



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 06:06 AM GMT

PDB ID : 2VX3  
Title : Crystal structure of the human dual specificity tyrosine- phosphorylation-regulated kinase 1A  
Authors : Roos, A.K.; Soundararajan, M.; Pike, A.C.W.; Federov, O.; King, O.; Burgess-Brown, N.; Philips, C.; Filippakopoulos, P.; Arrowsmith, C.H.; Wikstrom, M.; Edwards, A.; von Delft, F.; Bountra, C.; Knapp, S.  
Deposited on : 2008-06-30  
Resolution : 2.40 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

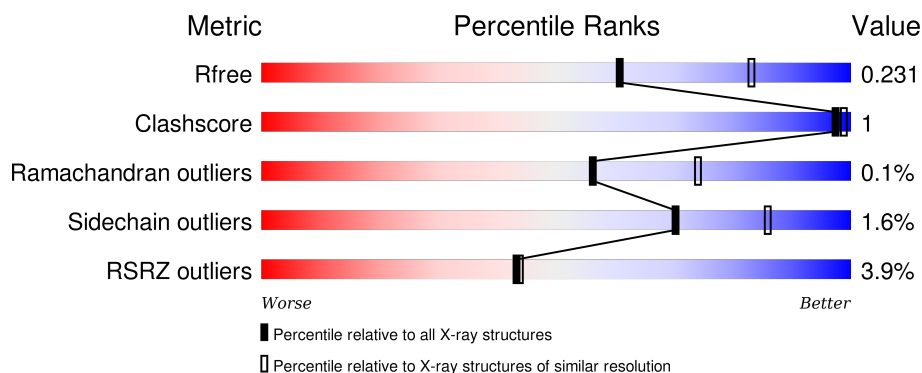
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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))
$R_{free}$	91344	2919 (2.40-2.40)
Clashscore	102246	3407 (2.40-2.40)
Ramachandran outliers	100387	3351 (2.40-2.40)
Sidechain outliers	100360	3352 (2.40-2.40)
RSRZ outliers	91569	2928 (2.40-2.40)

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	A	382	<div> <div> <div></div> <div>88%</div> <div>9%</div> </div> </div>
1	B	382	<div> <div> <div></div> <div>85%</div> <div>11%</div> </div> </div>
1	C	382	<div> <div> <div></div> <div>86%</div> <div>11%</div> </div> </div>
1	D	382	<div> <div> <div></div> <div>85%</div> <div>10%</div> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	P6G	A	700	-	-	-	X
4	P6G	B	700	-	-	-	X
4	P6G	C	700	-	-	-	X
5	SO4	B	1482	-	-	-	X
5	SO4	B	1483	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 11571 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 DUAL SPECIFICITY TYROSINE-PHOSPHORYLATION-REGULATED KINASE 1A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	347	Total	C	N	O	P	S	0	2	0
			2828	1815	489	506	1	17			
1	B	341	Total	C	N	O	P	S	0	1	0
			2760	1780	472	490	1	17			
1	C	339	Total	C	N	O	P	S	0	0	0
			2726	1760	462	487	1	16			
1	D	343	Total	C	N	O	P	S	0	2	0
			2772	1789	474	491	1	17			

There are 92 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	104	MET	-	EXPRESSION TAG	UNP Q13627
A	105	HIS	-	EXPRESSION TAG	UNP Q13627
A	106	HIS	-	EXPRESSION TAG	UNP Q13627
A	107	HIS	-	EXPRESSION TAG	UNP Q13627
A	108	HIS	-	EXPRESSION TAG	UNP Q13627
A	109	HIS	-	EXPRESSION TAG	UNP Q13627
A	110	HIS	-	EXPRESSION TAG	UNP Q13627
A	111	SER	-	EXPRESSION TAG	UNP Q13627
A	112	SER	-	EXPRESSION TAG	UNP Q13627
A	113	GLY	-	EXPRESSION TAG	UNP Q13627
A	114	VAL	-	EXPRESSION TAG	UNP Q13627
A	115	ASP	-	EXPRESSION TAG	UNP Q13627
A	116	LEU	-	EXPRESSION TAG	UNP Q13627
A	117	GLY	-	EXPRESSION TAG	UNP Q13627
A	118	THR	-	EXPRESSION TAG	UNP Q13627
A	119	GLU	-	EXPRESSION TAG	UNP Q13627
A	120	ASN	-	EXPRESSION TAG	UNP Q13627
A	121	LEU	-	EXPRESSION TAG	UNP Q13627
A	122	TYR	-	EXPRESSION TAG	UNP Q13627
A	123	PHE	-	EXPRESSION TAG	UNP Q13627

*Continued on next page...*

*Continued from previous page...*

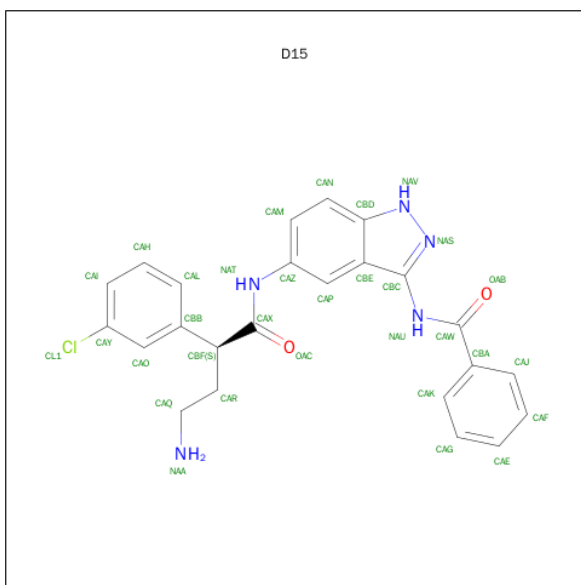
Chain	Residue	Modelled	Actual	Comment	Reference
A	124	GLN	-	EXPRESSION TAG	UNP Q13627
A	125	SER	-	EXPRESSION TAG	UNP Q13627
A	126	MET	-	EXPRESSION TAG	UNP Q13627
B	104	MET	-	EXPRESSION TAG	UNP Q13627
B	105	HIS	-	EXPRESSION TAG	UNP Q13627
B	106	HIS	-	EXPRESSION TAG	UNP Q13627
B	107	HIS	-	EXPRESSION TAG	UNP Q13627
B	108	HIS	-	EXPRESSION TAG	UNP Q13627
B	109	HIS	-	EXPRESSION TAG	UNP Q13627
B	110	HIS	-	EXPRESSION TAG	UNP Q13627
B	111	SER	-	EXPRESSION TAG	UNP Q13627
B	112	SER	-	EXPRESSION TAG	UNP Q13627
B	113	GLY	-	EXPRESSION TAG	UNP Q13627
B	114	VAL	-	EXPRESSION TAG	UNP Q13627
B	115	ASP	-	EXPRESSION TAG	UNP Q13627
B	116	LEU	-	EXPRESSION TAG	UNP Q13627
B	117	GLY	-	EXPRESSION TAG	UNP Q13627
B	118	THR	-	EXPRESSION TAG	UNP Q13627
B	119	GLU	-	EXPRESSION TAG	UNP Q13627
B	120	ASN	-	EXPRESSION TAG	UNP Q13627
B	121	LEU	-	EXPRESSION TAG	UNP Q13627
B	122	TYR	-	EXPRESSION TAG	UNP Q13627
B	123	PHE	-	EXPRESSION TAG	UNP Q13627
B	124	GLN	-	EXPRESSION TAG	UNP Q13627
B	125	SER	-	EXPRESSION TAG	UNP Q13627
B	126	MET	-	EXPRESSION TAG	UNP Q13627
C	104	MET	-	EXPRESSION TAG	UNP Q13627
C	105	HIS	-	EXPRESSION TAG	UNP Q13627
C	106	HIS	-	EXPRESSION TAG	UNP Q13627
C	107	HIS	-	EXPRESSION TAG	UNP Q13627
C	108	HIS	-	EXPRESSION TAG	UNP Q13627
C	109	HIS	-	EXPRESSION TAG	UNP Q13627
C	110	HIS	-	EXPRESSION TAG	UNP Q13627
C	111	SER	-	EXPRESSION TAG	UNP Q13627
C	112	SER	-	EXPRESSION TAG	UNP Q13627
C	113	GLY	-	EXPRESSION TAG	UNP Q13627
C	114	VAL	-	EXPRESSION TAG	UNP Q13627
C	115	ASP	-	EXPRESSION TAG	UNP Q13627
C	116	LEU	-	EXPRESSION TAG	UNP Q13627
C	117	GLY	-	EXPRESSION TAG	UNP Q13627
C	118	THR	-	EXPRESSION TAG	UNP Q13627
C	119	GLU	-	EXPRESSION TAG	UNP Q13627

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	120	ASN	-	EXPRESSION TAG	UNP Q13627
C	121	LEU	-	EXPRESSION TAG	UNP Q13627
C	122	TYR	-	EXPRESSION TAG	UNP Q13627
C	123	PHE	-	EXPRESSION TAG	UNP Q13627
C	124	GLN	-	EXPRESSION TAG	UNP Q13627
C	125	SER	-	EXPRESSION TAG	UNP Q13627
C	126	MET	-	EXPRESSION TAG	UNP Q13627
D	104	MET	-	EXPRESSION TAG	UNP Q13627
D	105	HIS	-	EXPRESSION TAG	UNP Q13627
D	106	HIS	-	EXPRESSION TAG	UNP Q13627
D	107	HIS	-	EXPRESSION TAG	UNP Q13627
D	108	HIS	-	EXPRESSION TAG	UNP Q13627
D	109	HIS	-	EXPRESSION TAG	UNP Q13627
D	110	HIS	-	EXPRESSION TAG	UNP Q13627
D	111	SER	-	EXPRESSION TAG	UNP Q13627
D	112	SER	-	EXPRESSION TAG	UNP Q13627
D	113	GLY	-	EXPRESSION TAG	UNP Q13627
D	114	VAL	-	EXPRESSION TAG	UNP Q13627
D	115	ASP	-	EXPRESSION TAG	UNP Q13627
D	116	LEU	-	EXPRESSION TAG	UNP Q13627
D	117	GLY	-	EXPRESSION TAG	UNP Q13627
D	118	THR	-	EXPRESSION TAG	UNP Q13627
D	119	GLU	-	EXPRESSION TAG	UNP Q13627
D	120	ASN	-	EXPRESSION TAG	UNP Q13627
D	121	LEU	-	EXPRESSION TAG	UNP Q13627
D	122	TYR	-	EXPRESSION TAG	UNP Q13627
D	123	PHE	-	EXPRESSION TAG	UNP Q13627
D	124	GLN	-	EXPRESSION TAG	UNP Q13627
D	125	SER	-	EXPRESSION TAG	UNP Q13627
D	126	MET	-	EXPRESSION TAG	UNP Q13627

- Molecule 2 is N-(5-{|(2S)-4-AMINO-2-(3-CHLOROPHENYL)BUTANOYL}AMINO}-1H-INDAZOL-3-YL)BENZAMIDE (three-letter code: D15) (formula: C<sub>24</sub>H<sub>22</sub>ClN<sub>5</sub>O<sub>2</sub>).

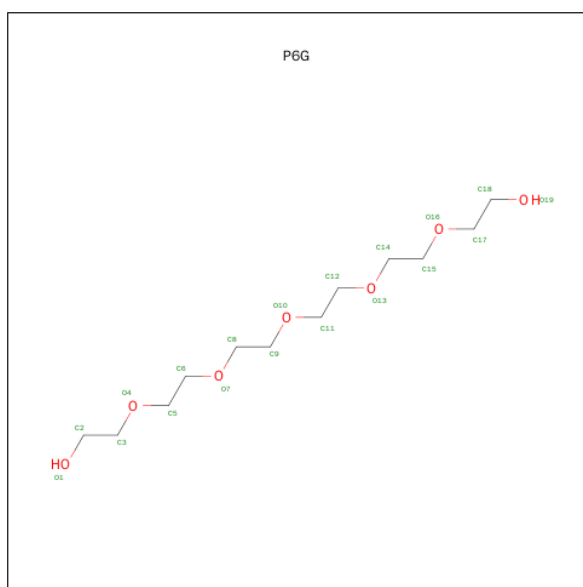


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total 32	C 24	Cl 1	N 5	O 2	0	0
2	B	1	Total 32	C 24	Cl 1	N 5	O 2	0	0
2	C	1	Total 32	C 24	Cl 1	N 5	O 2	0	0
2	D	1	Total 32	C 24	Cl 1	N 5	O 2	0	0

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	2	Total Cl 2 2	0	0
3	A	1	Total Cl 1 1	0	0
3	D	3	Total Cl 3 3	0	0
3	C	2	Total Cl 2 2	0	0

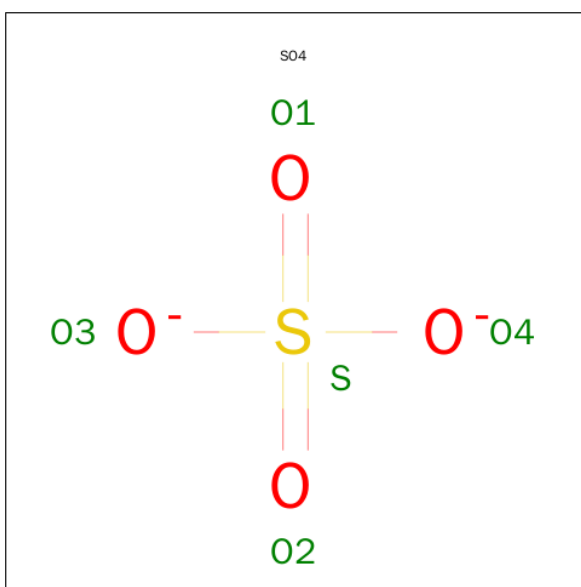
- Molecule 4 is HEXAETHYLENE GLYCOL (three-letter code: P6G) (formula:  $\text{C}_{12}\text{H}_{26}\text{O}_7$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 16	C 10	O 6	0	0
4	A	1	Total 13	C 8	O 5	0	0
4	A	1	Total 7	C 4	O 3	0	0
4	B	1	Total 16	C 10	O 6	0	0
4	B	1	Total 13	C 8	O 5	0	0
4	B	1	Total 7	C 4	O 3	0	0
4	C	1	Total 16	C 10	O 6	0	0
4	D	1	Total 16	C 10	O 6	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	C	1	Total	O	S	0	0
			5	4	1		
5	C	1	Total	O	S	0	0
			5	4	1		
5	C	1	Total	O	S	0	0
			5	4	1		
5	D	1	Total	O	S	0	0
			5	4	1		

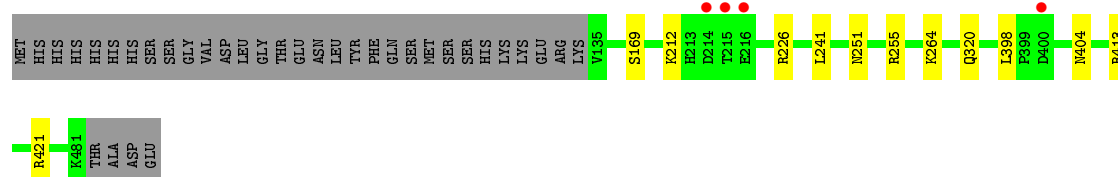
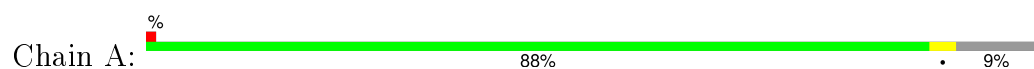
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	105	Total	O	0	0
			105	105		
6	B	52	Total	O	0	0
			52	52		
6	C	26	Total	O	0	0
			26	26		
6	D	22	Total	O	0	0
			22	22		

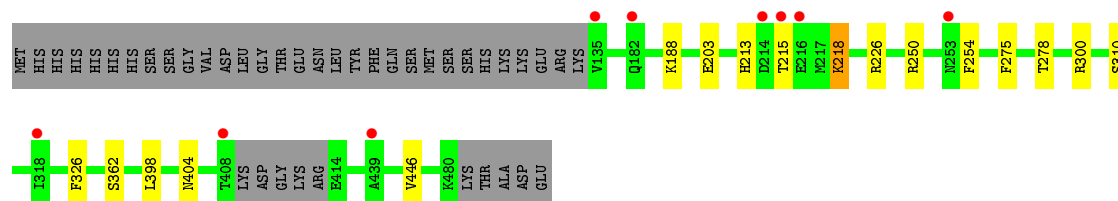
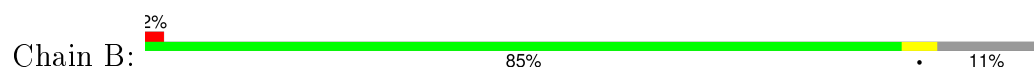
### 3 Residue-property plots [i](#)

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.

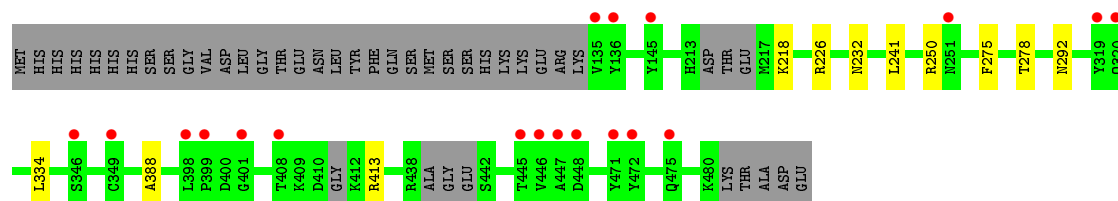
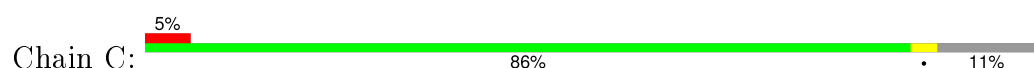
- Molecule 1: DUAL SPECIFICITY TYROSINE-PHOSPHORYLATION- REGULATED KINASE 1A



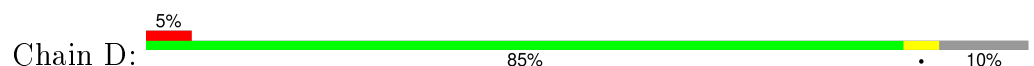
- Molecule 1: DUAL SPECIFICITY TYROSINE-PHOSPHORYLATION- REGULATED KINASE 1A

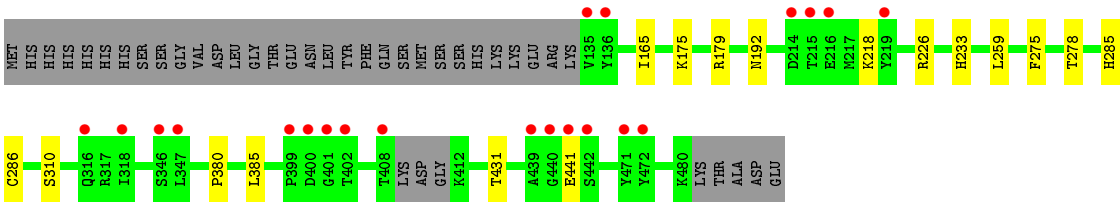


- Molecule 1: DUAL SPECIFICITY TYROSINE-PHOSPHORYLATION- REGULATED KINASE 1A



- Molecule 1: DUAL SPECIFICITY TYROSINE-PHOSPHORYLATION- REGULATED KINASE 1A





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	264.20Å 65.11Å 140.28Å 90.00° 115.44° 90.00°	Depositor
Resolution (Å)	26.00 – 2.40 25.96 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.9 (26.00-2.40) 99.9 (25.96-2.40)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.81 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.4.0066	Depositor
R, $R_{free}$	0.185 , 0.230 0.188 , 0.231	Depositor DCC
$R_{free}$ test set	1799 reflections (2.17%)	DCC
Wilson B-factor (Å <sup>2</sup> )	50.0	Xtriage
Anisotropy	0.040	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 44.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 84644 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	11571	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.37% 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

Bond lengths and bond angles in the following residue types are not validated in this section: D15, PTR, SO4, P6G, CL

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	0.74	0/2887	0.75	2/3895 (0.1%)
1	B	0.67	0/2811	0.71	0/3797
1	C	0.66	0/2772	0.71	1/3745 (0.0%)
1	D	0.65	0/2826	0.72	1/3816 (0.0%)
All	All	0.68	0/11296	0.72	4/15253 (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
1	B	0	1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	179	ARG	NE-CZ-NH1	5.67	123.14	120.30
1	A	421	ARG	NE-CZ-NH1	5.19	122.89	120.30
1	A	413	ARG	NE-CZ-NH2	-5.09	117.76	120.30
1	C	226	ARG	NE-CZ-NH1	-5.07	117.77	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	213	HIS	Peptide

## 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	A	2828	0	2803	3	0
1	B	2760	0	2733	8	0
1	C	2726	0	2662	4	0
1	D	2772	0	2731	7	0
2	A	32	0	22	3	0
2	B	32	0	22	2	0
2	C	32	0	22	2	0
2	D	32	0	22	0	0
3	A	1	0	0	0	0
3	B	2	0	0	0	0
3	C	2	0	0	0	0
3	D	3	0	0	0	0
4	A	36	0	47	0	0
4	B	36	0	47	0	0
4	C	16	0	21	1	0
4	D	16	0	21	0	0
5	A	5	0	0	0	0
5	B	15	0	0	0	0
5	C	15	0	0	0	0
5	D	5	0	0	0	0
6	A	105	0	0	1	0
6	B	52	0	0	1	0
6	C	26	0	0	0	0
6	D	22	0	0	0	0
All	All	11571	0	11153	27	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 1.

All (27) 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:275:PHE:O	1:B:278:THR:HG23	1.74	0.88
1:B:398:LEU:HD11	1:B:404:ASN:ND2	1.97	0.79
1:D:275:PHE:O	1:D:278:THR:HG23	1.83	0.78

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:600:D15:HAP	2:B:600:D15:OAC	1.85	0.73
2:A:600:D15:OAC	2:A:600:D15:HAP	1.99	0.60
1:B:398:LEU:HD11	1:B:404:ASN:HD22	1.65	0.59
4:C:700:P6G:O16	4:C:700:P6G:O1	2.23	0.56
2:A:600:D15:OAC	2:A:600:D15:CAP	2.56	0.53
1:D:380:PRO:HD2	1:D:385:LEU:HD11	1.91	0.53
1:D:285:HIS:O	1:D:286:CYS:HB2	2.10	0.52
1:A:398:LEU:HD11	1:A:404:ASN:ND2	2.25	0.51
1:B:326:PHE:CD1	1:B:362:SER:HA	2.47	0.50
1:B:215:THR:O	1:B:218:LYS:HG2	2.12	0.49
1:B:188:LYS:NZ	1:B:203:GLU:OE2	2.49	0.46
1:B:250:ARG:HG2	1:B:254:PHE:CZ	2.51	0.46
2:B:600:D15:CAP	2:B:600:D15:OAC	2.49	0.45
1:C:292:ASN:OD1	2:C:600:D15:NAA	2.46	0.45
1:A:241:LEU:O	2:A:600:D15:HAK	2.19	0.42
1:D:165:ILE:HG21	1:D:175:LYS:HB2	2.01	0.42
1:C:334:LEU:HB3	1:C:388:ALA:HB1	2.01	0.41
1:D:259:LEU:HD12	1:D:259:LEU:HA	1.90	0.41
1:C:241:LEU:O	2:C:600:D15:HAK	2.21	0.41
1:B:446:VAL:HG23	6:B:2045:HOH:O	2.20	0.41
1:A:264:LYS:NZ	6:A:2030:HOH:O	2.39	0.41
1:D:165:ILE:HD13	1:D:165:ILE:HG21	1.84	0.41
1:D:192:ASN:HB2	1:D:233:HIS:CE1	2.56	0.41
1:C:275:PHE:O	1:C:278:THR:HG23	2.20	0.41

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	A	346/382 (91%)	330 (95%)	16 (5%)	0	100 100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	337/382 (88%)	318 (94%)	19 (6%)	0	100	100
1	C	330/382 (86%)	315 (96%)	14 (4%)	1 (0%)	46	63
1	D	340/382 (89%)	325 (96%)	15 (4%)	0	100	100
All	All	1353/1528 (88%)	1288 (95%)	64 (5%)	1 (0%)	56	74

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	218	LYS

### 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	301/339 (89%)	295 (98%)	6 (2%)	63	81
1	B	292/339 (86%)	288 (99%)	4 (1%)	74	88
1	C	284/339 (84%)	281 (99%)	3 (1%)	80	92
1	D	289/339 (85%)	283 (98%)	6 (2%)	61	80
All	All	1166/1356 (86%)	1147 (98%)	19 (2%)	70	86

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	169	SER
1	A	212	LYS
1	A	226	ARG
1	A	251	ASN
1	A	255	ARG
1	A	320	GLN
1	B	218	LYS
1	B	226	ARG
1	B	300	ARG
1	B	310	SER

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	C	232	ASN
1	C	250	ARG
1	C	413	ARG
1	D	218	LYS
1	D	226[A]	ARG
1	D	226[B]	ARG
1	D	310	SER
1	D	431	THR
1	D	441	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	198	ASN
1	B	251	ASN
1	B	404	ASN
1	C	198	ASN
1	D	425	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

4 non-standard protein/DNA/RNA residues 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	$\# Z  > 2$
1	PTR	A	321	1	14,16,17	1.99	1 (7%)	18,22,24	1.41	2 (11%)
1	PTR	B	321	1	14,16,17	1.83	1 (7%)	18,22,24	1.28	3 (16%)
1	PTR	C	321	1	14,16,17	1.90	1 (7%)	18,22,24	1.34	3 (16%)
1	PTR	D	321	1	14,16,17	1.88	1 (7%)	18,22,24	1.31	2 (11%)

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
1	PTR	A	321	1	-	0/9/11/13	0/1/1/1
1	PTR	B	321	1	-	0/9/11/13	0/1/1/1
1	PTR	C	321	1	-	0/9/11/13	0/1/1/1
1	PTR	D	321	1	-	0/9/11/13	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	321	PTR	OH-CZ	-6.88	1.24	1.40
1	C	321	PTR	OH-CZ	-6.58	1.24	1.40
1	B	321	PTR	OH-CZ	-6.36	1.25	1.40
1	D	321	PTR	OH-CZ	-6.10	1.26	1.40

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	321	PTR	O3P-P-OH	2.20	113.10	105.22
1	B	321	PTR	O2P-P-OH	2.27	113.35	105.22
1	C	321	PTR	O3P-P-O2P	2.44	116.67	107.38
1	A	321	PTR	O3P-P-O2P	2.73	117.78	107.38
1	B	321	PTR	P-OH-CZ	2.82	131.88	123.76
1	B	321	PTR	O3P-P-O2P	2.85	118.22	107.38
1	D	321	PTR	O3P-P-OH	3.20	116.70	105.22
1	D	321	PTR	P-OH-CZ	3.31	133.29	123.76
1	C	321	PTR	P-OH-CZ	3.71	134.43	123.76
1	A	321	PTR	P-OH-CZ	3.84	134.81	123.76

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

Of 28 ligands modelled in this entry, 8 are monoatomic - leaving 20 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	$\# Z  > 2$
5	SO4	A	1483	-	4,4,4	0.35	0	6,6,6	0.49	0
2	D15	A	600	-	33,35,35	1.69	4 (12%)	36,48,48	2.48	10 (27%)
4	P6G	A	700	-	15,15,18	0.62	0	14,14,17	0.31	0
4	P6G	A	701	-	12,12,18	0.51	0	11,11,17	0.80	0
4	P6G	A	702	-	6,6,18	0.45	0	5,5,17	0.33	0
5	SO4	B	1482	-	4,4,4	0.37	0	6,6,6	0.31	0
5	SO4	B	1483	-	4,4,4	0.11	0	6,6,6	0.59	0
5	SO4	B	1484	-	4,4,4	0.10	0	6,6,6	0.56	0
2	D15	B	600	-	33,35,35	1.63	5 (15%)	36,48,48	2.31	8 (22%)
4	P6G	B	700	-	15,15,18	0.62	0	14,14,17	0.26	0
4	P6G	B	701	-	12,12,18	0.53	0	11,11,17	0.42	0
4	P6G	B	702	-	6,6,18	0.46	0	5,5,17	0.29	0
5	SO4	C	1482	-	4,4,4	0.14	0	6,6,6	0.31	0
5	SO4	C	1483	-	4,4,4	0.16	0	6,6,6	0.16	0
5	SO4	C	1484	-	4,4,4	0.14	0	6,6,6	0.36	0
2	D15	C	600	-	33,35,35	1.71	4 (12%)	36,48,48	1.86	7 (19%)
4	P6G	C	700	-	15,15,18	0.70	0	14,14,17	0.59	0
5	SO4	D	1482	-	4,4,4	0.10	0	6,6,6	0.41	0
2	D15	D	600	-	33,35,35	1.66	2 (6%)	36,48,48	1.92	8 (22%)
4	P6G	D	700	-	15,15,18	0.54	0	14,14,17	0.38	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
5	SO4	A	1483	-	-	0/0/0/0	0/0/0/0
2	D15	A	600	-	-	0/21/23/23	0/4/4/4
4	P6G	A	700	-	-	0/13/13/16	0/0/0/0
4	P6G	A	701	-	-	0/10/10/16	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	P6G	A	702	-	-	0/4/4/16	0/0/0/0
5	SO4	B	1482	-	-	0/0/0/0	0/0/0/0
5	SO4	B	1483	-	-	0/0/0/0	0/0/0/0
5	SO4	B	1484	-	-	0/0/0/0	0/0/0/0
2	D15	B	600	-	-	0/21/23/23	0/4/4/4
4	P6G	B	700	-	-	0/13/13/16	0/0/0/0
4	P6G	B	701	-	-	0/10/10/16	0/0/0/0
4	P6G	B	702	-	-	0/4/4/16	0/0/0/0
5	SO4	C	1482	-	-	0/0/0/0	0/0/0/0
5	SO4	C	1483	-	-	0/0/0/0	0/0/0/0
5	SO4	C	1484	-	-	0/0/0/0	0/0/0/0
2	D15	C	600	-	-	0/21/23/23	0/4/4/4
4	P6G	C	700	-	-	0/13/13/16	0/0/0/0
5	SO4	D	1482	-	-	0/0/0/0	0/0/0/0
2	D15	D	600	-	-	0/21/23/23	0/4/4/4
4	P6G	D	700	-	-	0/13/13/16	0/0/0/0

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	600	D15	NAV-NAS	-7.44	1.23	1.37
2	C	600	D15	NAV-NAS	-7.27	1.23	1.37
2	A	600	D15	NAV-NAS	-6.75	1.24	1.37
2	B	600	D15	NAV-NAS	-6.39	1.25	1.37
2	C	600	D15	CAZ-NAT	-2.87	1.36	1.41
2	A	600	D15	CAZ-NAT	-2.49	1.36	1.41
2	B	600	D15	CAR-CBF	-2.14	1.49	1.53
2	B	600	D15	CAN-CBD	-2.05	1.38	1.41
2	C	600	D15	CAP-CAZ	2.29	1.42	1.37
2	B	600	D15	CAP-CAZ	2.55	1.43	1.37
2	B	600	D15	CAO-CAY	2.58	1.42	1.38
2	D	600	D15	CAP-CAZ	2.78	1.43	1.37
2	A	600	D15	CBA-CAW	2.88	1.56	1.50
2	C	600	D15	CBA-CAW	3.11	1.56	1.50
2	A	600	D15	CAP-CAZ	3.36	1.44	1.37

All (33) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	600	D15	CAP-CBE-CBD	-8.08	117.15	120.34
2	D	600	D15	CAP-CBE-CBD	-5.94	118.00	120.34
2	B	600	D15	CAI-CAY-CL1	-5.34	110.59	119.35

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	600	D15	CAZ-NAT-CAX	-5.16	118.12	127.40
2	B	600	D15	CAZ-NAT-CAX	-5.14	118.16	127.40
2	B	600	D15	CAL-CBB-CBF	-4.57	112.20	120.93
2	C	600	D15	CAL-CBB-CBF	-4.50	112.33	120.93
2	B	600	D15	CAP-CBE-CBD	-4.37	118.61	120.34
2	C	600	D15	CAZ-NAT-CAX	-4.18	119.89	127.40
2	A	600	D15	CAL-CBB-CBF	-4.09	113.11	120.93
2	C	600	D15	CAI-CAY-CL1	-3.87	113.01	119.35
2	A	600	D15	CBC-CBE-CBD	-3.86	101.45	106.56
2	A	600	D15	CAI-CAY-CL1	-3.80	113.11	119.35
2	B	600	D15	CBC-CBE-CBD	-3.79	101.55	106.56
2	D	600	D15	CAZ-NAT-CAX	-3.77	120.62	127.40
2	D	600	D15	CAL-CBB-CBF	-3.46	114.33	120.93
2	D	600	D15	CBC-CBE-CBD	-3.43	102.02	106.56
2	A	600	D15	CAY-CAO-CBB	-3.43	115.36	119.38
2	B	600	D15	CAY-CAO-CBB	-3.20	115.62	119.38
2	C	600	D15	CAP-CBE-CBD	-3.06	119.13	120.34
2	D	600	D15	CAY-CAO-CBB	-2.67	116.24	119.38
2	C	600	D15	CBC-CBE-CBD	-2.61	103.11	106.56
2	D	600	D15	CAM-CAZ-CAP	-2.28	116.12	119.56
2	A	600	D15	CAL-CBB-CAO	2.15	121.38	118.79
2	A	600	D15	CAM-CAN-CBD	2.82	123.94	120.88
2	D	600	D15	CAO-CAY-CL1	2.92	122.77	119.14
2	A	600	D15	CAO-CBB-CBF	3.29	125.51	120.21
2	D	600	D15	CAL-CBB-CAO	3.65	123.17	118.79
2	A	600	D15	CAO-CAY-CL1	4.08	124.21	119.14
2	C	600	D15	CAO-CBB-CBF	4.22	127.01	120.21
2	B	600	D15	CAO-CBB-CBF	4.44	127.38	120.21
2	C	600	D15	CAO-CAY-CL1	4.88	125.20	119.14
2	B	600	D15	CAO-CAY-CL1	5.94	126.52	119.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	600	D15	3	0
2	B	600	D15	2	0
2	C	600	D15	2	0
4	C	700	P6G	1	0

## 5.7 Other polymers

There are no such residues in this entry.

## 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	346/382 (90%)	-0.29	4 (1%) 81 81	21, 29, 43, 59	0
1	B	340/382 (89%)	-0.23	9 (2%) 59 58	19, 30, 43, 62	0
1	C	338/382 (88%)	-0.02	19 (5%) 28 28	21, 30, 43, 54	0
1	D	342/382 (89%)	-0.04	21 (6%) 25 25	21, 30, 43, 86	0
All	All	1366/1528 (89%)	-0.14	53 (3%) 43 44	19, 30, 43, 86	0

All (53) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	440	GLY	7.2
1	A	215	THR	6.0
1	B	215	THR	5.4
1	B	214	ASP	4.9
1	D	401	GLY	4.6
1	D	399	PRO	4.3
1	D	136	TYR	4.3
1	D	214	ASP	4.3
1	D	439	ALA	4.1
1	C	399	PRO	4.0
1	C	471	TYR	3.9
1	D	442	SER	3.6
1	C	446	VAL	3.5
1	C	472	TYR	3.5
1	C	135	VAL	3.5
1	B	135	VAL	3.4
1	B	318	ILE	3.4
1	C	475	GLN	3.4
1	C	136	TYR	3.3
1	B	439	ALA	3.2
1	D	347	LEU	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	408	THR	3.0
1	C	408	THR	2.9
1	D	318	ILE	2.9
1	B	216	GLU	2.8
1	C	398	LEU	2.8
1	D	215	THR	2.8
1	A	214	ASP	2.7
1	D	346	SER	2.7
1	D	402	THR	2.6
1	C	447	ALA	2.5
1	D	400	ASP	2.5
1	C	319	TYR	2.4
1	C	445	THR	2.4
1	D	219	TYR	2.4
1	D	216	GLU	2.4
1	D	471	TYR	2.3
1	C	346	SER	2.3
1	C	401	GLY	2.3
1	C	145	TYR	2.2
1	D	135	VAL	2.2
1	C	251	ASN	2.2
1	C	448	ASP	2.2
1	B	182	GLN	2.2
1	D	316	GLN	2.2
1	C	320	GLN	2.2
1	C	349	CYS	2.1
1	D	408	THR	2.1
1	D	441	GLU	2.1
1	B	253	ASN	2.1
1	A	216	GLU	2.1
1	A	400	ASP	2.0
1	D	472	TYR	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [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. 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	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
1	PTR	B	321	16/17	0.94	0.25	-	37,47,60,62	0
1	PTR	A	321	16/17	0.96	0.13	-	34,44,57,58	0
1	PTR	D	321	16/17	0.91	0.27	-	37,47,60,61	0
1	PTR	C	321	16/17	0.92	0.18	-	34,45,59,60	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates 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. 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	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	P6G	C	700	16/19	0.90	0.21	5.09	47,66,83,85	0
4	P6G	A	700	16/19	0.92	0.19	3.68	62,72,76,78	0
5	SO4	B	1483	5/5	0.92	0.28	3.53	72,76,98,100	0
4	P6G	B	700	16/19	0.86	0.17	3.50	64,75,77,79	0
5	SO4	B	1482	5/5	0.96	0.25	2.83	65,66,88,91	0
4	P6G	A	702	7/19	0.93	0.17	1.02	55,55,64,70	0
5	SO4	D	1482	5/5	0.98	0.17	0.49	55,69,82,90	0
4	P6G	D	700	16/19	0.95	0.11	0.25	37,46,64,64	0
4	P6G	A	701	13/19	0.95	0.13	-0.20	32,46,55,65	0
5	SO4	C	1482	5/5	0.97	0.14	-0.21	75,86,88,92	0
4	P6G	B	701	13/19	0.94	0.12	-0.45	33,44,50,52	0
2	D15	D	600	32/32	0.95	0.11	-0.61	31,50,61,65	0
2	D15	C	600	32/32	0.95	0.11	-0.72	40,55,67,73	0
2	D15	B	600	32/32	0.97	0.09	-1.08	25,40,55,55	0
2	D15	A	600	32/32	0.96	0.10	-1.11	29,36,50,64	0
5	SO4	A	1483	5/5	0.98	0.09	-1.65	54,57,60,61	0
3	CL	D	1484	1/1	0.77	0.14	-	85,85,85,85	0
5	SO4	C	1483	5/5	0.96	0.31	-	99,106,109,112	0
5	SO4	C	1484	5/5	0.94	0.23	-	83,87,96,104	0
3	CL	D	1483	1/1	0.92	0.08	-	75,75,75,75	0
3	CL	C	620	1/1	0.98	0.09	-	61,61,61,61	0
3	CL	A	620	1/1	0.98	0.07	-	45,45,45,45	0
4	P6G	B	702	7/19	0.95	0.15	-	67,80,83,85	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	CL	D	620	1/1	0.98	0.09	-	56,56,56,56	0
5	SO4	B	1484	5/5	0.97	0.16	-	65,66,72,81	0
3	CL	B	1485	1/1	0.95	0.07	-	59,59,59,59	0
3	CL	B	620	1/1	0.98	0.09	-	47,47,47,47	0
3	CL	C	1485	1/1	0.95	0.09	-	75,75,75,75	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.