



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 05:27 AM GMT

PDB ID : 2QSH  
Title : Crystal structure of Rad4-Rad23 bound to a mismatch DNA  
Authors : Min, J.-H.; Pavletich, N.P.  
Deposited on : 2007-07-31  
Resolution : 2.81 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  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

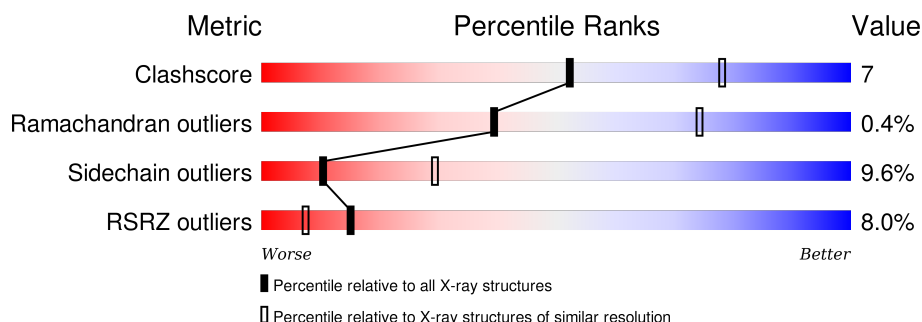
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.81 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))
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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. 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	W	24	
2	Y	24	
3	A	538	
4	X	171	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5485 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 DNA chain called top strand of the mismatch DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	W	24	Total	C	N	O	P	0	0	0
			481	233	82	143	23			

- Molecule 2 is a DNA chain called bottom strand of the mismatch DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	Y	23	Total	C	N	O	P	0	0	0
			456	219	82	134	21			

- Molecule 3 is a protein called DNA repair protein RAD4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	502	Total	C	N	O	S	0	0	0
			4132	2633	744	728	27			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	95	GLY	-	EXPRESSION TAG	UNP P14736
A	96	SER	-	EXPRESSION TAG	UNP P14736
A	97	SER	-	EXPRESSION TAG	UNP P14736
A	98	ARG	-	EXPRESSION TAG	UNP P14736
A	99	ALA	-	EXPRESSION TAG	UNP P14736
A	100	MET	-	EXPRESSION TAG	UNP P14736
A	223	GLU	VAL	SEE REMARK 999	UNP P14736
A	225	LEU	ILE	SEE REMARK 999	UNP P14736

- Molecule 4 is a protein called UV excision repair protein RAD23.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	X	54	Total	C	N	O	S	0	0	0
			416	265	70	79	2			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
X	228	GLY	-	EXPRESSION TAG	UNP P32628
X	229	SER	-	EXPRESSION TAG	UNP P32628

### 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: top strand of the mismatch DNA

Chain W: 




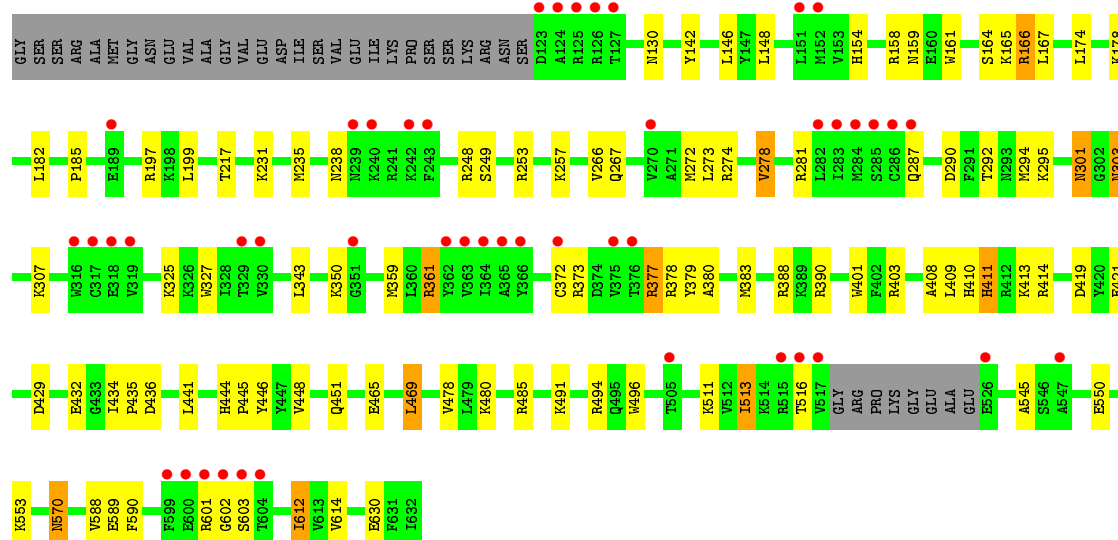
- Molecule 2: bottom strand of the mismatch DNA

Chain Y: 



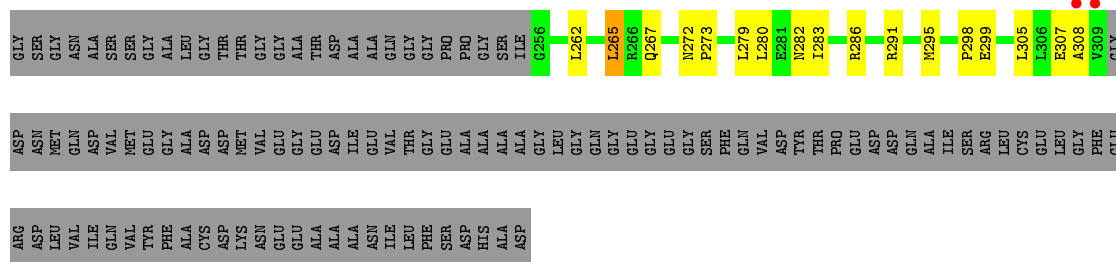
- Molecule 3: DNA repair protein RAD4

Chain A: 



- Molecule 4: UV excision repair protein RAD23

Chain X: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.64Å 79.64Å 403.97Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.81 29.79 – 2.80	Depositor EDS
% Data completeness (in resolution range)	91.3 (30.00-2.81) 90.4 (29.79-2.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.05 (at 2.80Å)	Xtriage
Refinement program	REFMAC 5.2	Depositor
R, $R_{free}$	0.209 , 0.244 0.256 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	61.7	Xtriage
Anisotropy	0.182	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 25.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 31753 reflections	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	5485	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.53% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	W	0.71	0/537	1.44	10/825 (1.2%)
2	Y	0.69	0/510	1.54	11/786 (1.4%)
3	A	0.43	0/4221	0.56	0/5673
4	X	0.40	0/422	0.54	0/575
All	All	0.49	0/5690	0.84	21/7859 (0.3%)

There are no bond length outliers.

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	Y	21	DG	O4'-C1'-N9	8.21	113.75	108.00
2	Y	12	DG	O4'-C1'-N9	7.67	113.37	108.00
2	Y	6	DA	O4'-C1'-N9	-7.00	103.10	108.00
2	Y	11	DT	O4'-C4'-C3'	-6.93	101.73	104.50
1	W	18	DG	P-O3'-C3'	6.53	127.53	119.70
1	W	22	DC	C1'-O4'-C4'	-6.31	103.79	110.10
2	Y	18	DT	O4'-C4'-C3'	-6.22	102.01	104.50
1	W	22	DC	O4'-C1'-N1	6.11	112.28	108.00
2	Y	11	DT	O4'-C1'-N1	6.11	112.28	108.00
2	Y	22	DT	O4'-C1'-N1	6.04	112.23	108.00
1	W	9	DA	O4'-C1'-C2'	-5.85	101.22	105.90
2	Y	13	DG	P-O3'-C3'	5.74	126.58	119.70
1	W	9	DA	C1'-O4'-C4'	-5.57	104.53	110.10
1	W	10	DC	O4'-C1'-N1	5.48	111.84	108.00
2	Y	18	DT	C1'-O4'-C4'	-5.46	104.64	110.10
1	W	3	DG	O4'-C1'-N9	5.46	111.82	108.00
2	Y	11	DT	C4'-C3'-C2'	-5.43	98.21	103.10
2	Y	14	DA	P-O3'-C3'	5.33	126.10	119.70
1	W	8	DA	O4'-C1'-N9	5.30	111.71	108.00
1	W	9	DA	O4'-C1'-N9	5.22	111.65	108.00
1	W	2	DT	C4-C5-C7	5.18	122.11	119.00

There are no chirality outliers.



There are no planarity outliers.

## 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	W	481	0	274	3	0
2	Y	456	0	253	6	0
3	A	4132	0	4214	59	0
4	X	416	0	427	9	0
All	All	5485	0	5168	71	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 7.

All (71) 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)
3:A:372:CYS:H	3:A:410:HIS:HD2	1.07	0.99
3:A:159:ASN:HD21	3:A:274:ARG:HH22	1.26	0.83
3:A:287:GLN:HE22	3:A:359:MET:H	1.30	0.80
3:A:444:HIS:HD2	3:A:446:TYR:H	1.30	0.77
1:W:9:DA:H2"	1:W:10:DC:H5"	1.67	0.75
2:Y:12:DG:H2"	2:Y:13:DG:C8	2.20	0.75
3:A:465:GLU:HG2	3:A:480:LYS:HD2	1.68	0.75
3:A:380:ALA:HB3	3:A:383:MET:HG2	1.69	0.74
3:A:359:MET:O	3:A:361:ARG:HD3	1.88	0.73
4:X:265:LEU:HD13	4:X:305:LEU:HD21	1.72	0.70
3:A:372:CYS:H	3:A:410:HIS:CD2	2.00	0.68
3:A:287:GLN:NE2	3:A:359:MET:H	1.93	0.66
3:A:545:ALA:H	3:A:570:ASN:HD21	1.45	0.65
3:A:377:ARG:HD3	3:A:377:ARG:O	2.00	0.62
4:X:267:GLN:HE21	4:X:272:ASN:ND2	1.98	0.61
3:A:292:THR:HG21	3:A:390:ARG:NH1	2.15	0.60
3:A:273:LEU:HB3	3:A:278:VAL:HG22	1.83	0.60
4:X:267:GLN:HE21	4:X:272:ASN:HD22	1.51	0.59
3:A:444:HIS:CD2	3:A:446:TYR:H	2.17	0.59
3:A:159:ASN:ND2	3:A:274:ARG:HH22	1.98	0.58

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:590:PHE:HD1	3:A:612:ILE:HD12	1.67	0.58
3:A:373:ARG:NH1	3:A:421:GLU:OE1	2.34	0.58
3:A:174:LEU:CB	3:A:272:MET:HE2	2.36	0.56
3:A:550:GLU:HA	3:A:589:GLU:HG2	1.87	0.56
3:A:414:ARG:NH1	3:A:419:ASP:OD1	2.39	0.55
2:Y:13:DG:H2"	2:Y:14:DA:OP2	2.06	0.55
3:A:465:GLU:CG	3:A:480:LYS:HD2	2.36	0.54
3:A:377:ARG:NH2	3:A:432:GLU:OE2	2.37	0.53
3:A:401:TRP:NE1	4:X:298:PRO:HG3	2.23	0.53
2:Y:14:DA:H1'	2:Y:15:DT:H5'	1.90	0.53
3:A:401:TRP:CD1	4:X:298:PRO:HG3	2.44	0.53
3:A:281:ARG:HD2	3:A:327:TRP:CZ3	2.44	0.53
3:A:511:LYS:HG2	3:A:513:ILE:CD1	2.39	0.53
2:Y:11:DT:H5"	3:A:602:GLY:H	1.75	0.52
3:A:142:TYR:CE1	4:X:291:ARG:HA	2.44	0.52
2:Y:11:DT:H5"	3:A:602:GLY:N	2.25	0.51
3:A:161:TRP:CE2	3:A:248:ARG:HG3	2.45	0.50
3:A:154:HIS:CE1	4:X:273:PRO:HG3	2.47	0.50
1:W:10:DC:H2"	1:W:11:DA:H8	1.76	0.49
3:A:444:HIS:CD2	3:A:445:PRO:HD2	2.46	0.49
3:A:590:PHE:CD1	3:A:612:ILE:HD12	2.46	0.49
3:A:174:LEU:HB2	3:A:272:MET:HE2	1.94	0.48
3:A:301:ASN:HD22	3:A:303:ASN:H	1.61	0.48
3:A:253:ARG:HG2	3:A:257:LYS:HE2	1.95	0.48
3:A:408:ALA:O	3:A:411:HIS:HE1	1.96	0.47
3:A:166:ARG:CZ	3:A:166:ARG:H	2.27	0.47
3:A:164:SER:OG	3:A:166:ARG:NH1	2.49	0.46
3:A:130:ASN:O	3:A:295:LYS:HA	2.15	0.46
3:A:377:ARG:HA	3:A:383:MET:HG3	1.97	0.46
4:X:280:LEU:HA	4:X:283:ILE:HG22	1.97	0.46
3:A:273:LEU:HB3	3:A:278:VAL:CG2	2.45	0.45
3:A:174:LEU:HB3	3:A:272:MET:HE2	1.97	0.45
3:A:436:ASP:C	3:A:469:LEU:HD21	2.37	0.45
4:X:282:ASN:HD21	4:X:286:ARG:HH21	1.66	0.44
1:W:10:DC:H2"	1:W:11:DA:C8	2.52	0.44
3:A:588:VAL:HG22	3:A:589:GLU:H	1.82	0.43
3:A:166:ARG:NH1	3:A:166:ARG:HB2	2.34	0.43
3:A:287:GLN:HE22	3:A:359:MET:N	2.07	0.43
3:A:491:LYS:HB2	3:A:496:TRP:CE2	2.53	0.43
3:A:465:GLU:HG2	3:A:480:LYS:HB3	2.01	0.42
3:A:570:ASN:HD22	3:A:570:ASN:C	2.23	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:434:ILE:HA	3:A:435:PRO:HD3	1.94	0.42
3:A:378:ARG:HD2	3:A:379:TYR:CZ	2.55	0.41
3:A:290:ASP:OD1	3:A:292:THR:HB	2.20	0.41
3:A:292:THR:HG21	3:A:390:ARG:CZ	2.50	0.41
3:A:182:LEU:O	3:A:185:PRO:HD3	2.21	0.41
2:Y:12:DG:H2"	2:Y:13:DG:N7	2.34	0.41
3:A:161:TRP:CD2	3:A:248:ARG:HG3	2.56	0.41
3:A:612:ILE:CG2	3:A:614:VAL:HG13	2.51	0.40
3:A:158:ARG:HH11	3:A:267:GLN:HE22	1.69	0.40
3:A:235:MET:O	3:A:238:ASN:HB3	2.22	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	
3	A	498/538 (93%)	474 (95%)	23 (5%)	1 (0%)	52	84
4	X	52/171 (30%)	50 (96%)	1 (2%)	1 (2%)	10	32
All	All	550/709 (78%)	524 (95%)	24 (4%)	2 (0%)	39	74

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	603	SER
4	X	308	ALA

### 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	
3	A	452/480 (94%)	410 (91%)	42 (9%)	11	32
4	X	47/129 (36%)	41 (87%)	6 (13%)	5	16
All	All	499/609 (82%)	451 (90%)	48 (10%)	10	29

All (48) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	146	LEU
3	A	148	LEU
3	A	165	LYS
3	A	166	ARG
3	A	167	LEU
3	A	178	LYS
3	A	197	ARG
3	A	199	LEU
3	A	217	THR
3	A	231	LYS
3	A	249	SER
3	A	266	VAL
3	A	278	VAL
3	A	294	MET
3	A	301	ASN
3	A	303	ASN
3	A	307	LYS
3	A	325	LYS
3	A	343	LEU
3	A	350	LYS
3	A	361	ARG
3	A	377	ARG
3	A	388	ARG
3	A	403	ARG
3	A	409	LEU
3	A	411	HIS
3	A	413	LYS
3	A	429	ASP
3	A	441	LEU
3	A	448	VAL
3	A	451	GLN
3	A	469	LEU
3	A	478	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	A	485	ARG
3	A	494	ARG
3	A	513	ILE
3	A	516	THR
3	A	553	LYS
3	A	570	ASN
3	A	601	ARG
3	A	612	ILE
3	A	630	GLU
4	X	262	LEU
4	X	265	LEU
4	X	279	LEU
4	X	295	MET
4	X	299	GLU
4	X	307	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	154	HIS
3	A	159	ASN
3	A	173	ASN
3	A	184	HIS
3	A	267	GLN
3	A	287	GLN
3	A	301	ASN
3	A	303	ASN
3	A	410	HIS
3	A	411	HIS
3	A	427	GLN
3	A	444	HIS
3	A	570	ASN
4	X	272	ASN

### 5.3.3 RNA ⓘ

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 ⓘ

### 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	W	24/24 (100%)	0.43	0	100 100	44, 61, 65, 79	0
2	Y	23/24 (95%)	0.34	0	100 100	42, 57, 94, 101	0
3	A	502/538 (93%)	0.40	46 (9%)	11 5	44, 56, 77, 101	0
4	X	54/171 (31%)	0.05	2 (3%)	45 33	43, 58, 68, 87	0
All	All	603/757 (79%)	0.37	48 (7%)	15 7	42, 57, 78, 101	0

All (48) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	517	VAL	5.5
3	A	603	SER	5.4
4	X	309	VAL	5.2
3	A	317	CYS	5.0
3	A	516	THR	4.4
3	A	526	GLU	4.3
3	A	601	ARG	4.0
3	A	125	ARG	3.6
3	A	604	THR	3.5
3	A	286	CYS	3.4
3	A	602	GLY	3.4
3	A	600	GLU	3.4
3	A	123	ASP	3.4
3	A	283	ILE	3.2
3	A	284	MET	3.1
3	A	127	THR	3.0
3	A	126	ARG	3.0
3	A	285	SER	3.0
3	A	329	THR	2.9
4	X	308	ALA	2.8
3	A	599	PHE	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	A	316	TRP	2.6
3	A	515	ARG	2.6
3	A	318	GLU	2.6
3	A	282	LEU	2.5
3	A	365	ALA	2.5
3	A	124	ALA	2.5
3	A	330	VAL	2.4
3	A	375	VAL	2.4
3	A	505	THR	2.4
3	A	364	ILE	2.4
3	A	547	ALA	2.3
3	A	363	VAL	2.3
3	A	366	TYR	2.3
3	A	242	LYS	2.3
3	A	240	LYS	2.3
3	A	151	LEU	2.3
3	A	239	ASN	2.2
3	A	319	VAL	2.2
3	A	351	GLY	2.2
3	A	287	GLN	2.2
3	A	152	MET	2.2
3	A	270	VAL	2.2
3	A	362	TYR	2.2
3	A	243	PHE	2.1
3	A	372	CYS	2.1
3	A	376	THR	2.1
3	A	189	GLU	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 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.