



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 12, 2021 – 12:03 PM EDT

PDB ID : 1ZMX  
Title : Crystal structure of D. melanogaster deoxyribonucleoside kinase N64D mutant in complex with thymidine  
Authors : Welin, M.; Skovgaard, T.; Knecht, W.; Berenstein, D.; Munch-Petersen, B.; Piskur, J.; Eklund, H.  
Deposited on : 2005-05-11  
Resolution : 3.10 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 : 2.23.2  
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) : 2.23.2

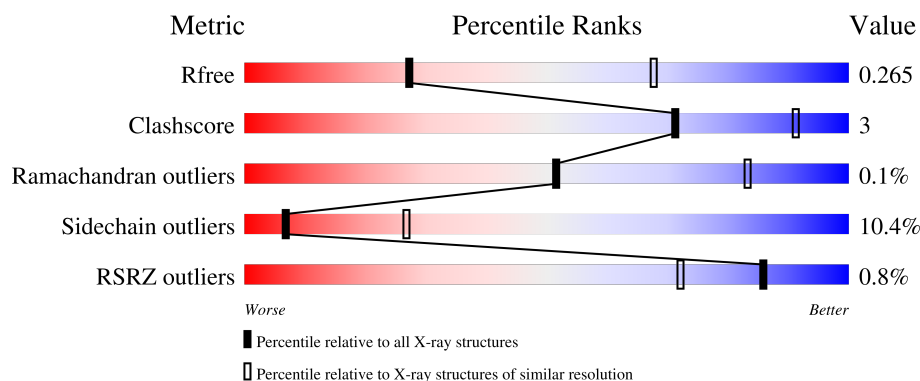
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


The reported resolution of this entry is 3.10 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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	230	
1	B	230	
1	C	230	
1	D	230	
1	E	230	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	230	<div><div><div>%</div><div><div></div><div></div><div></div></div><div>68%12%19%</div></div></div>
1	G	230	<div><div><div></div><div></div><div></div></div><div>69%12%18%</div></div>
1	H	230	<div><div><div>%</div><div><div></div><div></div><div></div></div><div>69%10%20%</div></div></div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 12712 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 Deoxynucleoside kinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	190	Total	C	N	O	S	0	0	0
			1588	1018	271	290	9			
1	B	185	Total	C	N	O	S	0	0	0
			1544	993	259	283	9			
1	C	191	Total	C	N	O	S	0	0	0
			1594	1021	272	291	10			
1	D	185	Total	C	N	O	S	0	0	0
			1544	993	259	283	9			
1	E	190	Total	C	N	O	S	0	0	0
			1588	1018	271	290	9			
1	F	187	Total	C	N	O	S	0	0	0
			1556	999	261	286	10			
1	G	189	Total	C	N	O	S	0	0	0
			1580	1014	270	287	9			
1	H	185	Total	C	N	O	S	0	0	0
			1542	992	259	281	10			

There are 8 discrepancies between the modelled and reference sequences:

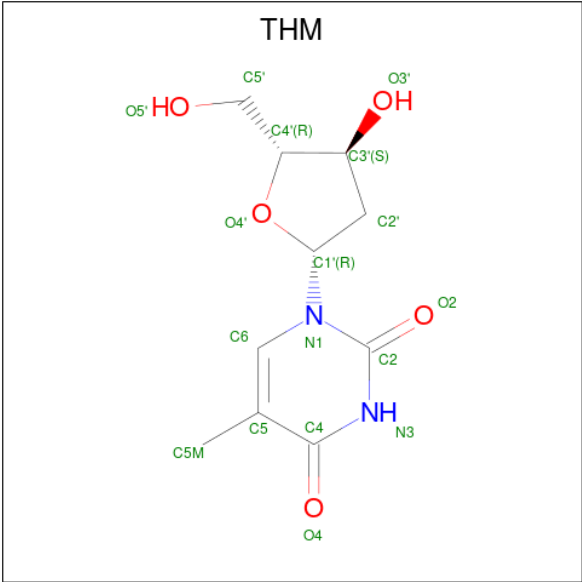
Chain	Residue	Modelled	Actual	Comment	Reference
A	64	ASP	ASN	engineered mutation	UNP Q9XZT6
B	64	ASP	ASN	engineered mutation	UNP Q9XZT6
C	64	ASP	ASN	engineered mutation	UNP Q9XZT6
D	64	ASP	ASN	engineered mutation	UNP Q9XZT6
E	64	ASP	ASN	engineered mutation	UNP Q9XZT6
F	64	ASP	ASN	engineered mutation	UNP Q9XZT6
G	64	ASP	ASN	engineered mutation	UNP Q9XZT6
H	64	ASP	ASN	engineered mutation	UNP Q9XZT6

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		
2	E	1	Total	O	S	0	0
			5	4	1		
2	F	1	Total	O	S	0	0
			5	4	1		
2	G	1	Total	O	S	0	0
			5	4	1		
2	H	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is THYMIDINE (three-letter code: THM) (formula: C<sub>10</sub>H<sub>14</sub>N<sub>2</sub>O<sub>5</sub>).

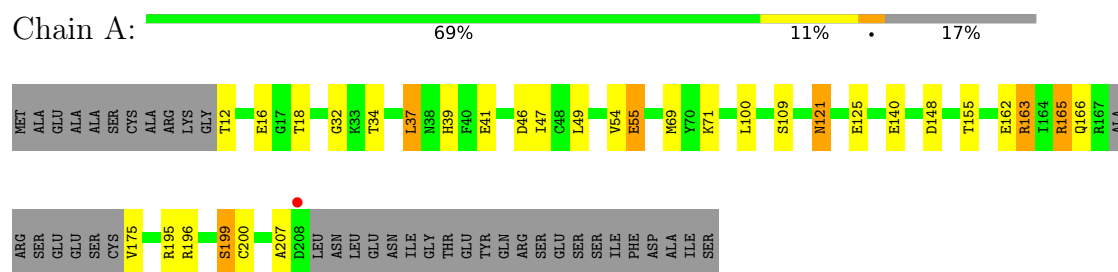


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			17	10	2	5		
3	B	1	Total	C	N	O	0	0
			17	10	2	5		
3	C	1	Total	C	N	O	0	0
			17	10	2	5		
3	D	1	Total	C	N	O	0	0
			17	10	2	5		
3	E	1	Total	C	N	O	0	0
			17	10	2	5		
3	F	1	Total	C	N	O	0	0
			17	10	2	5		
3	G	1	Total	C	N	O	0	0
			17	10	2	5		
3	H	1	Total	C	N	O	0	0
			17	10	2	5		

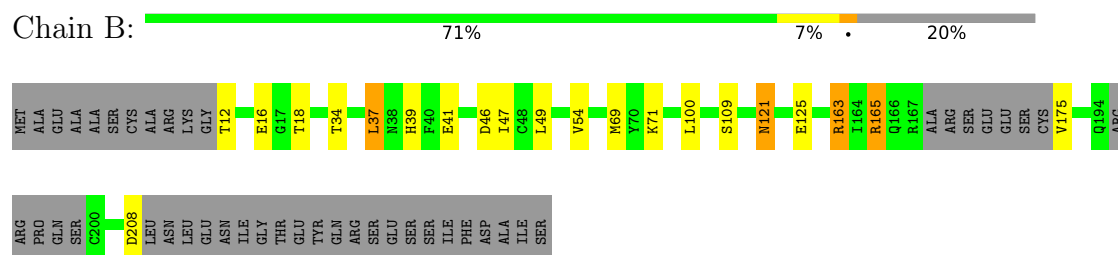
### 3 Residue-property plots

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.

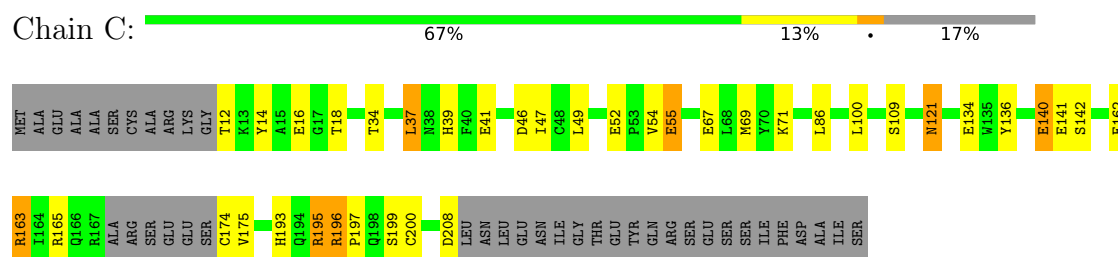
#### • Molecule 1: Deoxynucleoside kinase



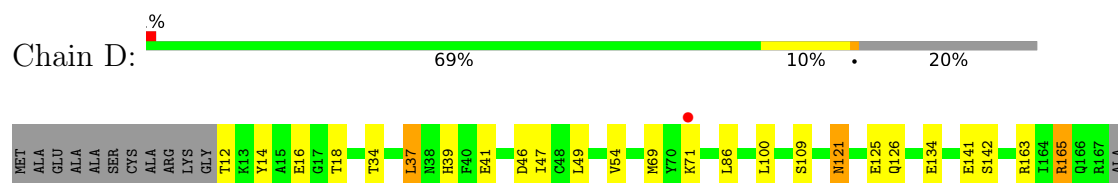
#### • Molecule 1: Deoxynucleoside kinase

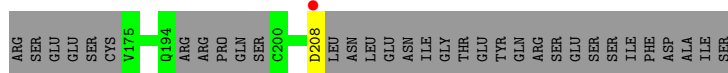


#### • Molecule 1: Deoxynucleoside kinase

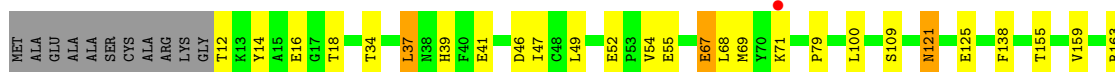


#### • Molecule 1: Deoxynucleoside kinase





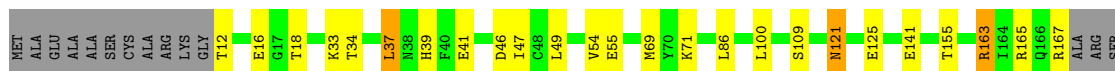
• Molecule 1: Deoxynucleoside kinase



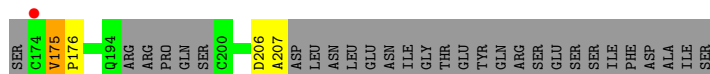
• Molecule 1: Deoxynucleoside kinase



• Molecule 1: Deoxynucleoside kinase



• Molecule 1: Deoxynucleoside kinase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.71Å 70.34Å 224.53Å 90.00° 90.69° 90.00°	Depositor
Resolution (Å)	41.17 – 3.10 41.13 – 3.10	Depositor EDS
% Data completeness (in resolution range)	100.0 (41.17-3.10) 100.0 (41.13-3.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.09 (at 3.12Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, $R_{free}$	0.269 , 0.288 0.237 , 0.265	Depositor DCC
$R_{free}$ test set	2005 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	54.9	Xtriage
Anisotropy	0.374	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 32.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.012 for -k,-h,-l 0.008 for k,h,-l 0.027 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	12712	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.08% 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: THM, SO4

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	0.45	0/1625	0.66	2/2196 (0.1%)
1	B	0.44	0/1579	0.65	1/2133 (0.0%)
1	C	0.47	0/1631	0.67	1/2204 (0.0%)
1	D	0.45	0/1579	0.67	1/2133 (0.0%)
1	E	0.46	0/1625	0.66	1/2196 (0.0%)
1	F	0.45	0/1591	0.64	1/2149 (0.0%)
1	G	0.47	0/1617	0.65	0/2185
1	H	0.47	0/1577	0.65	0/2130
All	All	0.46	0/12824	0.66	7/17326 (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	G	0	1

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	208	ASP	CB-CG-OD2	6.07	123.76	118.30
1	B	208	ASP	CB-CG-OD2	5.90	123.61	118.30
1	C	208	ASP	CB-CG-OD2	5.71	123.44	118.30
1	F	208	ASP	CB-CG-OD2	5.55	123.30	118.30
1	E	208	ASP	CB-CG-OD2	5.35	123.12	118.30
1	A	140	GLU	CA-CB-CG	5.27	125.00	113.40
1	A	148	ASP	CB-CG-OD2	5.04	122.83	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	G	199	SER	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	1588	0	1581	12	0
1	B	1544	0	1534	7	0
1	C	1594	0	1586	17	0
1	D	1544	0	1534	12	0
1	E	1588	0	1581	16	0
1	F	1556	0	1544	16	0
1	G	1580	0	1577	9	0
1	H	1542	0	1535	15	0
2	A	5	0	0	1	0
2	B	5	0	0	0	0
2	C	5	0	0	0	0
2	D	5	0	0	0	0
2	E	5	0	0	0	0
2	F	5	0	0	0	0
2	G	5	0	0	1	0
2	H	5	0	0	0	0
3	A	17	0	14	0	0
3	B	17	0	14	0	0
3	C	17	0	14	1	0
3	D	17	0	14	0	0
3	E	17	0	14	1	0
3	F	17	0	14	1	0
3	G	17	0	14	0	0
3	H	17	0	14	0	0
All	All	12712	0	12584	88	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 3.

All (88) 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:D:165:ARG:HH11	1:D:165:ARG:HG3	1.26	0.96
1:D:165:ARG:HG3	1:D:165:ARG:NH1	1.97	0.77
1:E:190:TRP:HE1	1:E:198:GLN:HB2	1.53	0.74
1:A:165:ARG:HH11	1:A:165:ARG:HB2	1.53	0.73
1:F:194:GLN:HB2	1:H:44:LYS:NZ	2.06	0.70
1:F:194:GLN:HB2	1:H:44:LYS:HZ2	1.57	0.67
1:A:121:ASN:HD22	1:A:121:ASN:C	2.00	0.65
1:E:121:ASN:HD22	1:E:121:ASN:C	2.00	0.65
1:F:121:ASN:C	1:F:121:ASN:HD22	2.00	0.65
1:G:121:ASN:C	1:G:121:ASN:HD22	2.00	0.65
1:C:121:ASN:C	1:C:121:ASN:HD22	2.01	0.65
1:E:155:THR:HG22	1:E:207:ALA:HB3	1.78	0.65
1:B:121:ASN:C	1:B:121:ASN:HD22	2.00	0.64
1:D:121:ASN:C	1:D:121:ASN:HD22	2.00	0.64
1:H:121:ASN:HD22	1:H:121:ASN:C	2.01	0.63
1:E:47:ILE:HG12	1:E:100:LEU:HD23	1.83	0.61
1:B:47:ILE:HG12	1:B:100:LEU:HD23	1.82	0.61
1:C:47:ILE:HG12	1:C:100:LEU:HD23	1.83	0.61
1:F:194:GLN:CB	1:H:44:LYS:HZ2	2.12	0.61
1:F:47:ILE:HG12	1:F:100:LEU:HD23	1.82	0.61
1:E:159:VAL:HG11	1:E:208:ASP:HB3	1.82	0.61
1:H:47:ILE:HG12	1:H:100:LEU:HD23	1.83	0.60
1:G:47:ILE:HG12	1:G:100:LEU:HD23	1.83	0.60
1:A:47:ILE:HG12	1:A:100:LEU:HD23	1.83	0.60
1:G:37:LEU:HD23	1:G:49:LEU:HD11	1.84	0.60
1:A:199:SER:O	1:A:200:CYS:HB3	2.02	0.60
1:D:47:ILE:HG12	1:D:100:LEU:HD23	1.82	0.60
1:D:37:LEU:HD23	1:D:49:LEU:HD11	1.85	0.59
1:B:37:LEU:HD23	1:B:49:LEU:HD11	1.84	0.58
1:F:37:LEU:HD23	1:F:49:LEU:HD11	1.85	0.58
1:C:37:LEU:HD23	1:C:49:LEU:HD11	1.85	0.58
1:A:37:LEU:HD23	1:A:49:LEU:HD11	1.85	0.58
1:H:37:LEU:HD23	1:H:49:LEU:HD11	1.85	0.57
1:E:37:LEU:HD23	1:E:49:LEU:HD11	1.85	0.57
1:C:52:GLU:OE2	3:C:1406:THM:HM51	2.05	0.56
1:A:32:GLY:N	2:A:1300:SO4:O1	2.36	0.56
1:C:55:GLU:H	1:C:55:GLU:CD	2.10	0.54
1:E:138:PHE:CB	1:F:86:LEU:HD11	2.39	0.51
1:E:190:TRP:HE1	1:E:198:GLN:CB	2.24	0.51
1:E:165:ARG:HB3	1:E:165:ARG:CZ	2.41	0.51
1:C:174:CYS:SG	1:C:175:VAL:N	2.83	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:193:HIS:HB3	1:C:195:ARG:HH11	1.75	0.50
1:C:142:SER:HB2	1:D:14:TYR:CD2	2.47	0.50
1:A:155:THR:HG22	1:A:207:ALA:HB3	1.93	0.49
1:C:196:ARG:HG3	1:C:197:PRO:HD2	1.95	0.49
1:C:134:GLU:HB3	1:D:86:LEU:HD21	1.94	0.48
1:G:155:THR:HG22	1:G:207:ALA:HB3	1.96	0.47
1:B:165:ARG:HA	1:B:165:ARG:HD2	1.67	0.47
1:G:33:LYS:N	2:G:1302:SO4:O4	2.48	0.47
1:C:86:LEU:HD21	1:D:134:GLU:HB3	1.97	0.46
1:C:136:TYR:O	1:C:140:GLU:HG2	2.17	0.46
1:G:196:ARG:HA	1:G:197:PRO:HD3	1.82	0.45
1:E:138:PHE:HB2	1:F:86:LEU:HD11	1.99	0.45
1:C:163:ARG:HE	1:C:163:ARG:HB2	1.60	0.44
1:E:14:TYR:CD2	1:F:142:SER:HB2	2.52	0.44
1:H:163:ARG:HE	1:H:163:ARG:HB2	1.60	0.44
1:F:194:GLN:CB	1:H:44:LYS:NZ	2.74	0.44
1:A:165:ARG:HB2	1:A:165:ARG:NH1	2.27	0.44
1:F:163:ARG:HE	1:F:163:ARG:HB2	1.60	0.44
1:C:86:LEU:CD2	1:D:134:GLU:HB3	2.47	0.43
1:H:155:THR:HG22	1:H:207:ALA:HB2	2.00	0.43
1:C:14:TYR:CD2	1:D:142:SER:HB2	2.54	0.43
1:E:52:GLU:OE2	3:E:1404:THM:HM51	2.19	0.43
1:A:16:GLU:O	1:A:18:THR:HG23	2.20	0.42
1:D:16:GLU:O	1:D:18:THR:HG23	2.20	0.42
1:A:155:THR:CG2	1:A:207:ALA:HB3	2.49	0.42
1:C:16:GLU:O	1:C:18:THR:HG23	2.20	0.42
1:C:142:SER:HB2	1:D:14:TYR:CG	2.55	0.42
1:E:79:PRO:HA	1:F:78:MET:CE	2.50	0.42
1:G:86:LEU:HD11	1:H:138:PHE:CB	2.50	0.42
1:F:193:HIS:ND1	1:H:41:GLU:OE2	2.45	0.41
1:H:16:GLU:O	1:H:18:THR:HG23	2.20	0.41
1:A:163:ARG:HE	1:A:163:ARG:HB2	1.60	0.41
1:E:67:GLU:HG3	1:E:68:LEU:N	2.36	0.41
1:F:175:VAL:HA	1:F:176:PRO:HD3	1.96	0.41
1:H:175:VAL:HA	1:H:176:PRO:HD2	1.88	0.41
1:G:16:GLU:O	1:G:18:THR:HG23	2.19	0.41
1:A:55:GLU:H	1:A:55:GLU:HG2	1.56	0.41
1:F:16:GLU:O	1:F:18:THR:HG23	2.20	0.41
1:H:121:ASN:C	1:H:121:ASN:ND2	2.73	0.41
1:H:206:ASP:O	1:H:207:ALA:C	2.59	0.41
1:B:37:LEU:CD2	1:B:49:LEU:HD11	2.51	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:16:GLU:O	1:E:18:THR:HG23	2.20	0.41
1:G:163:ARG:HE	1:G:163:ARG:HB2	1.61	0.41
1:B:16:GLU:O	1:B:18:THR:HG23	2.19	0.40
1:F:52:GLU:OE2	3:F:1405:THM:HM51	2.21	0.40
1:E:155:THR:CG2	1:E:207:ALA:HB3	2.46	0.40
1:B:163:ARG:HE	1:B:163:ARG:HB2	1.60	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	
1	A	186/230 (81%)	181 (97%)	5 (3%)	0	100	100
1	B	179/230 (78%)	175 (98%)	4 (2%)	0	100	100
1	C	187/230 (81%)	182 (97%)	5 (3%)	0	100	100
1	D	179/230 (78%)	175 (98%)	4 (2%)	0	100	100
1	E	186/230 (81%)	181 (97%)	4 (2%)	1 (0%)	29	64
1	F	181/230 (79%)	176 (97%)	5 (3%)	0	100	100
1	G	185/230 (80%)	178 (96%)	7 (4%)	0	100	100
1	H	179/230 (78%)	174 (97%)	5 (3%)	0	100	100
All	All	1462/1840 (80%)	1422 (97%)	39 (3%)	1 (0%)	51	83

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	198	GLN

### 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	176/208 (85%)	155 (88%)	21 (12%)	5	20
1	B	171/208 (82%)	156 (91%)	15 (9%)	10	36
1	C	177/208 (85%)	155 (88%)	22 (12%)	4	19
1	D	171/208 (82%)	155 (91%)	16 (9%)	8	32
1	E	176/208 (85%)	158 (90%)	18 (10%)	7	27
1	F	173/208 (83%)	157 (91%)	16 (9%)	9	33
1	G	175/208 (84%)	155 (89%)	20 (11%)	5	23
1	H	171/208 (82%)	155 (91%)	16 (9%)	8	32
All	All	1390/1664 (84%)	1246 (90%)	144 (10%)	7	27

All (144) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	12	THR
1	A	34	THR
1	A	37	LEU
1	A	39	HIS
1	A	41	GLU
1	A	46	ASP
1	A	54	VAL
1	A	55	GLU
1	A	69	MET
1	A	71	LYS
1	A	109	SER
1	A	121	ASN
1	A	125	GLU
1	A	162	GLU
1	A	163	ARG
1	A	165	ARG
1	A	166	GLN
1	A	175	VAL
1	A	195	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	196	ARG
1	A	199	SER
1	B	12	THR
1	B	34	THR
1	B	37	LEU
1	B	39	HIS
1	B	41	GLU
1	B	46	ASP
1	B	54	VAL
1	B	69	MET
1	B	71	LYS
1	B	109	SER
1	B	121	ASN
1	B	125	GLU
1	B	163	ARG
1	B	165	ARG
1	B	175	VAL
1	C	12	THR
1	C	34	THR
1	C	37	LEU
1	C	39	HIS
1	C	41	GLU
1	C	46	ASP
1	C	54	VAL
1	C	55	GLU
1	C	67	GLU
1	C	69	MET
1	C	71	LYS
1	C	109	SER
1	C	121	ASN
1	C	140	GLU
1	C	141	GLU
1	C	162	GLU
1	C	163	ARG
1	C	165	ARG
1	C	195	ARG
1	C	196	ARG
1	C	199	SER
1	C	200	CYS
1	D	12	THR
1	D	34	THR
1	D	37	LEU

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	D	39	HIS
1	D	41	GLU
1	D	46	ASP
1	D	54	VAL
1	D	69	MET
1	D	71	LYS
1	D	109	SER
1	D	121	ASN
1	D	125	GLU
1	D	126	GLN
1	D	141	GLU
1	D	163	ARG
1	D	165	ARG
1	E	12	THR
1	E	34	THR
1	E	37	LEU
1	E	39	HIS
1	E	41	GLU
1	E	46	ASP
1	E	54	VAL
1	E	55	GLU
1	E	67	GLU
1	E	69	MET
1	E	71	LYS
1	E	109	SER
1	E	121	ASN
1	E	125	GLU
1	E	163	ARG
1	E	165	ARG
1	E	198	GLN
1	E	208	ASP
1	F	12	THR
1	F	34	THR
1	F	37	LEU
1	F	39	HIS
1	F	41	GLU
1	F	46	ASP
1	F	54	VAL
1	F	67	GLU
1	F	69	MET
1	F	71	LYS
1	F	109	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	F	121	ASN
1	F	126	GLN
1	F	163	ARG
1	F	165	ARG
1	F	166	GLN
1	G	12	THR
1	G	34	THR
1	G	37	LEU
1	G	39	HIS
1	G	41	GLU
1	G	46	ASP
1	G	54	VAL
1	G	55	GLU
1	G	69	MET
1	G	71	LYS
1	G	109	SER
1	G	121	ASN
1	G	125	GLU
1	G	141	GLU
1	G	163	ARG
1	G	165	ARG
1	G	167	ARG
1	G	175	VAL
1	G	195	ARG
1	G	196	ARG
1	H	12	THR
1	H	34	THR
1	H	37	LEU
1	H	39	HIS
1	H	41	GLU
1	H	46	ASP
1	H	54	VAL
1	H	69	MET
1	H	71	LYS
1	H	109	SER
1	H	121	ASN
1	H	125	GLU
1	H	163	ARG
1	H	165	ARG
1	H	166	GLN
1	H	175	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (20)

such sidechains are listed below:

Mol	Chain	Res	Type
1	A	117	ASN
1	A	121	ASN
1	B	117	ASN
1	B	121	ASN
1	C	61	ASN
1	C	117	ASN
1	C	121	ASN
1	D	61	ASN
1	D	117	ASN
1	D	121	ASN
1	E	117	ASN
1	E	121	ASN
1	F	117	ASN
1	F	121	ASN
1	F	194	GLN
1	G	117	ASN
1	G	121	ASN
1	H	117	ASN
1	H	121	ASN
1	H	130	ASN

### 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](#)

16 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
2	SO4	A	1300	-	4,4,4	0.13	0	6,6,6	0.44	0
3	THM	E	1404	-	15,18,18	1.29	2 (13%)	16,26,26	2.23	5 (31%)
2	SO4	C	1306	-	4,4,4	0.29	0	6,6,6	0.48	0
2	SO4	D	1307	-	4,4,4	0.14	0	6,6,6	0.39	0
3	THM	A	1400	-	15,18,18	1.14	2 (13%)	16,26,26	2.42	4 (25%)
2	SO4	F	1305	-	4,4,4	0.18	0	6,6,6	0.33	0
2	SO4	G	1302	-	4,4,4	0.23	0	6,6,6	0.29	0
2	SO4	B	1301	-	4,4,4	0.20	0	6,6,6	0.38	0
3	THM	C	1406	-	15,18,18	1.52	2 (13%)	16,26,26	2.33	6 (37%)
3	THM	D	1407	-	15,18,18	1.10	2 (13%)	16,26,26	2.09	3 (18%)
3	THM	B	1401	-	15,18,18	1.26	1 (6%)	16,26,26	2.57	7 (43%)
2	SO4	E	1304	-	4,4,4	0.13	0	6,6,6	0.55	0
2	SO4	H	1303	-	4,4,4	0.13	0	6,6,6	0.33	0
3	THM	G	1402	-	15,18,18	1.26	1 (6%)	16,26,26	2.31	3 (18%)
3	THM	F	1405	-	15,18,18	1.23	1 (6%)	16,26,26	2.14	4 (25%)
3	THM	H	1403	-	15,18,18	1.16	1 (6%)	16,26,26	2.16	4 (25%)

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	THM	E	1404	-	-	0/3/18/18	0/2/2/2
3	THM	A	1400	-	-	0/3/18/18	0/2/2/2
3	THM	C	1406	-	-	2/3/18/18	0/2/2/2
3	THM	D	1407	-	-	0/3/18/18	0/2/2/2
3	THM	B	1401	-	-	0/3/18/18	0/2/2/2
3	THM	G	1402	-	-	2/3/18/18	0/2/2/2
3	THM	F	1405	-	-	2/3/18/18	0/2/2/2
3	THM	H	1403	-	-	2/3/18/18	0/2/2/2

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	1406	THM	C4-C5	4.57	1.51	1.41
3	G	1402	THM	C4-C5	4.15	1.50	1.41
3	B	1401	THM	C4-C5	4.09	1.50	1.41
3	E	1404	THM	C4-C5	4.01	1.50	1.41
3	F	1405	THM	C4-C5	3.81	1.49	1.41
3	H	1403	THM	C4-C5	3.53	1.49	1.41
3	A	1400	THM	C4-C5	3.30	1.48	1.41
3	C	1406	THM	C2-N3	-3.13	1.32	1.38
3	D	1407	THM	C4-C5	2.90	1.47	1.41
3	D	1407	THM	C2-N3	-2.64	1.32	1.38
3	A	1400	THM	C2-N3	-2.34	1.33	1.38
3	E	1404	THM	C2-N3	-2.34	1.33	1.38

All (36) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	G	1402	THM	C4-N3-C2	7.51	121.49	115.14
3	B	1401	THM	C4-N3-C2	7.07	121.11	115.14
3	A	1400	THM	C4-N3-C2	6.77	120.86	115.14
3	H	1403	THM	C4-N3-C2	6.59	120.70	115.14
3	D	1407	THM	C4-N3-C2	6.15	120.33	115.14
3	F	1405	THM	C4-N3-C2	6.06	120.26	115.14
3	E	1404	THM	C4-N3-C2	5.82	120.06	115.14
3	C	1406	THM	C4-N3-C2	5.63	119.89	115.14
3	E	1404	THM	C5-C6-N1	-4.43	117.42	122.19
3	F	1405	THM	C5-C6-N1	-3.84	118.06	122.19
3	A	1400	THM	C5-C6-N1	-3.75	118.15	122.19
3	B	1401	THM	C2'-C3'-C4'	3.61	110.28	102.76
3	H	1403	THM	C5-C6-N1	-3.45	118.47	122.19
3	D	1407	THM	C5-C6-N1	-3.29	118.65	122.19
3	C	1406	THM	C5-C6-N1	-3.27	118.67	122.19
3	B	1401	THM	C2'-C1'-N1	3.15	121.53	114.27
3	C	1406	THM	C5M-C5-C6	3.06	125.14	118.68
3	F	1405	THM	C2'-C3'-C4'	3.00	109.02	102.76
3	A	1400	THM	C2'-C3'-C4'	2.93	108.86	102.76
3	B	1401	THM	O3'-C3'-C2'	-2.83	100.77	110.90
3	E	1404	THM	C2'-C3'-C4'	2.77	108.54	102.76
3	G	1402	THM	C5-C6-N1	-2.67	119.32	122.19
3	B	1401	THM	C5-C6-N1	-2.65	119.33	122.19
3	C	1406	THM	C2'-C3'-C4'	2.63	108.25	102.76
3	G	1402	THM	C2'-C3'-C4'	2.63	108.24	102.76
3	C	1406	THM	C5'-C4'-C3'	-2.61	108.21	114.81
3	C	1406	THM	O5'-C5'-C4'	-2.59	102.41	111.29

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	H	1403	THM	C2'-C3'-C4'	2.50	107.97	102.76
3	D	1407	THM	C2'-C3'-C4'	2.48	107.92	102.76
3	H	1403	THM	O4'-C1'-C2'	2.34	110.68	106.25
3	B	1401	THM	O4'-C1'-C2'	2.27	110.53	106.25
3	E	1404	THM	O4'-C4'-C5'	2.17	113.91	109.21
3	F	1405	THM	O3'-C3'-C2'	-2.15	103.22	110.90
3	B	1401	THM	C6-N1-C1'	2.13	124.03	119.24
3	A	1400	THM	C2'-C1'-N1	2.12	119.16	114.27
3	E	1404	THM	C2'-C1'-N1	2.06	119.02	114.27

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	1406	THM	C3'-C4'-C5'-O5'
3	C	1406	THM	O4'-C4'-C5'-O5'
3	G	1402	THM	O4'-C4'-C5'-O5'
3	G	1402	THM	C3'-C4'-C5'-O5'
3	F	1405	THM	O4'-C4'-C5'-O5'
3	F	1405	THM	C3'-C4'-C5'-O5'
3	H	1403	THM	O4'-C4'-C5'-O5'
3	H	1403	THM	C3'-C4'-C5'-O5'

There are no ring outliers.

5 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1300	SO4	1	0
3	E	1404	THM	1	0
2	G	1302	SO4	1	0
3	C	1406	THM	1	0
3	F	1405	THM	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

### 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	190/230 (82%)	-0.05	1 (0%) 91 81	16, 36, 59, 69	0
1	B	185/230 (80%)	-0.23	0 100 100	16, 35, 58, 69	0
1	C	191/230 (83%)	-0.08	0 100 100	16, 36, 59, 69	0
1	D	185/230 (80%)	-0.10	2 (1%) 80 64	16, 36, 59, 69	0
1	E	190/230 (82%)	-0.21	3 (1%) 72 51	16, 36, 59, 69	0
1	F	187/230 (81%)	-0.28	3 (1%) 72 51	16, 36, 60, 69	0
1	G	189/230 (82%)	-0.16	0 100 100	16, 36, 60, 69	0
1	H	185/230 (80%)	-0.14	3 (1%) 72 51	16, 36, 58, 69	0
All	All	1502/1840 (81%)	-0.16	12 (0%) 86 72	16, 36, 59, 69	0

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	208	ASP	3.1
1	D	208	ASP	2.8
1	E	208	ASP	2.8
1	E	201	LYS	2.6
1	E	71	LYS	2.6
1	F	166	GLN	2.5
1	H	165	ARG	2.5
1	H	166	GLN	2.3
1	H	174	CYS	2.2
1	D	71	LYS	2.2
1	F	208	ASP	2.2
1	F	160	ALA	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	THM	H	1403	17/17	0.95	0.18	39,41,44,45	0
3	THM	C	1406	17/17	0.96	0.17	2,3,4,5	0
3	THM	E	1404	17/17	0.96	0.15	20,25,38,39	0
3	THM	G	1402	17/17	0.96	0.16	16,21,30,31	0
2	SO4	D	1307	5/5	0.96	0.12	39,41,41,42	0
3	THM	F	1405	17/17	0.97	0.14	24,25,31,32	0
3	THM	D	1407	17/17	0.97	0.18	20,24,28,28	0
3	THM	B	1401	17/17	0.97	0.14	7,11,14,15	0
2	SO4	G	1302	5/5	0.98	0.07	38,38,39,40	0
2	SO4	H	1303	5/5	0.98	0.08	49,49,50,50	0
3	THM	A	1400	17/17	0.98	0.17	3,6,14,15	0
2	SO4	C	1306	5/5	0.98	0.12	5,5,7,12	0
2	SO4	A	1300	5/5	0.98	0.10	9,11,14,14	0
2	SO4	E	1304	5/5	0.99	0.11	22,23,24,24	0
2	SO4	F	1305	5/5	0.99	0.09	27,28,29,29	0
2	SO4	B	1301	5/5	0.99	0.13	18,18,19,19	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.