Full wwPDB X-ray Structure Validation Report ⓘ

Aug 21, 2020 – 04:17 PM BST

PDB ID	:	6TPW
Title	:	Crystal structures of FNIII domain one through four of the human leucocyte common antigen-related protein (LAR)
Authors	:	Vilstrup, J.P.; Thirup, S.S.; Simonsen, A.; Birkefeldt, T.; Strandbygaard, D.
Deposited on	:	2019-12-14
Resolution	:	2.90 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 ①) were used in the production of this report:

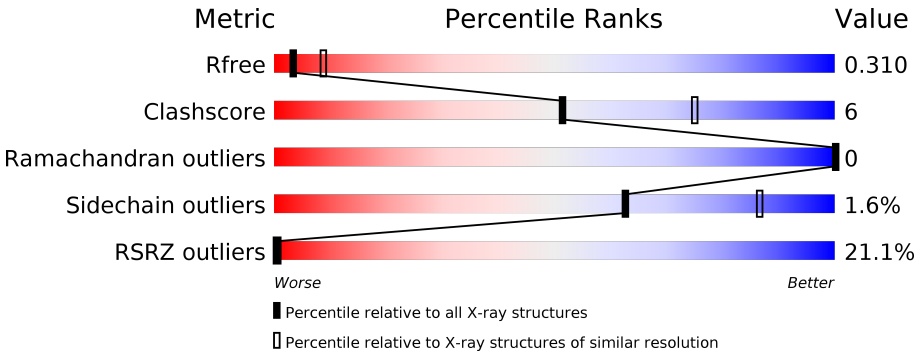
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

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 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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	402	20% 80% 12% • 7%

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
2	SO4	A	802	-	-	-	X

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2920 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.

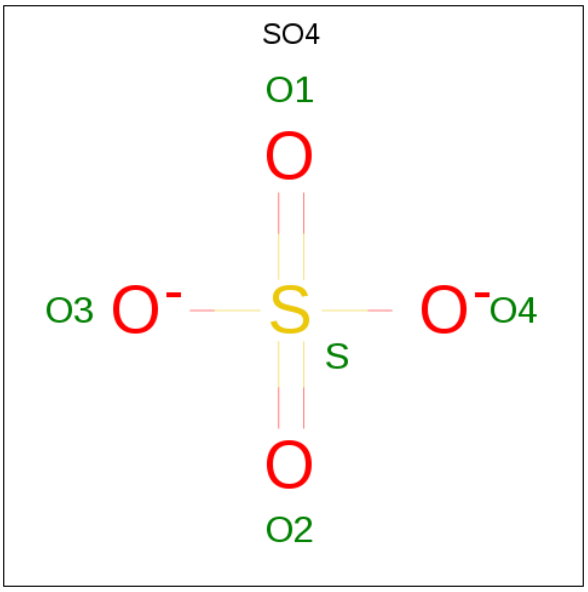
- Molecule 1 is a protein called Receptor-type tyrosine-protein phosphatase F.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	375	Total	C	N	O	S	0	0	0
			2909	1839	499	564	7			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	305	MET	-	initiating methionine	UNP P10586
A	306	HIS	-	expression tag	UNP P10586
A	307	HIS	-	expression tag	UNP P10586
A	308	HIS	-	expression tag	UNP P10586
A	309	HIS	-	expression tag	UNP P10586
A	310	HIS	-	expression tag	UNP P10586
A	311	HIS	-	expression tag	UNP P10586
A	312	GLU	-	expression tag	UNP P10586
A	313	ASN	-	expression tag	UNP P10586
A	314	LEU	-	expression tag	UNP P10586
A	315	TYR	-	expression tag	UNP P10586
A	316	PHE	-	expression tag	UNP P10586
A	317	GLN	-	expression tag	UNP P10586
A	318	GLY	-	expression tag	UNP P10586

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



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

- Molecule 3 is water.

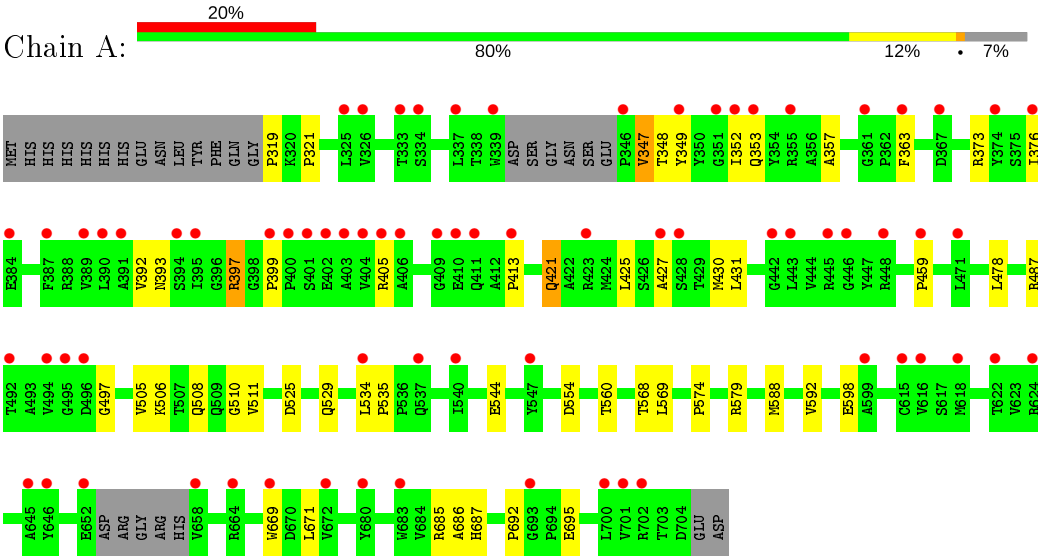
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O 1 1	0	0

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 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: Receptor-type tyrosine-protein phosphatase F

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 41 2 2	Depositor
Cell constants a, b, c, α, β, γ	74.86Å 74.86Å 461.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.98 – 2.90 49.45 – 2.90	Depositor EDS
% Data completeness (in resolution range)	98.7 (38.98-2.90) 99.9 (49.45-2.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.51 (at 2.91Å)	Xtriage
Refinement program	PHENIX 1.14_3260	Depositor
R, R_{free}	0.281 , 0.301 0.284 , 0.310	Depositor DCC
R_{free} test set	766 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	86.6	Xtriage
Anisotropy	0.766	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 104.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	2920	wwPDB-VP
Average B, all atoms (Å^2)	164.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.76% 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: 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	0.25	0/2990	0.45	0/4099

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	2909	0	2832	32	0
2	A	10	0	0	0	0
3	A	1	0	0	0	0
All	All	2920	0	2832	32	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 (32) 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:347:VAL:HG12	1:A:392:VAL:O	1.35	1.22

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:347:VAL:CG1	1:A:392:VAL:O	2.22	0.87
1:A:579:ARG:HG2	1:A:598:GLU:HG2	1.82	0.61
1:A:685:ARG:HD3	1:A:692:PRO:HB2	1.84	0.60
1:A:459:PRO:HB3	1:A:487:ARG:HH21	1.67	0.59
1:A:544:GLU:HG2	1:A:560:THR:HG22	1.86	0.57
1:A:510:GLY:HA3	1:A:588:MET:HG3	1.87	0.57
1:A:352:ILE:HG21	1:A:376:ILE:HD11	1.87	0.56
1:A:347:VAL:HG12	1:A:392:VAL:C	2.19	0.56
1:A:348:THR:HG22	1:A:393:ASN:HA	1.91	0.52
1:A:534:LEU:HD12	1:A:535:PRO:HD2	1.93	0.51
1:A:357:ALA:HB2	1:A:405:ARG:HH12	1.75	0.50
1:A:669:TRP:HZ3	1:A:671:LEU:HD23	1.77	0.49
1:A:425:LEU:HD11	1:A:431:LEU:HB2	1.94	0.49
1:A:321:PRO:HD3	1:A:399:PRO:HG2	1.94	0.48
1:A:529:GLN:HA	1:A:568:THR:HA	1.96	0.47
1:A:348:THR:HG23	1:A:349:TYR:CD1	2.50	0.47
1:A:421:GLN:HG2	1:A:506:LYS:HZ1	1.80	0.46
1:A:685:ARG:NH2	1:A:695:GLU:OE1	2.49	0.46
1:A:353:GLN:HG2	1:A:363:PHE:HB3	1.97	0.46
1:A:427:ALA:HB1	1:A:592:VAL:HG12	1.99	0.45
1:A:554:ASP:N	1:A:554:ASP:OD1	2.50	0.45
1:A:508:GLN:HB2	1:A:511:VAL:HB	1.99	0.43
1:A:687:HIS:CD2	1:A:692:PRO:HB3	2.53	0.43
1:A:321:PRO:HB3	1:A:399:PRO:O	2.19	0.43
1:A:413:PRO:HB2	1:A:497:GLY:HA3	2.01	0.42
1:A:319:PRO:HG3	1:A:393:ASN:CG	2.39	0.42
1:A:525:ASP:HB2	1:A:574:PRO:HB3	2.03	0.41
1:A:642:TYR:HD1	1:A:686:ALA:HA	1.86	0.41
1:A:569:LEU:HD23	1:A:569:LEU:HA	1.95	0.40
1:A:430:MET:HG2	1:A:478:LEU:HD11	2.03	0.40
1:A:397:ARG:HH11	1:A:397:ARG:H	1.69	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	369/402 (92%)	356 (96%)	13 (4%)	0	100 100

There are no Ramachandran outliers to report.

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	320/344 (93%)	315 (98%)	5 (2%)	62 86

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	347	VAL
1	A	373	ARG
1	A	397	ARG
1	A	421	GLN
1	A	505	VAL

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

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	801	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	A	802	-	4,4,4	0.14	0	6,6,6	0.05	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

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, 95^{th} 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	375/402 (93%)	1.13	79 (21%) 1 0	113, 161, 214, 258	0

All (79) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	701	VAL	9.5
1	A	669	TRP	5.7
1	A	351	GLY	5.4
1	A	391	ALA	5.1
1	A	387	PHE	5.0
1	A	361	GLY	4.3
1	A	540	ILE	4.3
1	A	488	VAL	4.3
1	A	615	CYS	4.2
1	A	702	ARG	4.0
1	A	496	ASP	3.9
1	A	374	TYR	3.8
1	A	443	LEU	3.8
1	A	346	PRO	3.8
1	A	402	GLU	3.8
1	A	382	PHE	3.6
1	A	618	MET	3.6
1	A	680	TYR	3.5
1	A	333	THR	3.4
1	A	442	GLY	3.4
1	A	599	ALA	3.2
1	A	658	VAL	3.2
1	A	494	VAL	3.1
1	A	411	GLN	3.1
1	A	400	PRO	3.1
1	A	376	ILE	3.0
1	A	491	PHE	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	495	GLY	3.0
1	A	395	ILE	3.0
1	A	537	GLN	3.0
1	A	413	PRO	2.8
1	A	627	TRP	2.8
1	A	401	SER	2.8
1	A	490	ALA	2.8
1	A	355	ARG	2.7
1	A	334	SER	2.7
1	A	339	TRP	2.7
1	A	353	GLN	2.7
1	A	459	PRO	2.7
1	A	492	THR	2.6
1	A	446	GLY	2.6
1	A	700	LEU	2.6
1	A	616	VAL	2.5
1	A	622	THR	2.5
1	A	445	ARG	2.5
1	A	693	GLY	2.5
1	A	352	ILE	2.5
1	A	403	ALA	2.4
1	A	410	GLU	2.4
1	A	404	VAL	2.3
1	A	683	TRP	2.3
1	A	384	GLU	2.3
1	A	377	GLY	2.3
1	A	428	SER	2.3
1	A	645	ALA	2.3
1	A	664	ARG	2.3
1	A	624	ARG	2.2
1	A	423	ARG	2.2
1	A	399	PRO	2.2
1	A	326	VAL	2.2
1	A	349	TYR	2.2
1	A	409	GLY	2.2
1	A	534	LEU	2.2
1	A	448	ARG	2.2
1	A	337	LEU	2.2
1	A	363	PHE	2.1
1	A	367	ASP	2.1
1	A	672	VAL	2.1
1	A	427	ALA	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	405	ARG	2.1
1	A	389	VAL	2.1
1	A	406	ALA	2.1
1	A	325	LEU	2.1
1	A	471	LEU	2.1
1	A	652	GLU	2.1
1	A	390	LEU	2.1
1	A	394	SER	2.1
1	A	547	TYR	2.0
1	A	646	TYR	2.0

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 monosaccharides 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, 95^{th} 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	SO4	A	802	5/5	0.77	0.42	169,179,190,197	0
2	SO4	A	801	5/5	0.84	0.37	142,168,188,198	0

6.5 Other polymers (i)

There are no such residues in this entry.