wwPDB X-ray Structure Validation Summary Report (i)

Oct 11, 2021 – 09:02 AM EDT

PDB ID	:	3BRH
Title	:	Protein Tyrosine Phosphatase PTPN-22 (Lyp) bound to the mono-Phosphorylated Lck active site peptide
Authors	:	Seidel, R.D.; Love, J.; Piserchio, A.; Cowburn, D.
Deposited on	:	2007-12-21
Resolution	:	2.20 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 (i) symbol.

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

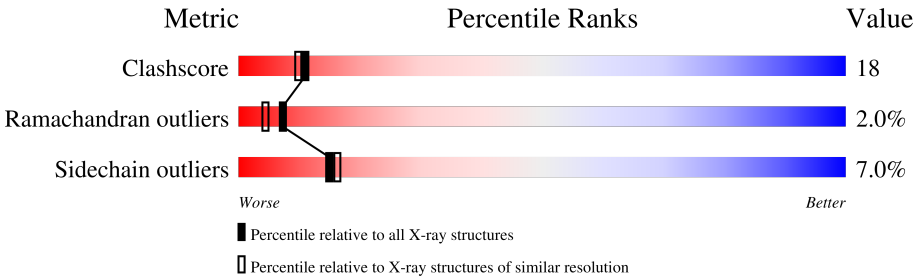
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 2.20 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))
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)

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 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	310	65% 27% • • 5%
1	B	310	67% 23% • • 5%
2	C	7	14% 29% 57%
2	D	7	29% 29% 43%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	PO4	A	400	-	X	-	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5335 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 22.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	296	Total	C	N	O	S	0	4	0
			2442	1575	397	454	16			
1	B	296	Total	C	N	O	S	0	5	0
			2475	1594	406	457	18			

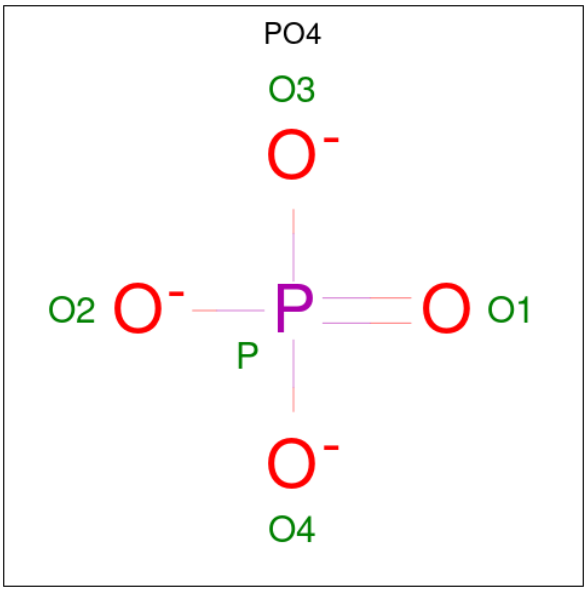
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	195	ALA	ASP	engineered mutation	UNP Q9Y2R2
A	227	SER	CYS	engineered mutation	UNP Q9Y2R2
B	195	ALA	ASP	engineered mutation	UNP Q9Y2R2
B	227	SER	CYS	engineered mutation	UNP Q9Y2R2

- Molecule 2 is a protein called Lck Active Site Peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	7	Total	C	N	O	0	1	0
			68	40	12	16			
2	D	4	Total	C	N	O	0	1	0
			45	27	6	12			

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O P 5 4 1	0	0
3	B	1	Total O P 5 4 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	139	Total O 139 139	0	0
4	B	149	Total O 150 150	0	1
4	C	5	Total O 5 5	0	0
4	D	1	Total O 1 1	0	0

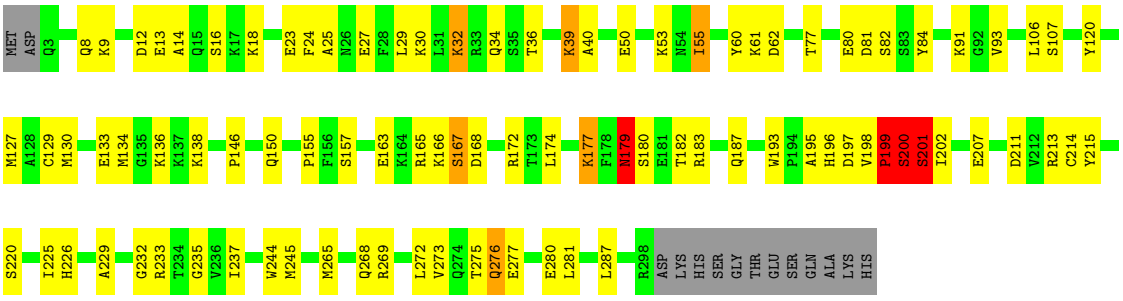
3 Residue-property plots (i)

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.

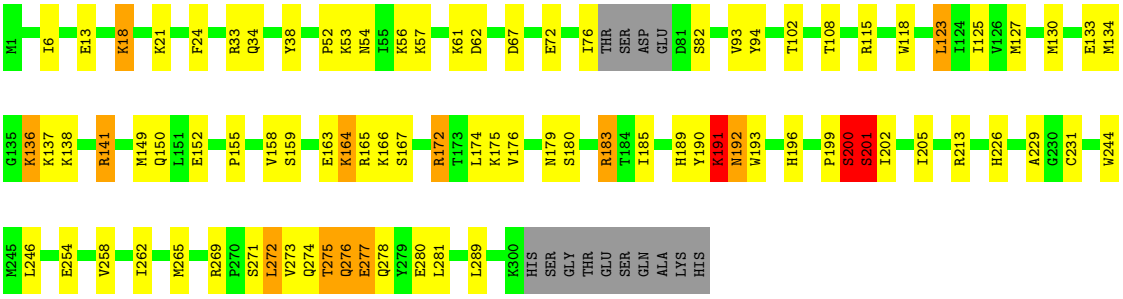
• Molecule 1: Tyrosine-protein phosphatase non-receptor type 22

Chain A:



• Molecule 1: Tyrosine-protein phosphatase non-receptor type 22

Chain B:



• Molecule 2: Lck Active Site Peptide

Chain C:

• Molecule 2: Lck Active Site Peptide

Chain D:


D391
N392
E393
Y394
THR
ALA
ARG

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	60.85Å 48.51Å 119.97Å 90.00° 103.49° 90.00°	Depositor
Resolution (Å)	29.59 – 2.20	Depositor
% Data completeness (in resolution range)	98.6 (29.59-2.20)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC	Depositor
R, R_{free}	0.188 , 0.264	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5335	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PO4

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	0.94	5/2507 (0.2%)	1.29	14/3387 (0.4%)
1	B	0.89	0/2542	1.34	20/3429 (0.6%)
2	C	1.06	0/68	2.13	2/90 (2.2%)
2	D	1.28	1/45 (2.2%)	1.12	0/59
All	All	0.92	6/5162 (0.1%)	1.33	36/6965 (0.5%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	3
1	B	0	3
2	C	3	5
2	D	0	1
All	All	3	12

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	84	TYR	CD1-CE1	-6.04	1.30	1.39
1	A	120	TYR	CD1-CE1	-5.68	1.30	1.39
1	A	84	TYR	CZ-OH	-5.45	1.28	1.37
1	A	84	TYR	CD2-CE2	-5.32	1.31	1.39
2	D	394	TYR	CB-CG	-5.12	1.44	1.51

The worst 5 of 36 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	394	TYR	N-CA-CB	15.35	138.23	110.60
2	C	394	TYR	CA-CB-CG	8.94	130.38	113.40
1	B	123	LEU	CA-CB-CG	8.78	135.49	115.30
1	B	201	SER	N-CA-C	8.40	133.69	111.00
1	B	165	ARG	NE-CZ-NH2	-8.14	116.23	120.30

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	C	394	TYR	CA
2	C	395	THR	CA
2	C	396	ALA	CA

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	199	PRO	Mainchain,Peptide
1	B	191	LYS	Peptide
1	B	199	PRO	Peptide
1	B	200	SER	Peptide
2	C	391	ASP	Peptide

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	2442	0	2422	90	0
1	B	2475	0	2469	71	0
2	C	68	0	53	47	0
2	D	45	0	28	15	0
3	A	5	0	0	1	0
3	B	5	0	0	1	0
4	A	139	0	0	12	0
4	B	150	0	0	10	0
4	C	5	0	0	7	0
4	D	1	0	0	0	0
All	All	5335	0	4972	181	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 18.

The worst 5 of 181 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:196:HIS:CE1	2:C:394:TYR:CE2	2.13	1.34
4:A:570:HOH:O	1:B:149:MET:HG3	1.32	1.27
1:B:196:HIS:HE1	2:D:394:TYR:CE1	1.53	1.25
1:A:196:HIS:NE2	2:C:394:TYR:HE2	1.37	1.22
1:A:280:GLU:HB3	4:A:329:HOH:O	1.47	1.13

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	295/310 (95%)	273 (92%)	17 (6%)	5 (2%)	9 6
1	B	298/310 (96%)	276 (93%)	19 (6%)	3 (1%)	15 14
2	C	6/7 (86%)	2 (33%)	1 (17%)	3 (50%)	0 0
2	D	3/7 (43%)	2 (67%)	0	1 (33%)	0 0
All	All	602/634 (95%)	553 (92%)	37 (6%)	12 (2%)	7 4

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	179	ASN
1	A	199	PRO
1	B	201	SER
2	C	396	ALA
2	C	392	ASN

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	269/284 (95%)	254 (94%)	15 (6%)	21 25
1	B	274/284 (96%)	256 (93%)	18 (7%)	16 19
2	C	7/6 (117%)	3 (43%)	4 (57%)	0 0
2	D	5/6 (83%)	1 (20%)	4 (80%)	0 0
All	All	555/580 (96%)	514 (93%)	41 (7%)	15 14

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	201	SER
2	C	394	TYR
1	B	272	LEU
1	B	277	GLU
2	D	392	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	276	GLN
1	B	192	ASN
2	D	392	ASN
1	B	196	HIS
1	A	268	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	$\#\|Z\| > 2$
3	PO4	B	400	-	4,4,4	1.27	0	6,6,6	1.48	1 (16%)
3	PO4	A	400	-	4,4,4	5.29	4 (100%)	6,6,6	1.20	0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	400	PO4	P-O3	-5.97	1.36	1.54
3	A	400	PO4	P-O2	-5.97	1.36	1.54
3	A	400	PO4	P-O4	-4.62	1.40	1.54
3	A	400	PO4	P-O1	-4.36	1.40	1.50

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	400	PO4	O3-P-O2	2.84	117.08	107.97

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	400	PO4	1	0
3	A	400	PO4	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

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.