Preliminary Full wwPDB X-ray Structure Validation Report (i)

Nov 10, 2014 – 09:04 AM EST

DISCLAIMER
This is a preliminary version of the new style of wwPDB validation report.
This report is produced by the wwPDB validation pipeline
before deposition or annotation of the structure.
This is not an official wwPDB validation report and is not a proof of deposition.
This report should not be submitted to journals.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at http://wwpdb.org/ValidationPDFNotes.html

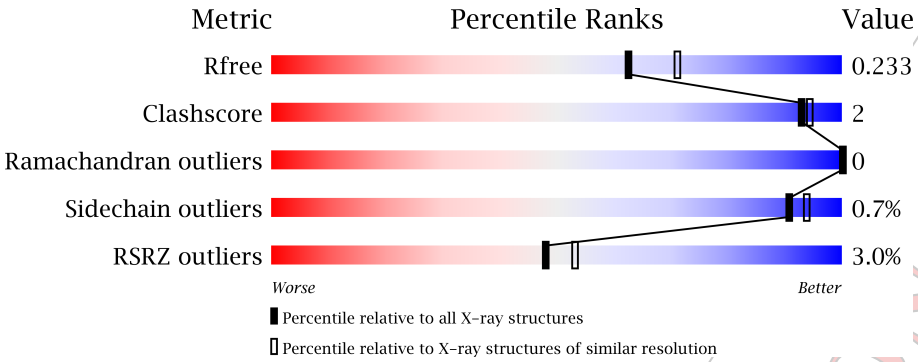
The following versions of software and data (see references) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.16 November 2013
Xtriage (Phenix)	:	dev-1439
EDS	:	stable24103
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable24103

1 Overall quality at a glance

The reported resolution of this entry is 2.10 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}	66092	3012 (2.10-2.10)
Clashscore	79885	3649 (2.10-2.10)
Ramachandran outliers	78287	3610 (2.10-2.10)
Sidechain outliers	78261	3611 (2.10-2.10)
RSRZ outliers	66119	3013 (2.10-2.10)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	364	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 2831 atoms, of which 51 are hydrogens and 0 are deuterium.

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 MITOGEN-ACTIVATED PROTEIN KINASE 10.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	333	2636	1700	442	475	19	0	1	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	MET	-	CLONING ARTIFACT	UNP P53779
A	?	ALA	-	CLONING ARTIFACT	UNP P53779

- Molecule 2 is PHOSPHOAMINOPHOSPHONICACID-ADENYLATE ESTER (three-letter code: ANP) (formula: unknown).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	H	N	O	P		
2	A	1	44	10	13	6	12	3	13	0

- Molecule 3 is CYCLOHEXYL-{4-[5-(3,4-DICHLOROPHENYL)-2-PIPERIDIN-4-YL-3-PROPYL-3H-IMIDAZOL-4-YL]-PYRIMIDIN-2-YL}AMINE (three-letter code: 880) (formula: unknown).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	Cl	H	N		
3	A	1	81	33	4	38	6	38	1

- Molecule 4 is UNKNOWN ATOM OR ION (three-letter code: UNX) (formula: unknown).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
4	A	3	3	3	0	0

- Molecule 5 is water.

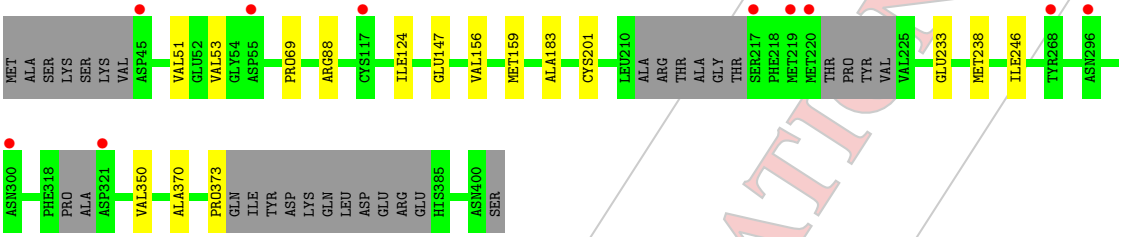
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	67	Total O 67 67	0	1

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 errors 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: MITOGEN-ACTIVATED PROTEIN KINASE 10

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	50.95Å 71.53Å 106.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	16.94 – 2.10 16.87 – 2.10	Depositor EDS
% Data completeness (in resolution range)	84.3 (16.94-2.10) 84.3 (16.87-2.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	10.55 (at 2.11Å)	Xtriage
Refinement program	BUSTER 2.11.6	Depositor
R, R_{free}	0.201 , 0.225 0.203 , 0.233	Depositor DCC
R_{free} test set	957 reflections (5.14%)	DCC
Wilson B-factor (Å^2)	22.3	Xtriage
Anisotropy	0.091	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 40.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 19586 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	2831	wwPDB-VP
Average B, all atoms (Å^2)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.25% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: UNX, 880, ANP

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	0.49	0/2698	0.61	0/3656

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	370	ALA	Mainchain

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2636	0	2585	9	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	31	13	0	0	0
3	A	43	38	0	2	0
4	A	3	0	0	0	0
5	A	67	0	0	0	0
All	All	2780	51	2585	9	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 2.

All (9) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:53:VAL:HG13	1:A:69:PRO:HG3	1.76	0.66
1:A:88:ARG:HH21	1:A:147:GLU:HG3	1.63	0.63
1:A:124:ILE:HD12	3:A:501[B]:880:CL45	2.36	0.62
1:A:124:ILE:CD1	3:A:501[B]:880:CL45	2.85	0.61
1:A:183:ALA:HB2	1:A:373:PRO:HD2	1.88	0.55
1:A:233:GLU:HA	1:A:238:MET:HE3	1.88	0.54
1:A:246:ILE:HD11	1:A:350:VAL:HA	1.97	0.47
1:A:238:MET:HE2	1:A:238:MET:HB2	1.79	0.43
1:A:156:VAL:O	1:A:159:MET:HG2	2.19	0.42

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	324/364 (89%)	313 (97%)	11 (3%)	0	100 100

There are no Ramachandran outliers to report.

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	282/328 (86%)	280 (99%)	2 (1%)	91 94

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	51	VAL
1	A	201	CYS

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

5.3.3 RNA

There are no RNA chains in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

Of 6 ligands modelled in this entry, 3 are unknown - leaving 3 for Mogul analysis.

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	880	A	501[A]	-	39,?,?	0.35	0	54,?,?	0.96	2 (3%)
3	880	A	501[B]	-	39,?,?	0.35	0	54,?,?	0.91	2 (3%)
2	ANP	A	502	-	33,?,?	0.88	2 (6%)	48,?,?	0.84	2 (4%)

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
3	880	A	501[A]	-	-	0/19/?/?	0/5/?/?
3	880	A	501[B]	-	-	0/19/?/?	0/5/?/?
2	ANP	A	502	-	-	0/18/?/?	0/3/?/?

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	502	ANP	PG-N3B	3.53	1.67	1.64
2	A	502	ANP	PG-O1G	2.31	1.48	1.46

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	501[B]	880	C50-N54-C55	4.71	128.43	123.42
3	A	501[A]	880	C50-N54-C55	4.71	128.43	123.42
2	A	502	ANP	O2G-PG-O3G	-2.95	99.23	107.69
3	A	501[B]	880	C16-C2-N1	2.41	126.64	123.47
3	A	501[A]	880	C16-C2-N1	2.41	126.64	123.47
2	A	502	ANP	PB-N3B-PG	2.14	133.66	130.07

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues

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	333/364 (91%)	-0.04	10 (3%) 48 53	9, 28, 53, 84	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	117	CYS	6.1
1	A	217	SER	5.9
1	A	219	MET	5.2
1	A	55	ASP	3.5
1	A	268	TYR	3.2
1	A	45	ASP	2.7
1	A	296	ASN	2.6
1	A	321	ASP	2.6
1	A	300	ASN	2.4
1	A	220	MET	2.1

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

There are no carbohydrates in this entry.

6.4 Ligands

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(Å^2)	Q<0.9
2	ANP	A	502	31/?	0.15	1.52	24,35,42,44	44
4	UNX	A	2001	1/?	0.08	-0.74	52,52,52,52	0
3	880	A	501[A]	35/?	0.07	-1.38	4,13,22,24	43
3	880	A	501[B]	35/?	0.07	-1.47	6,13,22,24	43
4	UNX	A	2002	1/?	0.06	-1.78	59,59,59,59	0
4	UNX	A	2003	1/?	0.06	-1.96	53,53,53,53	0

6.5 Other polymers (i)

There are no such residues in this entry.