



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 21, 2020 – 10:06 AM BST

PDB ID : 6ZRC  
Title : Structure of the human RBAP48 in complex with a macrocyclic peptide cyclized via a xylene linker attached to two cysteines  
Authors : Vetter, I.R.; Porfetye, A.T.  
Deposited on : 2020-07-13  
Resolution : 2.60 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.14.6  
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) : 2.14.6

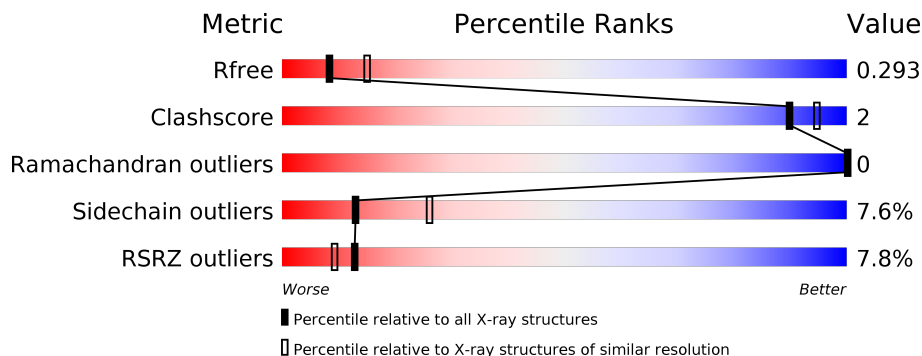
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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 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	427	<div style="display: flex; align-items: center;"> <div style="width: 6%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 78%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 10%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 11%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">6%      78%      10% • 11%</p>
1	B	427	<div style="display: flex; align-items: center;"> <div style="width: 8%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 79%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 10%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 11%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">8%      79%      10% • 11%</p>
2	P	16	<div style="display: flex; align-items: center;"> <div style="width: 88%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 13%; height: 10px; background-color: yellow;"></div> </div> <p style="text-align: center;">88%      13%</p>
2	Q	16	<div style="display: flex; align-items: center;"> <div style="width: 13%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 88%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 13%; height: 10px; background-color: yellow;"></div> </div> <p style="text-align: center;">13%      88%      13%</p>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 6335 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 Histone-binding protein RBBP4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	381	3042	1917	519	596	10	0	0	0
1	B	381	3046	1919	519	598	10	0	0	0

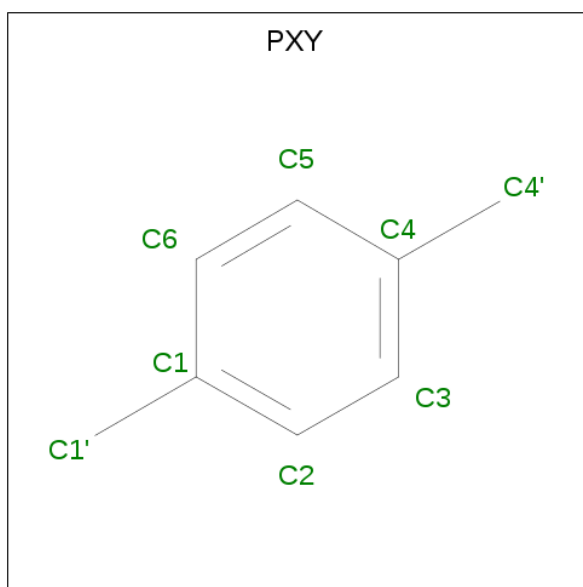
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP Q09028
A	0	PRO	-	expression tag	UNP Q09028
B	-1	GLY	-	expression tag	UNP Q09028
B	0	PRO	-	expression tag	UNP Q09028

- Molecule 2 is a protein called macrocyclic peptide based on residues 659-672 of the metastasis-associated protein MTA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	P	16	115	70	26	17	2	0	0	1
2	Q	16	115	70	26	17	2	0	0	1

- Molecule 3 is PARA-XYLENE (three-letter code: PXY) (formula: C<sub>8</sub>H<sub>10</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	P	1	Total C 8 8	0	0
3	Q	1	Total C 8 8	0	0

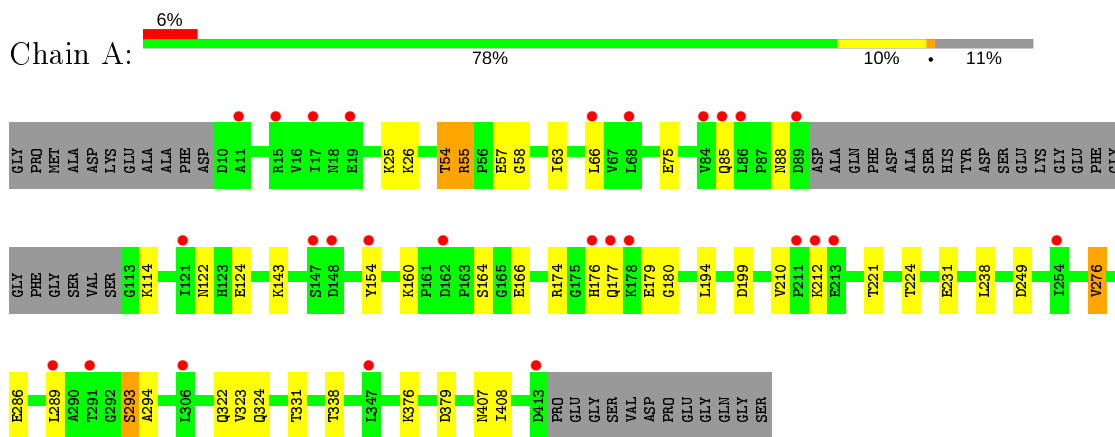
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total O 1 1	0	0

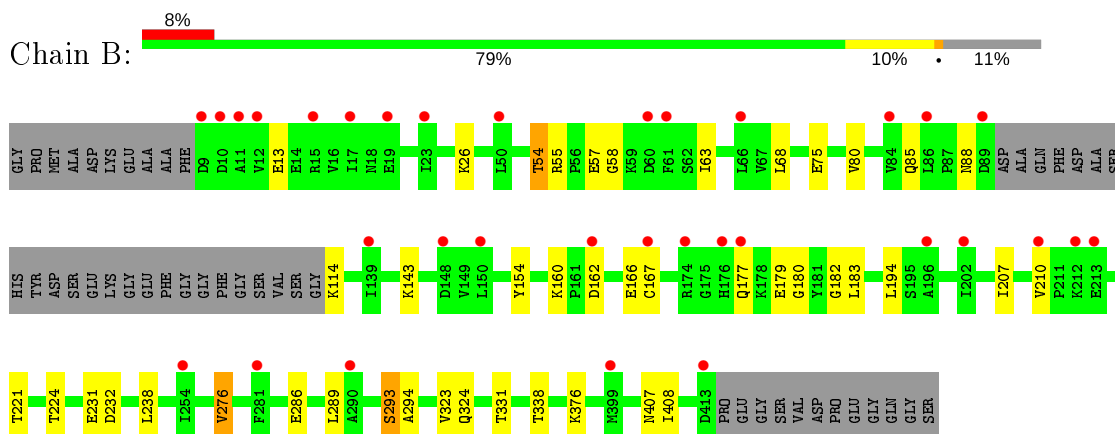
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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.

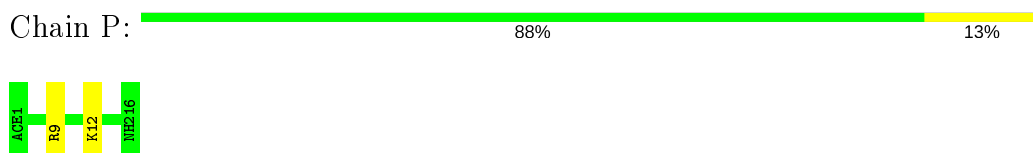
- Molecule 1: Histone-binding protein RBBP4



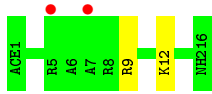
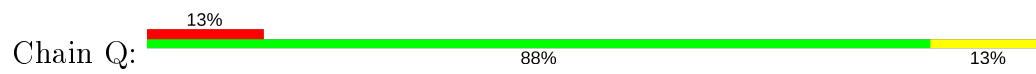
- Molecule 1: Histone-binding protein RBBP4



- Molecule 2: macrocyclic peptide based on residues 659-672 of the metastasis-associated protein MTA1



- Molecule 2: macrocyclic peptide based on residues 659-672 of the metastasis-associated protein MTA1



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	71.96Å 96.33Å 149.30Å 90.00° 90.01° 90.00°	Depositor
Resolution (Å)	48.16 – 2.60 48.16 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.0 (48.16-2.60) 99.5 (48.16-2.60)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.46 (at 2.61Å)	Xtrriage
Refinement program	PHENIX 1.18rc7_3834	Depositor
R, $R_{free}$	0.263 , 0.293 0.263 , 0.293	Depositor DCC
$R_{free}$ test set	1567 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	64.0	Xtrriage
Anisotropy	0.930	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 66.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.458 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	6335	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	115.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.43% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PXY, ACE, NH2

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	0.56	0/3125	0.80	4/4261 (0.1%)
1	B	0.58	0/3129	0.79	3/4267 (0.1%)
2	P	0.71	0/114	0.83	0/152
2	Q	0.65	0/114	0.87	0/152
All	All	0.57	0/6482	0.79	7/8832 (0.1%)

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
1	B	0	1
All	All	0	2

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	58	GLY	C-N-CA	6.91	138.98	121.70
1	B	58	GLY	C-N-CA	6.58	138.14	121.70
1	A	194	LEU	CB-CG-CD1	-5.70	101.31	111.00
1	A	249	ASP	CB-CG-OD1	5.43	123.19	118.30
1	B	194	LEU	CB-CG-CD1	-5.26	102.05	111.00
1	B	154	TYR	CA-CB-CG	5.24	123.35	113.40
1	A	154	TYR	CA-CB-CG	5.01	122.91	113.40

There are no chirality outliers.

All (2) planarity outliers are listed below:



Mol	Chain	Res	Type	Group
1	A	55	ARG	Peptide
1	B	162	ASP	Peptide

## 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	3042	0	2895	12	0
1	B	3046	0	2896	10	1
2	P	115	0	121	0	0
2	Q	115	0	121	0	0
3	P	8	0	7	0	0
3	Q	8	0	7	0	0
4	B	1	0	0	0	0
All	All	6335	0	6047	22	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 2.

All (22) 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:B:54:THR:HG23	1:B:63:ILE:HB	1.93	0.50
1:A:276:VAL:HA	1:A:293:SER:HB2	1.94	0.50
1:B:180:GLY:HA3	1:B:231:GLU:HA	1.95	0.49
1:B:276:VAL:HA	1:B:293:SER:HB2	1.95	0.49
1:A:176:HIS:O	1:A:177:GLN:NE2	2.45	0.49
1:A:338:THR:HG23	1:A:376:LYS:HD2	1.97	0.47
1:B:338:THR:HG23	1:B:376:LYS:HD2	1.97	0.47
1:B:323:VAL:O	1:B:324:GLN:NE2	2.48	0.46
1:A:122:ASN:N	1:A:122:ASN:OD1	2.49	0.45
1:A:293:SER:OG	1:A:294:ALA:N	2.49	0.45
1:A:54:THR:HG23	1:A:63:ILE:HB	1.99	0.44
1:A:63:ILE:HD13	1:A:85:GLN:HG2	2.00	0.43
1:B:63:ILE:HD13	1:B:85:GLN:HG2	2.00	0.43
1:A:238:LEU:HD12	1:A:286:GLU:HB3	2.00	0.43
1:B:68:LEU:HB2	1:B:80:VAL:HB	2.00	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:199:ASP:N	1:A:199:ASP:OD1	2.50	0.42
1:B:238:LEU:HD12	1:B:286:GLU:HB3	2.00	0.42
1:A:180:GLY:HA3	1:A:231:GLU:HA	2.01	0.42
1:A:322:GLN:NE2	1:A:379:ASP:OD1	2.48	0.42
1:B:293:SER:OG	1:B:294:ALA:N	2.52	0.41
1:B:182:GLY:HA3	1:B:232:ASP:HA	2.03	0.40
1:A:323:VAL:O	1:A:324:GLN:NE2	2.54	0.40

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:B:167:CYS:CB	1:B:167:CYS:SG 2_556	2.18	0.02

## 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	377/427 (88%)	363 (96%)	14 (4%)	0	100	100
1	B	377/427 (88%)	363 (96%)	14 (4%)	0	100	100
2	P	14/16 (88%)	13 (93%)	1 (7%)	0	100	100
2	Q	14/16 (88%)	14 (100%)	0	0	100	100
All	All	782/886 (88%)	753 (96%)	29 (4%)	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	343/376 (91%)	317 (92%)	26 (8%)	13	26
1	B	344/376 (92%)	320 (93%)	24 (7%)	15	30
2	P	11/11 (100%)	9 (82%)	2 (18%)	1	2
2	Q	11/11 (100%)	9 (82%)	2 (18%)	1	2
All	All	709/774 (92%)	655 (92%)	54 (8%)	13	26

All (54) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	25	LYS
1	A	26	LYS
1	A	54	THR
1	A	55	ARG
1	A	57	GLU
1	A	66	LEU
1	A	75	GLU
1	A	88	ASN
1	A	114	LYS
1	A	124	GLU
1	A	143	LYS
1	A	160	LYS
1	A	164	SER
1	A	166	GLU
1	A	174	ARG
1	A	179	GLU
1	A	210	VAL
1	A	212	LYS
1	A	221	THR
1	A	224	THR
1	A	276	VAL
1	A	289	LEU
1	A	293	SER
1	A	331	THR
1	A	407	ASN
1	A	408	ILE
2	P	9	ARG
2	P	12	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	Q	9	ARG
2	Q	12	LYS
1	B	13	GLU
1	B	26	LYS
1	B	54	THR
1	B	55	ARG
1	B	57	GLU
1	B	75	GLU
1	B	88	ASN
1	B	114	LYS
1	B	143	LYS
1	B	160	LYS
1	B	166	GLU
1	B	177	GLN
1	B	179	GLU
1	B	183	LEU
1	B	207	ILE
1	B	210	VAL
1	B	221	THR
1	B	224	THR
1	B	276	VAL
1	B	289	LEU
1	B	293	SER
1	B	331	THR
1	B	407	ASN
1	B	408	ILE

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

### 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](#)

2 ligands are modelled in this entry.

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	PXY	Q	101	2	8,8,8	0.69	0	10,10,10	1.15	0
3	PXY	P	101	2	8,8,8	0.50	0	10,10,10	1.26	0

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	PXY	Q	101	2	-	-	0/1/1/1
3	PXY	P	101	2	-	-	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	A	381/427 (89%)	0.53	27 (7%) <span style="border: 1px solid red; padding: 2px;">16</span> <span style="border: 1px solid red; padding: 2px;">11</span>	92, 114, 141, 175	0
1	B	381/427 (89%)	0.65	33 (8%) <span style="border: 1px solid red; padding: 2px;">10</span> <span style="border: 1px solid red; padding: 2px;">7</span>	92, 112, 143, 168	0
2	P	14/16 (87%)	0.50	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	109, 114, 119, 120	0
2	Q	14/16 (87%)	0.70	2 (14%) <span style="border: 1px solid red; padding: 2px;">2</span> <span style="border: 1px solid red; padding: 2px;">1</span>	110, 116, 120, 125	0
All	All	790/886 (89%)	0.59	62 (7%) <span style="border: 1px solid red; padding: 2px;">13</span> <span style="border: 1px solid red; padding: 2px;">9</span>	92, 114, 142, 175	0

All (62) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	413	ASP	6.3
1	A	11	ALA	5.6
1	A	413	ASP	5.5
1	A	213	GLU	5.4
1	B	9	ASP	5.3
1	B	176	HIS	4.5
1	A	15	ARG	4.5
1	A	177	GLN	4.4
1	A	178	LYS	4.3
1	B	60	ASP	4.2
1	B	12	VAL	4.2
1	B	212	LYS	4.1
1	B	162	ASP	3.9
1	B	89	ASP	3.7
1	A	17	ILE	3.6
1	A	162	ASP	3.6
1	A	254	ILE	3.4
1	A	291	THR	3.4
1	A	147	SER	3.2
1	A	289	LEU	3.2
1	A	19	GLU	3.2

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	A	212	LYS	3.1
1	B	15	ARG	3.0
1	B	23	ILE	3.0
1	B	281	PHE	2.9
1	B	19	GLU	2.9
1	B	86	LEU	2.9
1	A	176	HIS	2.8
1	B	167	CYS	2.8
1	B	254	ILE	2.8
1	B	177	GLN	2.7
1	B	202	ILE	2.7
1	A	68	LEU	2.7
1	A	306	LEU	2.6
1	A	347	LEU	2.6
1	B	17	ILE	2.6
1	A	89	ASP	2.5
1	B	11	ALA	2.5
1	B	148	ASP	2.5
1	B	50	LEU	2.4
1	B	174	ARG	2.4
2	Q	7	ALA	2.4
1	B	399	MET	2.3
1	A	148	ASP	2.3
1	A	86	LEU	2.2
1	A	121	ILE	2.2
1	B	150	LEU	2.2
1	A	84	VAL	2.2
1	B	84	VAL	2.1
2	Q	5	ARG	2.1
1	A	66	LEU	2.1
1	B	10	ASP	2.1
1	B	290	ALA	2.1
1	A	211	PRO	2.1
1	B	66	LEU	2.1
1	B	210	VAL	2.1
1	B	213	GLU	2.1
1	B	61	PHE	2.1
1	B	139	ILE	2.0
1	B	196	ALA	2.0
1	A	154	TYR	2.0
1	A	85	GLN	2.0

## 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 [i](#)

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. 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	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	PXY	P	101	8/8	0.89	0.29	118,121,125,125	0
3	PXY	Q	101	8/8	0.92	0.18	125,127,129,130	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.