



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 20, 2023 – 06:05 PM JST

PDB ID : 7D3T  
Title : Crystal structure of OSPHR2 in complex with DNA  
Authors : Guan, Z.Y.; Zhang, Z.F.; Liu, Z.  
Deposited on : 2020-09-20  
Resolution : 2.70 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](mailto: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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

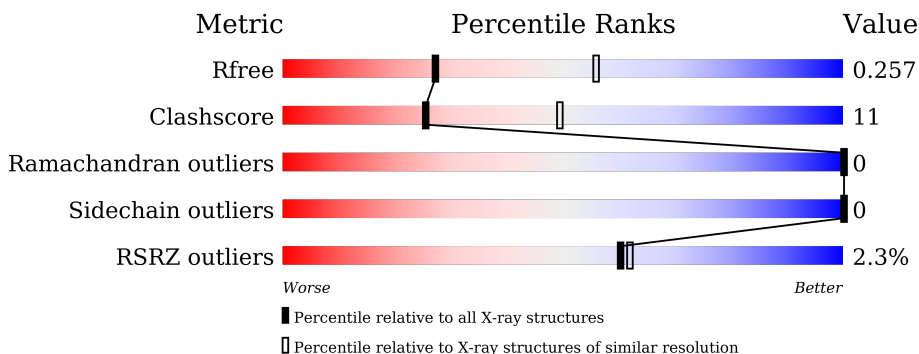
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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.




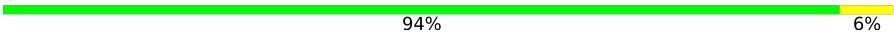
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$ . 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	144	
1	B	144	
1	C	144	
1	D	144	
2	E	16	
2	G	16	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	16	 62% 38%
3	H	16	 94% 6%

## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 3289 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 Protein PHOSPHATE STARVATION RESPONSE 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	59	487	310	93	82	2	0	0	0
1	B	59	487	310	93	82	2	0	0	0
1	C	59	487	310	93	82	2	0	0	0
1	D	62	510	324	96	88	2	0	0	0

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	314	LEU	-	expression tag	UNP B8B5N8
A	315	GLU	-	expression tag	UNP B8B5N8
A	316	HIS	-	expression tag	UNP B8B5N8
A	317	HIS	-	expression tag	UNP B8B5N8
A	318	HIS	-	expression tag	UNP B8B5N8
A	319	HIS	-	expression tag	UNP B8B5N8
A	320	HIS	-	expression tag	UNP B8B5N8
A	321	HIS	-	expression tag	UNP B8B5N8
A	322	HIS	-	expression tag	UNP B8B5N8
A	323	HIS	-	expression tag	UNP B8B5N8
B	314	LEU	-	expression tag	UNP B8B5N8
B	315	GLU	-	expression tag	UNP B8B5N8
B	316	HIS	-	expression tag	UNP B8B5N8
B	317	HIS	-	expression tag	UNP B8B5N8
B	318	HIS	-	expression tag	UNP B8B5N8
B	319	HIS	-	expression tag	UNP B8B5N8
B	320	HIS	-	expression tag	UNP B8B5N8
B	321	HIS	-	expression tag	UNP B8B5N8
B	322	HIS	-	expression tag	UNP B8B5N8
B	323	HIS	-	expression tag	UNP B8B5N8
C	314	LEU	-	expression tag	UNP B8B5N8

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	315	GLU	-	expression tag	UNP B8B5N8
C	316	HIS	-	expression tag	UNP B8B5N8
C	317	HIS	-	expression tag	UNP B8B5N8
C	318	HIS	-	expression tag	UNP B8B5N8
C	319	HIS	-	expression tag	UNP B8B5N8
C	320	HIS	-	expression tag	UNP B8B5N8
C	321	HIS	-	expression tag	UNP B8B5N8
C	322	HIS	-	expression tag	UNP B8B5N8
C	323	HIS	-	expression tag	UNP B8B5N8
D	314	LEU	-	expression tag	UNP B8B5N8
D	315	GLU	-	expression tag	UNP B8B5N8
D	316	HIS	-	expression tag	UNP B8B5N8
D	317	HIS	-	expression tag	UNP B8B5N8
D	318	HIS	-	expression tag	UNP B8B5N8
D	319	HIS	-	expression tag	UNP B8B5N8
D	320	HIS	-	expression tag	UNP B8B5N8
D	321	HIS	-	expression tag	UNP B8B5N8
D	322	HIS	-	expression tag	UNP B8B5N8
D	323	HIS	-	expression tag	UNP B8B5N8

- Molecule 2 is a DNA chain called DNA (5'-D(P\*GP\*CP\*TP\*TP\*GP\*AP\*GP\*GP\*AP\*TP\*AP\*TP\*CP\*CP\*GP\*A)-3').

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
2	E	16	Total	C	N	O	P	0	0	0
			331	157	62	96	16			
2	G	16	Total	C	N	O	P	0	0	0
			331	157	62	96	16			

- Molecule 3 is a DNA chain called DNA (5'-D(P\*CP\*TP\*CP\*GP\*GP\*AP\*TP\*AP\*TP\*CP\*CP\*TP\*CP\*AP\*AP\*G)-3').

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
3	F	16	Total	C	N	O	P	0	0	0
			325	155	58	96	16			
3	H	16	Total	C	N	O	P	0	0	0
			322	155	58	94	15			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	2	Total O 2 2	0	0
4	B	1	Total O 1 1	0	0
4	D	1	Total O 1 1	0	0
4	E	1	Total O 1 1	0	0
4	F	1	Total O 1 1	0	0
4	H	3	Total O 3 3	0	0



- Molecule 2: DNA (5'-D(P\*GP\*CP\*TP\*TP\*GP\*AP\*GP\*GP\*AP\*TP\*AP\*TP\*CP\*CP\*GP\*A)-3')

Chain E:  44% 38% 19%



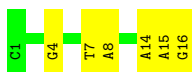
- Molecule 2: DNA (5'-D(P\*GP\*CP\*TP\*TP\*GP\*AP\*GP\*GP\*AP\*TP\*AP\*TP\*CP\*CP\*GP\*A)-3')

Chain G:  50% 50%



- Molecule 3: DNA (5'-D(P\*CP\*TP\*CP\*GP\*GP\*AP\*TP\*AP\*TP\*CP\*CP\*TP\*CP\*AP\*AP\*G)-3')

Chain F:  62% 38%



- Molecule 3: DNA (5'-D(P\*CP\*TP\*CP\*GP\*GP\*AP\*TP\*AP\*TP\*CP\*CP\*TP\*CP\*AP\*AP\*G)-3')

Chain H:  94% 6%





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	57.41Å 100.82Å 79.35Å 90.00° 102.46° 90.00°	Depositor
Resolution (Å)	49.00 – 2.70 49.00 – 2.70	Depositor EDS
% Data completeness (in resolution range)	94.8 (49.00-2.70) 84.3 (49.00-2.70)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.15 (at 2.69Å)	Xtrriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, $R_{free}$	0.213 , 0.258 0.213 , 0.257	Depositor DCC
$R_{free}$ test set	1173 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	53.8	Xtrriage
Anisotropy	0.876	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 41.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3289	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	79.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.54% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	0.59	0/498	0.90	3/670 (0.4%)
1	B	0.43	0/498	0.61	0/670
1	C	0.48	0/498	0.70	0/670
1	D	0.60	0/521	0.94	4/701 (0.6%)
2	E	1.08	2/371 (0.5%)	1.04	1/571 (0.2%)
2	G	0.94	0/371	0.98	0/571
3	F	1.10	0/363	1.06	0/557
3	H	0.85	0/360	0.99	1/553 (0.2%)
All	All	0.76	2/3480 (0.1%)	0.91	9/4963 (0.2%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	11	DA	C3'-O3'	-5.87	1.36	1.44
2	E	5	DG	C3'-O3'	-5.49	1.36	1.44

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	279	LYS	CD-CE-NZ	-10.03	88.64	111.70
1	D	280	LEU	CB-CG-CD2	-6.77	99.49	111.00
1	A	279	LYS	N-CA-CB	-5.79	100.18	110.60
1	D	279	LYS	N-CA-CB	-5.75	100.24	110.60
2	E	4	DT	O4'-C4'-C3'	-5.58	102.27	104.50
1	A	261	ASP	CB-CG-OD1	-5.56	113.30	118.30
1	D	279	LYS	CB-CA-C	5.34	121.07	110.40
3	H	7	DT	C1'-O4'-C4'	-5.17	104.93	110.10
1	A	250	ARG	CA-CB-CG	5.01	124.42	113.40

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts [i](#)

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	487	0	505	11	1
1	B	487	0	505	12	0
1	C	487	0	505	10	0
1	D	510	0	527	16	1
2	E	331	0	181	9	0
2	G	331	0	181	8	0
3	F	325	0	181	7	0
3	H	322	0	182	0	0
4	A	2	0	0	0	0
4	B	1	0	0	0	0
4	D	1	0	0	0	0
4	E	1	0	0	0	0
4	F	1	0	0	0	0
4	H	3	0	0	0	0
All	All	3289	0	2767	64	1

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 11.

All (64) 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:275:LYS:O	1:A:279:LYS:HB2	1.63	0.97
1:A:249:MET:HG2	1:A:250:ARG:H	1.30	0.95
1:C:286:LEU:HD21	1:C:291:VAL:HG23	1.54	0.88
1:A:249:MET:HG2	1:A:250:ARG:N	1.97	0.79
1:B:278:LEU:HD13	1:B:286:LEU:HD22	1.64	0.79
1:D:296:GLN:HE22	2:G:9:DA:H62	1.32	0.76
1:A:302:ARG:HA	1:B:303:TYR:HB2	1.69	0.74
2:E:1:DG:H2''	2:E:2:DC:H5''	1.70	0.73
1:B:292:LYS:NZ	3:F:4:DG:O6	2.25	0.68
2:E:3:DT:H2'	2:E:4:DT:H71	1.78	0.66
1:D:292:LYS:NZ	2:G:7:DG:N7	2.45	0.65
1:D:256:HIS:HD2	1:D:294:HIS:CE1	2.15	0.64
1:D:266:LEU:HD21	1:D:280:LEU:HD12	1.79	0.64

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:252:THR:HG23	1:B:255:LEU:H	1.63	0.64
1:C:278:LEU:HD13	1:C:286:LEU:HD22	1.79	0.64
1:C:252:THR:HG22	1:C:255:LEU:HG	1.80	0.63
1:D:275:LYS:O	1:D:279:LYS:HB2	1.99	0.62
2:E:4:DT:H2''	2:E:5:DG:H5'	1.81	0.62
3:F:15:DA:H2'	3:F:16:DG:C8	2.36	0.59
1:D:254:GLU:OE1	1:D:254:GLU:N	2.33	0.57
1:D:277:VAL:HG12	1:D:281:MET:HE2	1.86	0.57
1:D:296:GLN:HE22	2:G:9:DA:N6	2.00	0.57
1:A:273:THR:OG1	2:E:6:DA:OP1	2.23	0.57
1:B:297:LYS:O	1:B:300:THR:HG22	2.04	0.57
2:E:3:DT:H2''	2:E:4:DT:O5'	2.07	0.55
2:E:4:DT:H2'	2:E:5:DG:C8	2.42	0.55
1:D:251:TRP:CD2	1:D:294:HIS:ND1	2.75	0.54
1:C:252:THR:CG2	1:C:255:LEU:HG	2.38	0.54
3:F:15:DA:H2''	3:F:16:DG:H5'	1.90	0.53
2:G:4:DT:H2''	2:G:5:DG:O5'	2.09	0.53
1:A:301:ALA:HA	1:A:304:ARG:HH21	1.74	0.53
1:B:295:LEU:O	1:B:299:ARG:HG3	2.09	0.53
1:C:295:LEU:HD21	1:C:299:ARG:NH2	2.24	0.52
1:D:252:THR:HG22	1:D:254:GLU:OE1	2.09	0.52
1:D:266:LEU:CD2	1:D:280:LEU:HD12	2.39	0.52
1:B:295:LEU:HD21	1:B:299:ARG:HH21	1.75	0.52
1:C:287:THR:HG22	1:C:290:HIS:CE1	2.46	0.50
2:G:1:DG:H2''	2:G:2:DC:H5'	1.93	0.50
1:B:249:MET:HG2	1:B:250:ARG:N	2.27	0.50
1:A:273:THR:HG23	1:A:276:GLY:H	1.76	0.49
3:F:15:DA:H2''	3:F:16:DG:C5'	2.42	0.49
1:A:296:GLN:HE22	2:E:9:DA:H62	1.61	0.48
1:B:278:LEU:CD1	1:B:286:LEU:HD22	2.40	0.47
1:D:256:HIS:HD2	1:D:294:HIS:NE2	2.12	0.47
1:D:277:VAL:HG12	1:D:281:MET:CE	2.45	0.47
1:D:296:GLN:NE2	2:G:9:DA:H62	2.07	0.47
3:F:15:DA:H2'	3:F:16:DG:H8	1.79	0.47
1:B:257:GLU:HA	1:B:260:VAL:HG12	1.96	0.46
1:A:273:THR:HG23	1:A:276:GLY:HA3	1.97	0.46
1:B:252:THR:CG2	1:B:255:LEU:HD13	2.45	0.46
2:G:1:DG:H2'	2:G:2:DC:C6	2.49	0.46
2:E:10:DT:H2''	2:E:11:DA:O5'	2.16	0.46
1:C:287:THR:HG22	1:C:290:HIS:ND1	2.29	0.46
3:F:14:DA:H2'	3:F:15:DA:C8	2.51	0.46

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:F:7:DT:H2''	3:F:8:DA:O5'	2.16	0.45
1:C:257:GLU:HA	1:C:260:VAL:HG12	1.99	0.45
1:A:275:LYS:HE3	2:E:5:DG:H5''	1.99	0.44
1:A:304:ARG:HG2	1:A:305:PRO:HD3	2.00	0.44
1:B:286:LEU:HD21	1:B:291:VAL:HG23	2.00	0.44
1:C:277:VAL:O	1:C:281:MET:HB2	2.19	0.43
1:D:284:ASP:N	1:D:284:ASP:OD1	2.51	0.43
1:C:259:PHE:O	1:C:263:VAL:HG23	2.19	0.42
2:G:14:DC:H2''	2:G:15:DG:C8	2.55	0.42
1:D:303:TYR:O	1:D:306:GLU:HB2	2.21	0.41

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:279:LYS:O	1:D:279:LYS:NZ[2_646]	1.95	0.25

## 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	57/144 (40%)	52 (91%)	5 (9%)	0	100	100
1	B	57/144 (40%)	57 (100%)	0	0	100	100
1	C	57/144 (40%)	56 (98%)	1 (2%)	0	100	100
1	D	60/144 (42%)	54 (90%)	6 (10%)	0	100	100
All	All	231/576 (40%)	219 (95%)	12 (5%)	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	52/124 (42%)	52 (100%)	0	100	100
1	B	52/124 (42%)	52 (100%)	0	100	100
1	C	52/124 (42%)	52 (100%)	0	100	100
1	D	55/124 (44%)	55 (100%)	0	100	100
All	All	211/496 (42%)	211 (100%)	0	100	100

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

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	296	GLN
1	D	256	HIS
1	D	296	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 monosaccharides 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, 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	59/144 (40%)	0.51	4 (6%) 17 15	57, 73, 109, 118	0
1	B	59/144 (40%)	0.25	3 (5%) 28 26	52, 72, 127, 158	0
1	C	59/144 (40%)	0.11	0 100 100	59, 90, 122, 133	0
1	D	62/144 (43%)	0.32	0 100 100	58, 72, 96, 102	0
2	E	16/16 (100%)	-0.35	0 100 100	59, 67, 96, 98	0
2	G	16/16 (100%)	-0.56	0 100 100	69, 81, 96, 102	0
3	F	16/16 (100%)	-0.37	0 100 100	56, 70, 91, 95	0
3	H	16/16 (100%)	-0.47	0 100 100	67, 83, 94, 99	0
All	All	303/640 (47%)	0.14	7 (2%) 60 62	52, 76, 110, 158	0

All (7) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	303	TYR	5.5
1	B	304	ARG	4.1
1	B	305	PRO	3.8
1	A	303	TYR	3.1
1	A	305	PRO	2.4
1	A	279	LYS	2.3
1	A	301	ALA	2.1

### 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

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.