



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 17, 2014 – 09:59 PM EDT

PDB ID : 4PW7  
Title : structure of UHRF2-SRA in complex with a 5mC-containing DNA  
Authors : ZHou, T.; Xiong, J.; Wang, M.; Yang, N.; Wong, J.; Zhu, B.; Xu, R.M.  
Deposited on : 2014-03-19  
Resolution : 2.00 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

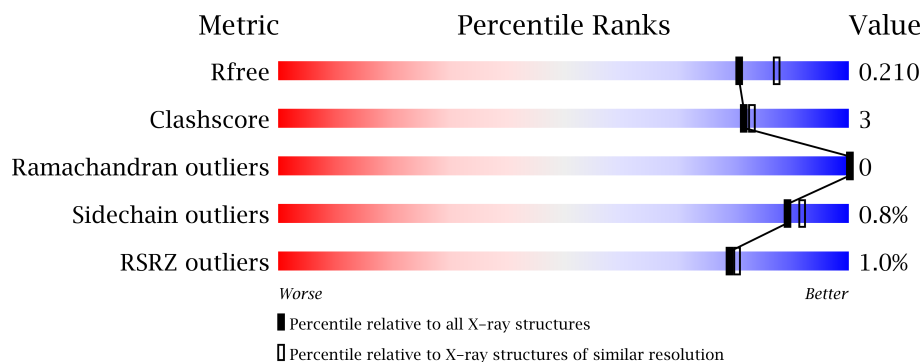
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : **FAILED**  
Xtriage (Phenix) : dev-1439  
EDS : stable23161  
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) : stable23161

# 1 Overall quality at a glance

The reported resolution of this entry is 2.00 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	4888 (2.00-2.00)
Clashscore	79885	6188 (2.00-2.00)
Ramachandran outliers	78287	6102 (2.00-2.00)
Sidechain outliers	78261	6100 (2.00-2.00)
RSRZ outliers	66119	4890 (2.00-2.00)

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  $\geq 3$ , 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	230	
1	B	230	
1	E	230	
1	F	230	
2	C	12	
2	G	12	
3	D	12	
3	H	12	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8173 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 E3 ubiquitin-protein ligase UHRF2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	197	Total	C	N	O	S	0	10	0
			1602	1016	300	283	3			
1	B	189	Total	C	N	O	S	0	8	0
			1532	967	284	278	3			
1	E	198	Total	C	N	O	S	0	11	0
			1614	1020	304	286	4			
1	F	191	Total	C	N	O	S	0	4	0
			1519	956	280	279	4			

- Molecule 2 is a DNA chain called 5mC-containing DNA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	12	Total	C	N	O	P	0	0	0
			249	118	45	74	12			
2	G	12	Total	C	N	O	P	0	0	0
			249	118	45	74	12			

- Molecule 3 is a DNA chain called 5mC-containing DNA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	12	Total	C	N	O	P	0	0	0
			238	114	45	68	11			
3	H	12	Total	C	N	O	P	0	0	0
			238	114	45	68	11			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	194	Total	O	0	0
			194	194		
4	B	201	Total	O	0	0
			201	201		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	50	Total 50	O 50	0	0
4	D	24	Total 24	O 24	0	0
4	E	199	Total 199	O 199	0	0
4	F	203	Total 203	O 203	0	0
4	G	46	Total 46	O 46	0	0
4	H	15	Total 15	O 15	0	0



Chain C: 



- Molecule 2: 5mC-containing DNA1

Chain G: 



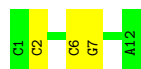
- Molecule 3: 5mC-containing DNA2

Chain D: 



- Molecule 3: 5mC-containing DNA2

Chain H: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	42.33Å 83.69Å 83.69Å 89.97° 89.94° 90.02°	Depositor
Resolution (Å)	28.05 – 2.00 29.77 – 2.00	Depositor EDS
% Data completeness (in resolution range)	97.8 (28.05-2.00) 97.7 (29.77-2.00)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.82 (at 2.00Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4.1491)	Depositor
R, $R_{free}$	0.163 , 0.211 0.164 , 0.210	Depositor DCC
$R_{free}$ test set	3802 reflections (5.02%)	DCC
Wilson B-factor (Å <sup>2</sup> )	17.3	Xtriage
Anisotropy	0.116	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 45.8	EDS
Estimated twinning fraction	0.447 for h,l,-k 0.447 for h,-l,k 0.427 for h,-k,-l 0.048 for -h,-k,l 0.047 for -h,k,-l 0.047 for -h,-l,-k 0.048 for -h,l,k	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 75800 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	8173	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.15% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>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: 5CM

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	0/1674	0.49	0/2266
1	B	0.33	0/1597	0.46	0/2163
1	E	0.33	0/1689	0.49	0/2285
1	F	0.33	0/1572	0.47	0/2131
2	C	0.71	0/255	0.93	0/390
2	G	0.71	0/255	0.99	1/390 (0.3%)
3	D	0.71	0/266	0.79	0/407
3	H	0.66	0/266	0.82	0/407
All	All	0.40	0/7574	0.56	1/10439 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	G	12	DG	O4'-C1'-N9	5.47	111.83	108.00

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	1602	0	1611	17	0
1	B	1532	0	1510	6	0
1	E	1614	0	1621	10	0
1	F	1519	0	1481	7	0
2	C	249	0	138	1	0
2	G	249	0	138	2	0
3	D	238	0	135	4	0
3	H	238	0	135	4	0
4	A	194	0	0	6	0
4	B	201	0	0	3	0
4	C	50	0	0	0	0
4	D	24	0	0	1	0
4	E	199	0	0	2	0
4	F	203	0	0	3	0
4	G	46	0	0	0	0
4	H	15	0	0	0	0
All	All	8173	0	6769	44	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 3.

All (44) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:555:ASN:OD1	4:A:839:HOH:O	1.99	0.80
1:F:555:ASN:OD1	4:F:848:HOH:O	2.03	0.76
3:H:6:DC:H2"	3:H:7:DG:H5"	1.66	0.76
1:A:520:ARG:HE	3:D:7:DG:H21	1.34	0.75
1:E:520:ARG:HE	3:H:7:DG:H21	1.34	0.74
2:G:12:DG:O6	3:H:2:DC:N4	2.20	0.67
1:B:555:ASN:OD1	4:B:834:HOH:O	2.12	0.66
1:E:601:HIS:NE2	4:E:870:HOH:O	2.30	0.64
1:E:544:PRO:HD2	1:E:554[B]:ARG:HH11	1.64	0.61
1:B:554[B]:ARG:NH2	4:B:891:HOH:O	2.33	0.60
1:A:520:ARG:HE	3:D:7:DG:N2	1.99	0.58
1:E:520:ARG:HE	3:H:7:DG:N2	2.02	0.58
1:A:441:ILE:HG22	1:A:442:VAL:HG13	1.85	0.57
1:F:634:ARG:NH1	4:F:807:HOH:O	2.38	0.56
1:F:571:ARG:HD3	1:F:577:ALA:HB3	1.87	0.55
1:A:472:ARG:NH1	4:A:876:HOH:O	2.40	0.53
1:A:520:ARG:NH1	4:A:887:HOH:O	2.42	0.52
1:F:504:GLU:OE2	1:F:611:ARG:NE	2.39	0.51
1:A:571[A]:ARG:NH1	4:A:885:HOH:O	2.46	0.48
1:E:571[A]:ARG:NH1	4:E:808:HOH:O	2.35	0.48

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:E:562[B]:VAL:HG12	1:E:588:TYR:O	2.14	0.48
1:E:521:ILE:HD11	2:G:7:5CM:H4'	1.96	0.47
1:E:571[B]:ARG:HG3	1:E:577:ALA:HB3	1.97	0.46
1:F:542:ASP:HB3	1:F:553[B]:SER:OG	2.16	0.45
1:A:470:VAL:HG12	1:A:565:ILE:HD11	2.00	0.44
1:E:542:ASP:HB3	1:E:553[B]:SER:OG	2.18	0.44
1:B:542:ASP:HB3	1:B:553[B]:SER:OG	2.19	0.43
1:A:476:GLY:HA3	4:A:811:HOH:O	2.18	0.43
1:A:491:LEU:HD11	1:A:505:PHE:CE1	2.52	0.43
3:D:12:DA:N3	4:D:117:HOH:O	2.36	0.43
1:E:489:LEU:HD11	1:E:564:VAL:HG22	2.00	0.43
1:F:572:LYS:HB2	4:F:881:HOH:O	2.18	0.42
3:D:1:DC:C2'	3:D:2:DC:H5'	2.50	0.41
1:A:542:ASP:OD1	1:A:554[B]:ARG:NH2	2.54	0.41
1:A:641:TYR:HA	1:A:642:PRO:HD3	1.81	0.41
1:F:571:ARG:NH1	1:F:579:GLU:OE1	2.53	0.41
1:A:629:ARG:NH2	4:A:761:HOH:O	2.27	0.41
1:A:513:LYS:HG2	1:A:524:PRO:HA	2.01	0.41
1:B:462:ARG:NH1	1:B:476:GLY:O	2.50	0.41
1:A:488:SER:HA	1:A:562[B]:VAL:HG23	2.02	0.41
1:B:473:PRO:CD	1:B:479:HIS:HB2	2.51	0.41
1:A:571[A]:ARG:HG2	1:A:577:ALA:HB3	2.02	0.41
1:B:554[A]:ARG:HD3	4:B:898:HOH:O	2.21	0.40
1:A:520:ARG:O	2:C:6:DC:H2''	2.22	0.40

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	203/230 (88%)	199 (98%)	4 (2%)	0	100	100
1	B	193/230 (84%)	190 (98%)	3 (2%)	0	100	100
1	E	205/230 (89%)	198 (97%)	7 (3%)	0	100	100

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	191/230 (83%)	187 (98%)	4 (2%)	0	100	100
All	All	792/920 (86%)	774 (98%)	18 (2%)	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	170/186 (91%)	165 (97%)	5 (3%)	55	52
1	B	163/186 (88%)	163 (100%)	0	100	100
1	E	172/186 (92%)	172 (100%)	0	100	100
1	F	161/186 (87%)	159 (99%)	2 (1%)	82	84
All	All	666/744 (90%)	659 (99%)	7 (1%)	89	86

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	554[A]	ARG
1	A	554[B]	ARG
1	A	571[A]	ARG
1	A	571[B]	ARG
1	A	634	ARG
1	F	440	THR
1	F	571	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	601	HIS
1	A	636	GLN
1	B	601	HIS
1	B	636	GLN
1	E	636	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	F	601	HIS

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

Mogul failed to run properly - this section will therefore be empty.

## 5.5 Carbohydrates ⓘ

Mogul failed to run properly - this section will therefore be empty.

## 5.6 Ligand geometry ⓘ

Mogul failed to run properly - this section will therefore be empty.

## 5.7 Other polymers ⓘ

Mogul failed to run properly - this section will therefore be empty.

## 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	197/230 (85%)	-0.34	1 (0%) 88 89	8, 15, 45, 53	0
1	B	189/230 (82%)	-0.38	0 100 100	9, 15, 39, 54	0
1	E	198/230 (86%)	-0.28	5 (2%) 54 54	8, 15, 47, 58	0
1	F	191/230 (83%)	-0.37	2 (1%) 79 80	9, 15, 38, 54	0
2	C	12/12 (100%)	0.06	0 100 100	16, 28, 38, 44	0
2	G	12/12 (100%)	-0.10	0 100 100	20, 29, 40, 42	0
3	D	12/12 (100%)	-0.07	0 100 100	26, 36, 55, 59	0
3	H	12/12 (100%)	0.01	0 100 100	29, 37, 58, 59	0
All	All	823/968 (85%)	-0.33	8 (0%) 79 80	8, 15, 46, 59	0

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	441	ILE	4.4
1	E	601	HIS	3.6
1	E	441	ILE	3.4
1	E	523	ALA	3.3
1	E	524	PRO	2.7
1	E	440	THR	2.6
1	F	441	ILE	2.5
1	F	525	SER	2.2

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

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<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	5CM	C	7	20/21	0.12	0.46	11,15,20,20	0
2	5CM	G	7	20/21	0.10	-0.10	14,18,26,28	0

### 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.