1	-6.87	116.86	120.30
1	A	295	ARG	NE-CZ-NH2	-6.84	116.88	120.30
1	A	233	GLN	CA-CB-CG	-6.82	98.40	113.40
1	A	236	ARG	NE-CZ-NH1	6.76	123.68	120.30
1	A	431	ASP	CB-CG-OD2	-6.71	112.26	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	423	ARG	NE-CZ-NH2	6.66	123.63	120.30
1	A	278	ARG	NE-CZ-NH2	6.59	123.60	120.30
1	A	337	ARG	NE-CZ-NH2	-6.59	117.01	120.30
1	A	265	SER	CA-CB-OG	6.55	128.88	111.20
1	A	272	ARG	NE-CZ-NH2	-6.49	117.06	120.30
1	A	354	TRP	CE2-CD2-CG	-6.39	102.19	107.30
1	A	383	TYR	CB-CG-CD1	-6.24	117.26	121.00
1	A	433	VAL	CG1-CB-CG2	-6.21	100.96	110.90
1	A	391	VAL	CG1-CB-CG2	-6.20	100.99	110.90
1	A	256	THR	CA-CB-CG2	6.12	120.97	112.40
1	A	234	CYS	CA-CB-SG	5.87	124.56	114.00
1	A	231	ASP	CB-CG-OD1	5.71	123.44	118.30
1	A	202	THR	CA-C-N	5.50	129.31	117.20
1	A	414	ILE	CA-C-N	5.48	127.17	116.20
1	A	298	MET	CA-CB-CG	-5.47	104.00	113.30
1	A	453	VAL	CG1-CB-CG2	-5.42	102.23	110.90
1	A	267	LEU	CB-CG-CD1	-5.41	101.81	111.00
1	A	429	VAL	CG1-CB-CG2	-5.36	102.33	110.90
1	A	404	ARG	NE-CZ-NH2	-5.33	117.63	120.30
1	A	359	ALA	CA-C-N	5.32	128.90	117.20
1	A	445	VAL	CA-C-N	-5.31	105.53	117.20
1	A	194	ARG	NE-CZ-NH1	5.28	122.94	120.30
1	A	298	MET	CG-SD-CE	5.25	108.60	100.20
1	A	397	LEU	CD1-CG-CD2	-5.24	94.77	110.50
1	A	250	GLN	CA-CB-CG	-5.14	102.09	113.40
1	A	303	ARG	NE-CZ-NH1	5.13	122.87	120.30
1	A	318	THR	CA-C-N	5.11	128.44	117.20
1	A	400	VAL	CA-C-N	5.09	128.41	117.20
1	A	194	ARG	NE-CZ-NH2	-5.03	117.78	120.30
1	A	303	ARG	NE-CZ-NH2	-5.03	117.78	120.30

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	2137	0	2144	24	0
2	A	10	0	0	0	0
3	A	60	0	0	0	0
All	All	2207	0	2144	24	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 (24) 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:237:GLN:HG3	1:A:238:THR:HG23	1.62	0.80
1:A:264:GLN:HG2	1:A:299:PRO:HG3	1.78	0.64
1:A:251:VAL:HA	1:A:435:GLN:OE1	2.05	0.56
1:A:388:SER:HB3	1:A:391:VAL:HG23	1.90	0.54
1:A:415:GLY:O	1:A:419:MET:HG3	2.08	0.53
1:A:279:THR:HG23	1:A:400:VAL:HG23	1.91	0.52
1:A:327:ILE:HG13	1:A:364:VAL:HG11	1.91	0.52
1:A:205:ARG:NH2	1:A:446:GLN:HA	2.28	0.48
1:A:354:TRP:CE2	1:A:409:ARG:HG2	2.49	0.48
1:A:268:GLU:O	1:A:272:ARG:HG3	2.15	0.47
1:A:244:LEU:HD21	1:A:269:SER:HB2	1.96	0.47
1:A:292:ALA:O	1:A:294:GLN:NE2	2.48	0.47
1:A:316:LYS:HE2	1:A:318:THR:HG22	1.98	0.45
1:A:236:ARG:HE	1:A:236:ARG:HB3	1.56	0.44
1:A:268:GLU:OE1	1:A:309:GLY:HA3	2.18	0.44
1:A:261:TYR:HE1	1:A:299:PRO:HD2	1.83	0.43
1:A:268:GLU:HG3	1:A:308:TYR:O	2.19	0.43
1:A:205:ARG:HH22	1:A:446:GLN:HA	1.84	0.42
1:A:357:GLN:O	1:A:447:LYS:HE2	2.19	0.42
1:A:332:TYR:HB2	1:A:348:VAL:HB	2.01	0.42
1:A:388:SER:OG	1:A:390:ALA:HB3	2.20	0.41
1:A:205:ARG:HG3	1:A:442:GLY:O	2.21	0.41
1:A:261:TYR:CE1	1:A:299:PRO:HD2	2.57	0.40
1:A:316:LYS:O	1:A:332:TYR:HA	2.20	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	276/278 (99%)	259 (94%)	16 (6%)	1 (0%)	34 54

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	445	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	235/235 (100%)	219 (93%)	16 (7%)	16 30

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	192	GLU
1	A	213	ASN
1	A	221	CYS
1	A	236	ARG
1	A	241	ARG
1	A	264	GLN
1	A	289	SER
1	A	304	GLN
1	A	319	GLN
1	A	325	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	363	GLU
1	A	377	GLU
1	A	403	SER
1	A	418	CYS
1	A	449	GLU
1	A	459	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	264	GLN
1	A	294	GLN
1	A	320	GLN

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)

2 ligands 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	SO4	A	1	-	4,4,4	0.76	0	6,6,6	0.99	0
2	SO4	A	2	-	4,4,4	0.58	0	6,6,6	0.82	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.