



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 24, 2023 – 02:31 AM EDT

PDB ID : 5TH9  
Title : Structure determination of a potent, selective antibody inhibitor of human MMP9 (GS-5745 bound to MMP-9)  
Authors : Appleby, T.C.; Greenstein, A.E.; Kwon, H.J.  
Deposited on : 2016-09-29  
Resolution : 3.00 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35.1  
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) : 2.35.1

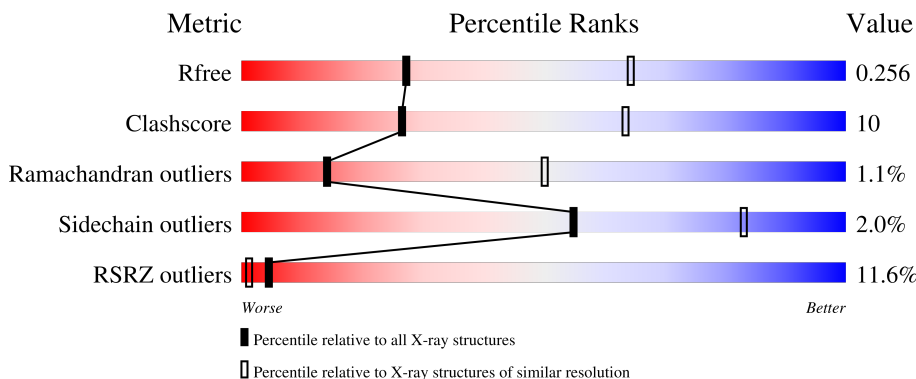
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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.





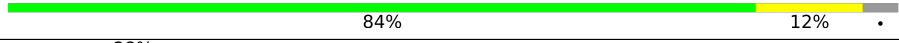

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 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	L	214	
1	M	214	
1	N	214	
2	H	230	
2	I	230	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	J	230	
3	A	231	
3	B	231	
3	C	231	

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 13578 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 GS-5745 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	L	214	Total	C	N	O	S	0	1	0
			1656	1032	281	337	6			
1	M	214	Total	C	N	O	S	0	0	0
			1646	1026	278	336	6			
1	N	106	Total	C	N	O	S	0	0	0
			810	507	136	164	3			

- Molecule 2 is a protein called GS-5745 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	H	208	Total	C	N	O	S	0	1	0
			1593	1017	258	311	7			
2	I	208	Total	C	N	O	S	0	1	0
			1593	1017	258	311	7			
2	J	106	Total	C	N	O	S	0	0	0
			839	539	135	160	5			

- Molecule 3 is a protein called Matrix metalloproteinase-9,Matrix metalloproteinase-9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	222	Total	C	N	O	S	0	0	0
			1767	1135	301	327	4			
3	B	222	Total	C	N	O	S	0	0	0
			1767	1135	301	327	4			
3	C	217	Total	C	N	O	S	0	0	0
			1734	1114	296	320	4			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	38	MET	-	initiating methionine	UNP P14780
A	39	ALA	-	expression tag	UNP P14780

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	38	MET	-	initiating methionine	UNP P14780
B	39	ALA	-	expression tag	UNP P14780
C	38	MET	-	initiating methionine	UNP P14780
C	39	ALA	-	expression tag	UNP P14780

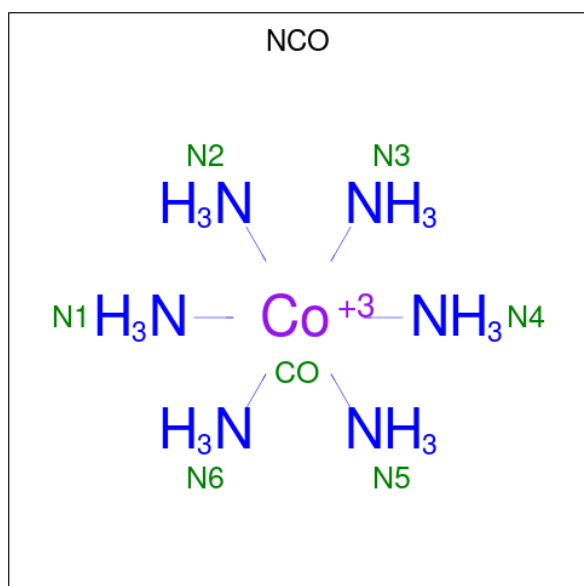
- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	2	Total Zn 2 2	0	0
4	B	2	Total Zn 2 2	0	0
4	C	2	Total Zn 2 2	0	0

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	2	Total Ca 2 2	0	0
5	B	2	Total Ca 2 2	0	0
5	C	2	Total Ca 2 2	0	0

- Molecule 6 is COBALT HEXAMMINE(III) (three-letter code: NCO) (formula:  $\text{CoH}_{18}\text{N}_6$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	Co	N	0	0
			7	1	6		
6	B	1	Total	Co	N	0	0
			7	1	6		
6	C	1	Total	Co	N	0	0
			7	1	6		

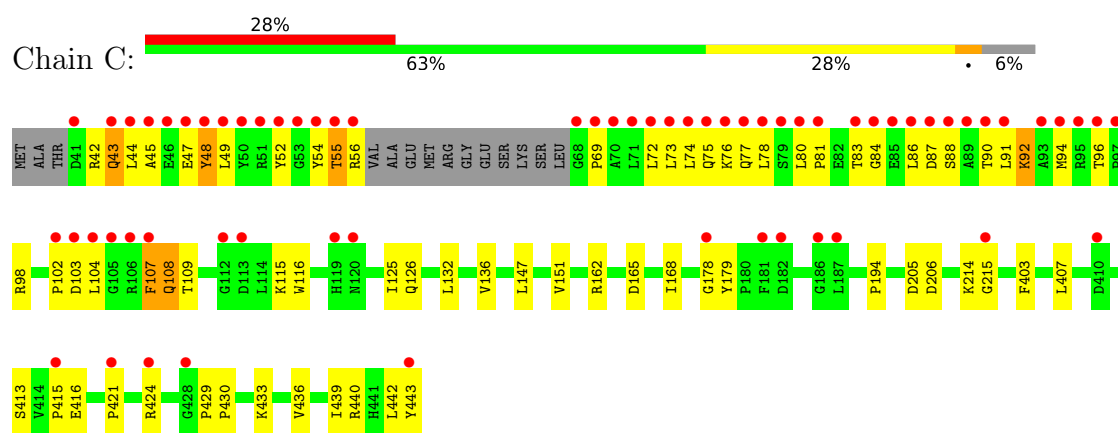
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	L	23	Total	O	0	0
			23	23		
7	H	26	Total	O	0	0
			26	26		
7	A	19	Total	O	0	0
			19	19		
7	M	27	Total	O	0	0
			27	27		
7	I	15	Total	O	0	0
			15	15		
7	B	30	Total	O	0	0
			30	30		









## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	168.17Å 168.17Å 234.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.64 – 3.00 46.64 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.8 (46.64-3.00) 94.5 (46.64-3.00)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.74 (at 3.01Å)	Xtrriage
Refinement program	PHENIX 1.10.1_2155	Depositor
R, $R_{free}$	0.218 , 0.257 0.219 , 0.256	Depositor DCC
$R_{free}$ test set	2000 reflections (2.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	36.2	Xtrriage
Anisotropy	0.013	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 27.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	13578	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.16% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, CA, NCO

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	L	0.67	0/1691	0.57	0/2295
1	M	0.69	0/1680	0.63	0/2280
1	N	0.47	0/827	0.75	1/1124 (0.1%)
2	H	0.71	0/1635	0.62	0/2233
2	I	0.74	1/1635 (0.1%)	0.63	1/2233 (0.0%)
2	J	0.78	7/862 (0.8%)	0.89	3/1172 (0.3%)
3	A	0.72	0/1822	0.61	0/2478
3	B	0.63	0/1822	0.59	1/2478 (0.0%)
3	C	0.33	0/1789	0.60	1/2432 (0.0%)
All	All	0.65	8/13763 (0.1%)	0.64	7/18725 (0.0%)

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
2	J	0	3
3	C	0	1
All	All	0	4

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	J	71	LYS	CE-NZ	6.96	1.66	1.49
2	J	66	ARG	NE-CZ	-6.67	1.24	1.33
2	J	66	ARG	CZ-NH1	-6.59	1.24	1.33
2	J	66	ARG	CZ-NH2	-6.39	1.24	1.33
2	J	71	LYS	CD-CE	-5.88	1.36	1.51

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	J	18	LEU	CB-CG-CD1	8.66	125.73	111.00
2	J	71	LYS	CD-CE-NZ	-7.75	93.86	111.70
3	C	86	LEU	CA-CB-CG	7.42	132.38	115.30
2	I	80	LEU	CB-CG-CD1	-6.78	99.48	111.00
2	J	11	LEU	CA-CB-CG	6.61	130.51	115.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	43	GLN	Peptide
2	J	47	TRP	Peptide
2	J	72	ASP	Peptide
2	J	81	LYS	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	L	1656	0	1605	6	0
1	M	1646	0	1599	12	0
1	N	810	0	787	34	0
2	H	1593	0	1563	20	0
2	I	1593	0	1563	21	0
2	J	839	0	813	73	0
3	A	1767	0	1672	31	0
3	B	1767	0	1672	20	0
3	C	1734	0	1635	75	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
4	C	2	0	0	0	0
5	A	2	0	0	0	0
5	B	2	0	0	0	0
5	C	2	0	0	0	0
6	A	7	0	0	0	0
6	B	7	0	0	0	0
6	C	7	0	0	1	0
7	A	19	0	0	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	B	30	0	0	2	0
7	H	26	0	0	0	0
7	I	15	0	0	0	0
7	L	23	0	0	0	0
7	M	27	0	0	0	0
All	All	13578	0	12909	270	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 10.

The worst 5 of 270 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)
2:J:14:PRO:C	2:J:16:GLU:OE1	1.79	1.20
3:C:49:LEU:HD21	3:C:55:THR:HG23	1.18	1.10
3:A:440:ARG:HD3	3:C:440:ARG:HH12	1.26	1.00
2:J:29:LEU:HG	2:J:71:LYS:HE2	1.43	0.99
3:C:75:GLN:HE22	3:C:83:THR:HG22	1.28	0.98

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
1	L	213/214 (100%)	200 (94%)	12 (6%)	1 (0%)	29 68
1	M	212/214 (99%)	200 (94%)	12 (6%)	0	100 100
1	N	104/214 (49%)	89 (86%)	9 (9%)	6 (6%)	1 10
2	H	205/230 (89%)	196 (96%)	9 (4%)	0	100 100
2	I	205/230 (89%)	199 (97%)	6 (3%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	J	104/230 (45%)	81 (78%)	17 (16%)	6 (6%)	1	10
3	A	218/231 (94%)	205 (94%)	11 (5%)	2 (1%)	17	55
3	B	218/231 (94%)	209 (96%)	8 (4%)	1 (0%)	29	68
3	C	213/231 (92%)	186 (87%)	25 (12%)	2 (1%)	17	55
All	All	1692/2025 (84%)	1565 (92%)	109 (6%)	18 (1%)	14	50

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	N	19	VAL
1	N	30	ARG
2	J	14	PRO
3	C	55	THR
1	L	127	SER

### 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	L	189/188 (100%)	181 (96%)	8 (4%)	30	66
1	M	188/188 (100%)	187 (100%)	1 (0%)	88	96
1	N	91/188 (48%)	87 (96%)	4 (4%)	28	65
2	H	182/201 (90%)	179 (98%)	3 (2%)	62	86
2	I	182/201 (90%)	180 (99%)	2 (1%)	73	90
2	J	92/201 (46%)	87 (95%)	5 (5%)	22	57
3	A	181/188 (96%)	178 (98%)	3 (2%)	60	85
3	B	181/188 (96%)	180 (99%)	1 (1%)	86	95
3	C	177/188 (94%)	174 (98%)	3 (2%)	60	85
All	All	1463/1731 (84%)	1433 (98%)	30 (2%)	55	82

5 of 30 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	M	65	SER
3	C	48	TYR
3	B	426	THR
3	C	107	PHE
2	J	64	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	M	89	GLN
3	C	126	GLN
3	B	199	GLN
3	C	401	HIS
3	C	75	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](#)

Of 15 ligands modelled in this entry, 12 are monoatomic - leaving 3 for Mogul analysis.

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
6	NCO	C	505	-	6,6,6	4.48	6 (100%)	-		
6	NCO	B	505	-	6,6,6	4.87	6 (100%)	-		
6	NCO	A	505	-	6,6,6	5.57	6 (100%)	-		

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	505	NCO	CO-N2	-5.80	1.76	1.96
6	A	505	NCO	CO-N4	-5.66	1.77	1.96
6	A	505	NCO	CO-N3	-5.52	1.77	1.96
6	A	505	NCO	CO-N5	-5.50	1.77	1.96
6	A	505	NCO	CO-N1	-5.48	1.77	1.96

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	C	505	NCO	1	0

## 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	L	214/214 (100%)	-0.29	2 (0%) 84 63	6, 23, 44, 96	0
1	M	214/214 (100%)	-0.42	1 (0%) 91 75	3, 18, 43, 87	0
1	N	106/214 (49%)	2.47	64 (60%) 0 0	85, 106, 128, 140	0
2	H	208/230 (90%)	-0.32	0 100 100	5, 17, 66, 86	0
2	I	208/230 (90%)	-0.27	0 100 100	4, 16, 50, 72	0
2	J	106/230 (46%)	3.04	68 (64%) 0 0	96, 123, 141, 151	0
3	A	222/231 (96%)	-0.35	0 100 100	5, 18, 56, 77	0
3	B	222/231 (96%)	-0.37	0 100 100	5, 20, 46, 72	0
3	C	217/231 (93%)	1.47	65 (29%) 0 0	61, 92, 126, 142	0
All	All	1717/2025 (84%)	0.27	200 (11%) 4 1	3, 25, 122, 151	0

The worst 5 of 200 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	J	10	GLY	9.9
2	J	49	GLY	9.1
1	N	80	ALA	9.0
1	N	83	VAL	8.6
2	J	11	LEU	8.5

### 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(Å <sup>2</sup> )	Q<0.9
6	NCO	C	505	7/7	0.85	0.19	92,112,119,149	0
5	CA	C	504	1/1	0.92	0.17	101,101,101,101	0
5	CA	B	503	1/1	0.92	0.09	24,24,24,24	0
4	ZN	C	501	1/1	0.93	0.06	81,81,81,81	0
5	CA	C	503	1/1	0.93	0.10	75,75,75,75	0
4	ZN	C	502	1/1	0.97	0.07	78,78,78,78	0
5	CA	A	504	1/1	0.98	0.08	13,13,13,13	0
5	CA	B	504	1/1	0.99	0.06	12,12,12,12	0
5	CA	A	503	1/1	0.99	0.10	26,26,26,26	0
4	ZN	A	502	1/1	0.99	0.09	5,5,5,5	0
6	NCO	A	505	7/7	0.99	0.13	10,13,17,22	0
6	NCO	B	505	7/7	0.99	0.15	27,31,50,70	0
4	ZN	B	501	1/1	0.99	0.12	13,13,13,13	0
4	ZN	B	502	1/1	1.00	0.08	2,2,2,2	0
4	ZN	A	501	1/1	1.00	0.09	15,15,15,15	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.