Full wwPDB X-ray Structure Validation Report (i)

May 23, 2020 – 04:32 pm BST

PDB ID	:	3WPE
Title	:	Crystal structure of bovine TLR9 in complex with agonistic DNA1668_12mer
Authors	:	Ohto, U.; Tanji, H.; Shimizu, T.
Deposited on	:	2014-01-11
Resolution	:	2.38 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 following versions of software and data (see references (i)) were used in the production of this report:

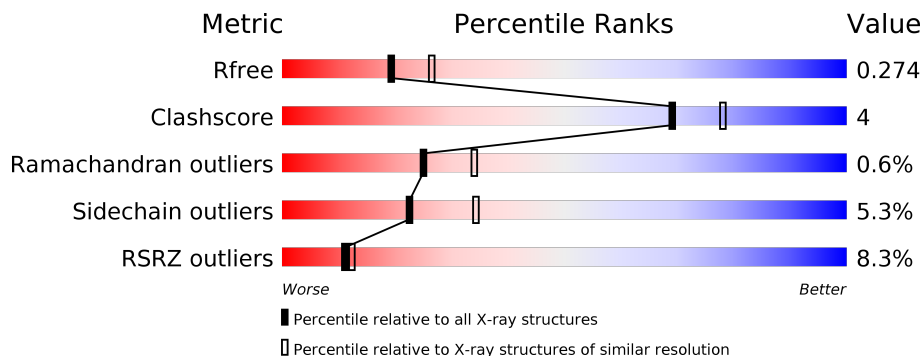|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.38 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	Percentile Ranks	Value
Rfree		0.274
Clashscore		4
Ramachandran outliers		0.6%
Sidechain outliers		5.3%
RSRZ outliers		8.3%

Worse ——————————————————— *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	5509 (2.40-2.36)
Clashscore	141614	6082 (2.40-2.36)
Ramachandran outliers	138981	5973 (2.40-2.36)
Sidechain outliers	138945	5975 (2.40-2.36)
RSRZ outliers	127900	5397 (2.40-2.36)

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	801	7% · 78% · 10% · 11%
2	C	12	17% · 33% · 50%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5783 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 Toll-like receptor 9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	714	Total 5628	C 3596	N 996	O 1013	S 23	0	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	21	ARG	-	EXPRESSION TAG	UNP Q5I2M5
A	22	SER	-	EXPRESSION TAG	UNP Q5I2M5
A	23	PRO	-	EXPRESSION TAG	UNP Q5I2M5
A	24	TRP	-	EXPRESSION TAG	UNP Q5I2M5
A	816	GLU	-	EXPRESSION TAG	UNP Q5I2M5
A	817	PHE	-	EXPRESSION TAG	UNP Q5I2M5
A	818	LEU	-	EXPRESSION TAG	UNP Q5I2M5
A	819	VAL	-	EXPRESSION TAG	UNP Q5I2M5
A	820	PRO	-	EXPRESSION TAG	UNP Q5I2M5
A	821	ARG	-	EXPRESSION TAG	UNP Q5I2M5

- Molecule 2 is a DNA chain called DNA (5'-D(*CP*AP*TP*GP*AP*CP*GP*TP*TP*CP*CP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	6	Total 124	C 59	N 22	O 37	P 6	0	0	0

- Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	31	Total 31	O 31	0	0

3 Residue-property plots (i)

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: Toll-like receptor 9
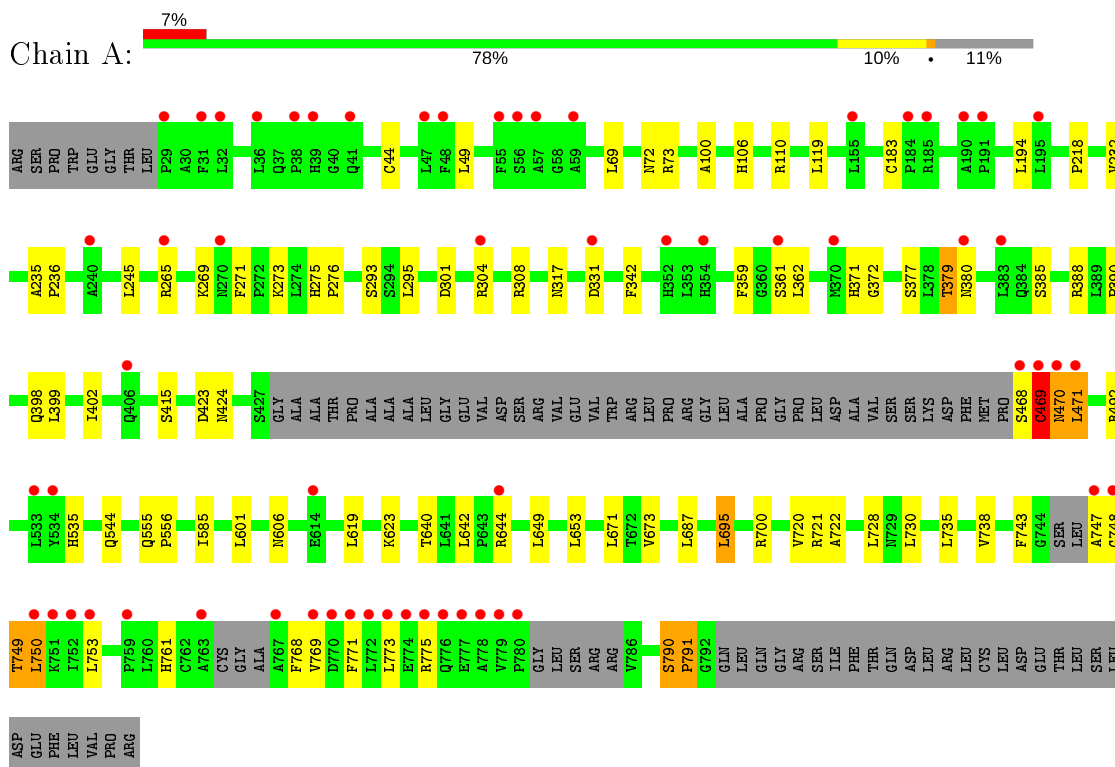
- Molecule 2: DNA (5'-D(*CP*AP*TP*GP*AP*CP*GP*TP*TP*CP*CP*T)-3')
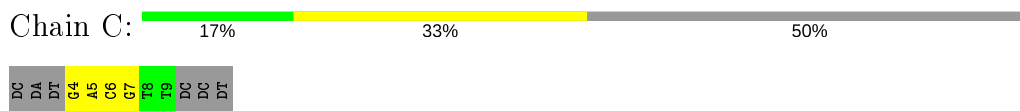
4   Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	115.82Å 113.08Å 67.43Å 90.00° 100.61° 90.00°	Depositor
Resolution (Å)	43.05 – 2.38 43.01 – 2.38	Depositor EDS
% Data completeness (in resolution range)	99.7 (43.05-2.38) 99.8 (43.01-2.38)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.92 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.217 , 0.274 0.218 , 0.274	Depositor DCC
R_{free} test set	1717 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	37.8	Xtriage
Anisotropy	0.260	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 48.3	EDS
L-test for twinning[2]	$< |L| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	5783	wwPDB-VP
Average B, all atoms (Å^2)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.07% 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 ⓘ

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	0.33	0/5759	0.58	0/7827
2	C	0.29	0/138	0.69	0/211
All	All	0.33	0/5897	0.58	0/8038

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	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	469	CYS	Peptide

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	5628	0	5661	40	0
2	C	124	0	69	4	0
3	A	31	0	0	2	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	5783	0	5730	41	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 (41) 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:73:ARG:NH1	2:C:7:DG:O6	2.16	0.78
1:A:470:ASN:O	1:A:471:LEU:HD23	1.93	0.68
1:A:232:VAL:HG12	1:A:271:PHE:CD1	2.34	0.62
1:A:738:VAL:O	1:A:768:PHE:HB3	2.00	0.62
1:A:73:ARG:NH2	2:C:6:DC:OP2	2.36	0.59
1:A:275:HIS:HB2	1:A:276:PRO:HD2	1.85	0.58
1:A:642:LEU:HD13	1:A:644:ARG:HD2	1.85	0.58
1:A:390:PRO:O	1:A:415:SER:OG	2.18	0.58
1:A:379:THR:OG1	1:A:380:ASN:N	2.38	0.56
1:A:304:ARG:NH1	1:A:331:ASP:OD2	2.39	0.54
1:A:695:LEU:HD23	1:A:722:ALA:HB2	1.89	0.54
1:A:232:VAL:CG1	1:A:271:PHE:CD1	2.91	0.54
1:A:469:CYS:O	1:A:469:CYS:SG	2.66	0.53
1:A:44:CYS:HB2	1:A:69:LEU:HD23	1.91	0.52
1:A:295:LEU:H	1:A:317:ASN:HD22	1.59	0.51
1:A:73:ARG:HH21	2:C:5:DA:H3'	1.76	0.50
1:A:235:ALA:HB1	1:A:236:PRO:HD2	1.93	0.50
1:A:399:LEU:HD23	1:A:423:ASP:HB3	1.94	0.50
1:A:555:GLN:HB3	1:A:556:PRO:HD3	1.94	0.49
1:A:769:VAL:O	1:A:773:LEU:HD13	2.13	0.49
1:A:276:PRO:O	3:A:917:HOH:O	2.20	0.48
1:A:276:PRO:HB3	1:A:301:ASP:CB	2.43	0.48
1:A:293:SER:H	1:A:317:ASN:HD21	1.61	0.48
1:A:100:ALA:O	1:A:106:HIS:HA	2.14	0.47
1:A:359:PHE:O	1:A:362:LEU:HB2	2.15	0.46
1:A:194:LEU:O	1:A:218:PRO:HG3	2.14	0.46
1:A:49:LEU:H	1:A:72:ASN:HD22	1.66	0.44
2:C:4:DG:H2"	2:C:5:DA:C8	2.53	0.44
1:A:402:ILE:H	1:A:424:ASN:HD22	1.65	0.43
1:A:585:ILE:H	1:A:606:ASN:HD22	1.67	0.43
1:A:790:SER:HA	1:A:791:PRO:HA	1.89	0.43
1:A:775:ARG:HD2	1:A:775:ARG:HA	1.77	0.43
1:A:468:SER:OG	1:A:469:CYS:N	2.50	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:342:PHE:HA	1:A:372:GLY:O	2.19	0.42
1:A:761:HIS:ND1	1:A:790:SER:HB3	2.34	0.42
1:A:750:LEU:HD12	1:A:753:LEU:HB2	2.00	0.42
1:A:271:PHE:CE1	1:A:273:LYS:HD2	2.55	0.42
1:A:747:ALA:O	1:A:749:THR:N	2.53	0.42
1:A:728:LEU:HG	1:A:730:LEU:HD13	2.02	0.41
1:A:371:HIS:HB2	1:A:398:GLN:O	2.20	0.41
1:A:535:HIS:HE1	3:A:922:HOH:O	2.05	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	704/801 (88%)	653 (93%)	47 (7%)	4 (1%)	25	34

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	720	VAL
1	A	361	SER
1	A	748	GLY
1	A	791	PRO

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	641/712 (90%)	607 (95%)	34 (5%)	22 34

All (34) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	110	ARG
1	A	119	LEU
1	A	183	CYS
1	A	245	LEU
1	A	265	ARG
1	A	269	LYS
1	A	308	ARG
1	A	377	SER
1	A	379	THR
1	A	385	SER
1	A	388	ARG
1	A	469	CYS
1	A	470	ASN
1	A	471	LEU
1	A	492	ARG
1	A	544	GLN
1	A	601	LEU
1	A	619	LEU
1	A	623	LYS
1	A	640	THR
1	A	649	LEU
1	A	653	LEU
1	A	671	LEU
1	A	673	VAL
1	A	687	LEU
1	A	695	LEU
1	A	700	ARG
1	A	721	ARG
1	A	735	LEU
1	A	743	PHE
1	A	749	THR
1	A	750	LEU
1	A	771	PHE
1	A	790	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	72	ASN
1	A	132	ASN
1	A	317	ASN
1	A	330	ASN
1	A	334	GLN
1	A	396	HIS
1	A	400	ASN
1	A	424	ASN
1	A	544	GLN
1	A	583	ASN
1	A	606	ASN
1	A	627	ASN
1	A	702	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 carbohydrates in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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	714/801 (89%)	0.57	60 (8%)	11	12	23, 47, 78, 111	0
2	C	6/12 (50%)	-0.18	0	100	100	54, 56, 60, 65	0
All	All	720/813 (88%)	0.57	60 (8%)	11	12	23, 47, 78, 111	0

All (60) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	773	LEU	6.6
1	A	778	ALA	6.3
1	A	36	LEU	6.2
1	A	767	ALA	5.3
1	A	468	SER	4.8
1	A	469	CYS	4.8
1	A	772	LEU	4.6
1	A	777	GLU	4.5
1	A	779	VAL	4.4
1	A	184	PRO	4.3
1	A	770	ASP	4.2
1	A	57	ALA	4.2
1	A	380	ASN	4.1
1	A	39	HIS	4.0
1	A	29	PRO	4.0
1	A	240	ALA	3.7
1	A	771	PHE	3.5
1	A	32	LEU	3.5
1	A	751	LYS	3.4
1	A	470	ASN	3.4
1	A	780	PRO	3.3
1	A	471	LEU	3.3
1	A	774	GLU	3.3
1	A	331	ASP	3.2

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	750	LEU	3.2
1	A	406	GLN	3.1
1	A	265	ARG	3.1
1	A	747	ALA	3.1
1	A	59	ALA	3.0
1	A	769	VAL	3.0
1	A	41	GLN	2.9
1	A	190	ALA	2.8
1	A	533	LEU	2.8
1	A	55	PHE	2.8
1	A	56	SER	2.7
1	A	191	PRO	2.6
1	A	354	HIS	2.6
1	A	748	GLY	2.6
1	A	48	PHE	2.5
1	A	775	ARG	2.5
1	A	361	SER	2.4
1	A	763	ALA	2.4
1	A	753	LEU	2.3
1	A	776	GLN	2.3
1	A	752	ILE	2.3
1	A	38	PRO	2.3
1	A	534	TYR	2.3
1	A	31	PHE	2.3
1	A	47	LEU	2.3
1	A	352	HIS	2.3
1	A	759	PRO	2.3
1	A	185	ARG	2.2
1	A	195	LEU	2.2
1	A	383	LEU	2.2
1	A	155	LEU	2.1
1	A	644	ARG	2.1
1	A	270	ASN	2.1
1	A	304	ARG	2.0
1	A	370	MET	2.0
1	A	614	GLU	2.0

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

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.