Full wwPDB X-ray Structure Validation Report (i)

Sep 14, 2015 – 10:58 PM EDT

PDB ID	:	5DNO
Title	:	Structure of a protein-RNA complex
Authors	:	Wang, C.Y.; Zhu, Y.W.; Wu, J.H.; Shi, Y.Y.
Deposited on	:	2015-09-10
Resolution	:	1.80 Å(reported)

DISCLAIMER
This is a preliminary version of the new style of wwPDB validation report.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at http://wwpdb.org/ValidationPDFNotes.html

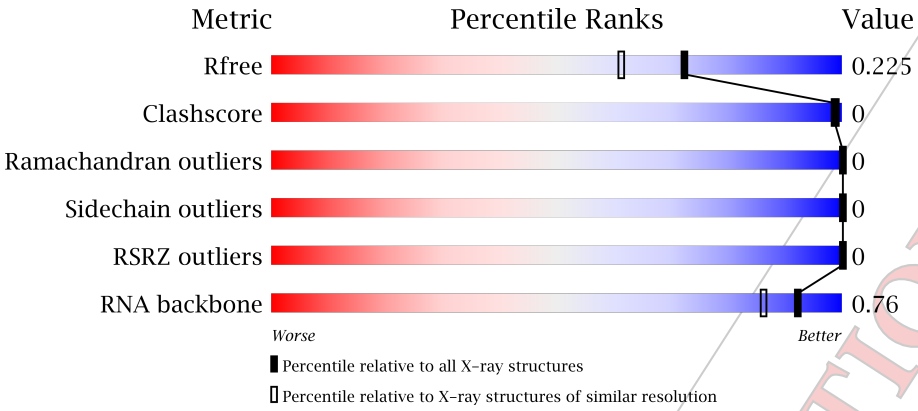
The following versions of software and data (see references) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.17 November 2013
Xtriage (Phenix)	:	dev-1439
EDS	:	stable24195
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)	:	stable24195

1 Overall quality at a glance

The reported resolution of this entry is 1.80 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	3513 (1.80-1.80)
Clashscore	79885	4461 (1.80-1.80)
Ramachandran outliers	78287	4404 (1.80-1.80)
Sidechain outliers	78261	4403 (1.80-1.80)
RSRZ outliers	66119	3515 (1.80-1.80)
RNA backbone	1838	1001 (2.92-0.96)

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	178	
2	B	7	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 1563 atoms, of which 0 are hydrogen 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 YTH domain-containing protein mmi1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	165	1304	816	232	245	11	0	0	1

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	311	MET	-	expression tag	UNP O74958
A	312	GLY	-	expression tag	UNP O74958
A	313	SER	-	expression tag	UNP O74958
A	314	SER	-	expression tag	UNP O74958
A	315	HIS	-	expression tag	UNP O74958
A	316	HIS	-	expression tag	UNP O74958
A	317	HIS	-	expression tag	UNP O74958
A	318	HIS	-	expression tag	UNP O74958
A	319	HIS	-	expression tag	UNP O74958
A	320	ALA	-	expression tag	UNP O74958
A	321	MET	-	expression tag	UNP O74958

- Molecule 2 is a RNA chain called RNA (5'-R(*CP*UP*UP*AP*AP*AP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	7	143	66	25	46	6	0	0	0

- Molecule 3 is water.

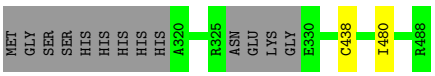
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	106	106	106	0	0
			Total	O		
3	B	10	10	10	0	0

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: YTH domain-containing protein mmi1

Chain A:



- Molecule 2: RNA (5'-R(*CP*UP*UP*AP*AP*AP*C)-3')

Chain B:

There are no outlier residues recorded for this chain.

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α, β, γ	77.53Å 77.53Å 65.53Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	38.76 – 1.80 38.76 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.9 (38.76-1.80) 99.9 (38.76-1.80)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.44 (at 1.79Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.176 , 0.216 0.186 , 0.225	Depositor DCC
R_{free} test set	1084 reflections (5.47%)	DCC
Wilson B-factor (Å^2)	10.2	Xtriage
Anisotropy	0.179	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 35.8	EDS
Estimated twinning fraction	0.199 for h,-h-k,-l	Xtriage
L-test for twinning	$< \|L\| > = 0.43$, $< L^2 > = 0.26$	Xtriage
Outliers	0 of 20917 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	1563	wwPDB-VP
Average B, all atoms (Å^2)	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.50% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

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	0.47	0/1328	0.68	0/1782
2	B	0.44	0/159	0.76	0/245
All	All	0.47	0/1487	0.69	0/2027

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	1304	0	1273	1	0
2	B	143	0	77	0	0
3	A	106	0	0	0	0
3	B	10	0	0	0	0
All	All	1563	0	1350	1	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 0.

All (1) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:438:CYS:HB3	1:A:480:ILE:HD11	1.90	0.54

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	161/178 (90%)	161 (100%)	0	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	140/154 (91%)	140 (100%)	0	100 100

There are no protein residues with a non-rotameric sidechain to report.

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

5.3.3 RNA

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	0/7	-	-

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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

There are no ligands in this entry.

5.7 Other polymers

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	165/178 (92%)	-0.44	0	100	100	6, 10, 26, 37	0
2	B	7/7 (100%)	-0.40	0	100	100	8, 11, 35, 50	0
All	All	172/185 (92%)	-0.44	0	100	100	6, 11, 26, 50	0

There are no RSRZ outliers to report.

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.