



# Full wwPDB X-ray Structure Validation Report

Nov 23, 2014 – 11:25 AM EST

PDB ID : 4R5L  
Title : Crystal structure of the DnaK C-terminus (Dnak-SBD-C)  
Authors : Leu, J.I.; Zhang, P.; Murphy, M.E.; Marmorstein, R.; George, D.L.  
Deposited on : 2014-08-21  
Resolution : 2.97 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

---

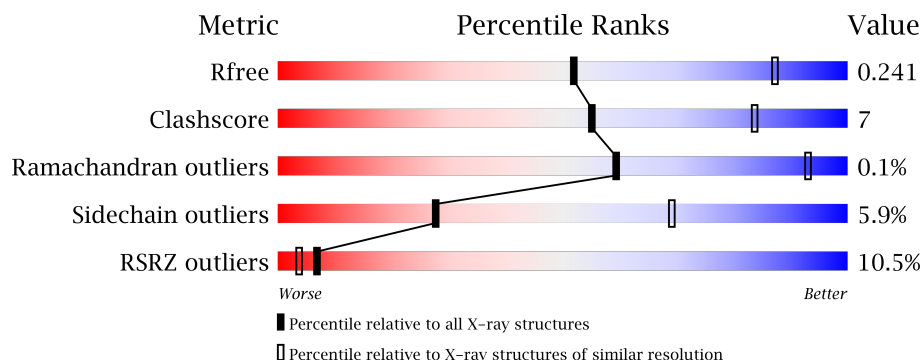
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.16 November 2013  
Xtriage (Phenix) : dev-1439  
EDS : stable24195  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.1.3  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable24195

# 1 Overall quality at a glance

The reported resolution of this entry is 2.97 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))
$R_{free}$	66092	1468 (3.00-2.96)
Clashscore	79885	1894 (3.00-2.96)
Ramachandran outliers	78287	1826 (3.00-2.96)
Sidechain outliers	78261	1829 (3.00-2.96)
RSRZ outliers	66119	1469 (3.00-2.96)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

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

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	SO4	A	702	-	X
2	SO4	B	701	-	X
2	SO4	B	703	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 6651 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Chaperone protein DnaK.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	222	Total	C	N	O	S	0	0	0
			1660	1021	297	336	6			
1	B	226	Total	C	N	O	S	0	0	0
			1583	976	288	314	5			
1	C	223	Total	C	N	O	S	0	0	0
			1624	997	292	330	5			
1	D	227	Total	C	N	O	S	0	0	0
			1696	1040	300	350	6			

There are 44 discrepancies between the modelled and reference sequences:

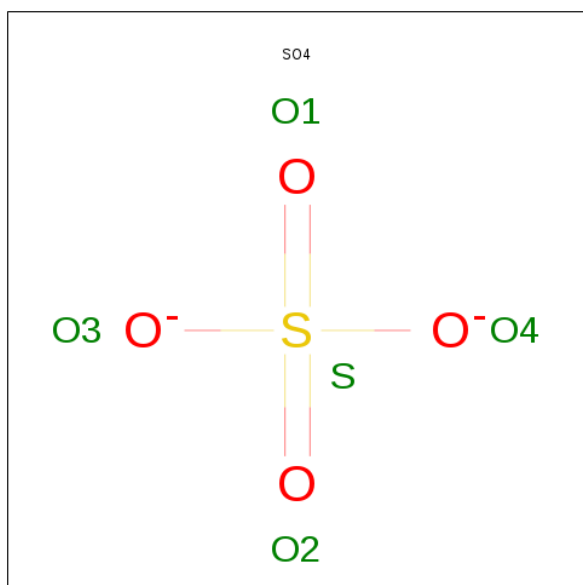
Chain	Residue	Modelled	Actual	Comment	Reference
A	378	MET	-	EXPRESSION TAG	UNP P0A6Y8
A	379	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	380	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	381	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	382	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	383	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	384	HIS	-	EXPRESSION TAG	UNP P0A6Y8
A	385	ILE	-	EXPRESSION TAG	UNP P0A6Y8
A	386	GLU	-	EXPRESSION TAG	UNP P0A6Y8
A	387	GLY	-	EXPRESSION TAG	UNP P0A6Y8
A	388	ARG	-	EXPRESSION TAG	UNP P0A6Y8
B	378	MET	-	EXPRESSION TAG	UNP P0A6Y8
B	379	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	380	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	381	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	382	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	383	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	384	HIS	-	EXPRESSION TAG	UNP P0A6Y8
B	385	ILE	-	EXPRESSION TAG	UNP P0A6Y8
B	386	GLU	-	EXPRESSION TAG	UNP P0A6Y8
B	387	GLY	-	EXPRESSION TAG	UNP P0A6Y8

*Continued on next page...*

Continued from previous page...

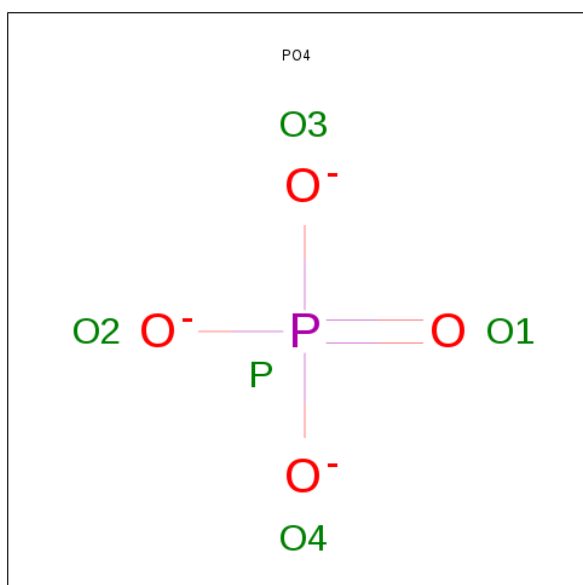
Chain	Residue	Modelled	Actual	Comment	Reference
B	388	ARG	-	EXPRESSION TAG	UNP P0A6Y8
C	378	MET	-	EXPRESSION TAG	UNP P0A6Y8
C	379	HIS	-	EXPRESSION TAG	UNP P0A6Y8
C	380	HIS	-	EXPRESSION TAG	UNP P0A6Y8
C	381	HIS	-	EXPRESSION TAG	UNP P0A6Y8
C	382	HIS	-	EXPRESSION TAG	UNP P0A6Y8
C	383	HIS	-	EXPRESSION TAG	UNP P0A6Y8
C	384	HIS	-	EXPRESSION TAG	UNP P0A6Y8
C	385	ILE	-	EXPRESSION TAG	UNP P0A6Y8
C	386	GLU	-	EXPRESSION TAG	UNP P0A6Y8
C	387	GLY	-	EXPRESSION TAG	UNP P0A6Y8
C	388	ARG	-	EXPRESSION TAG	UNP P0A6Y8
D	378	MET	-	EXPRESSION TAG	UNP P0A6Y8
D	379	HIS	-	EXPRESSION TAG	UNP P0A6Y8
D	380	HIS	-	EXPRESSION TAG	UNP P0A6Y8
D	381	HIS	-	EXPRESSION TAG	UNP P0A6Y8
D	382	HIS	-	EXPRESSION TAG	UNP P0A6Y8
D	383	HIS	-	EXPRESSION TAG	UNP P0A6Y8
D	384	HIS	-	EXPRESSION TAG	UNP P0A6Y8
D	385	ILE	-	EXPRESSION TAG	UNP P0A6Y8
D	386	GLU	-	EXPRESSION TAG	UNP P0A6Y8
D	387	GLY	-	EXPRESSION TAG	UNP P0A6Y8
D	388	ARG	-	EXPRESSION TAG	UNP P0A6Y8

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



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

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O P 5 4 1	0	0
3	A	1	Total O P 5 4 1	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	C	1	Total	O	P	0	0
			5	4	1		

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	1	Total	Ca	0	0
			1	1		

- Molecule 5 is water.

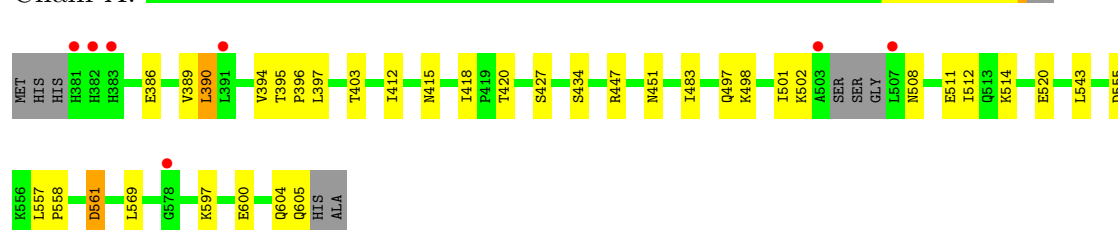
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	10	Total	O	0	0
			10	10		
5	B	8	Total	O	0	0
			8	8		
5	C	1	Total	O	0	0
			1	1		
5	D	3	Total	O	0	0
			3	3		

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

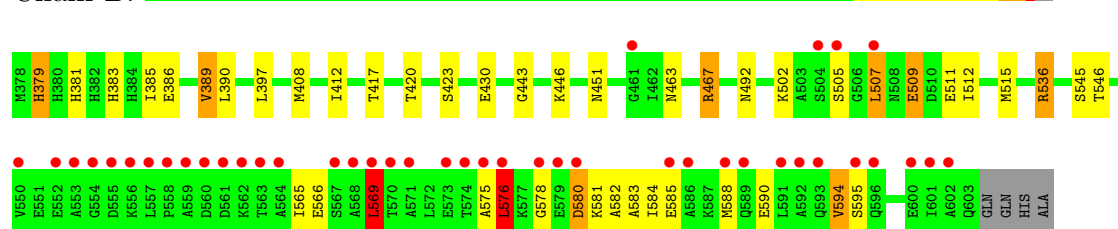
#### • Molecule 1: Chaperone protein DnaK

Chain A:



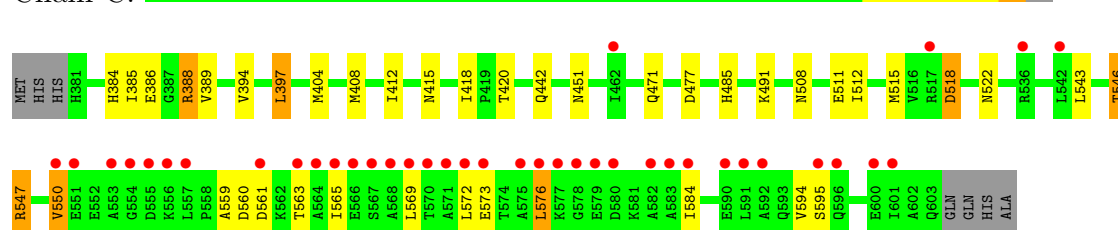
#### • Molecule 1: Chaperone protein DnaK

Chain B:



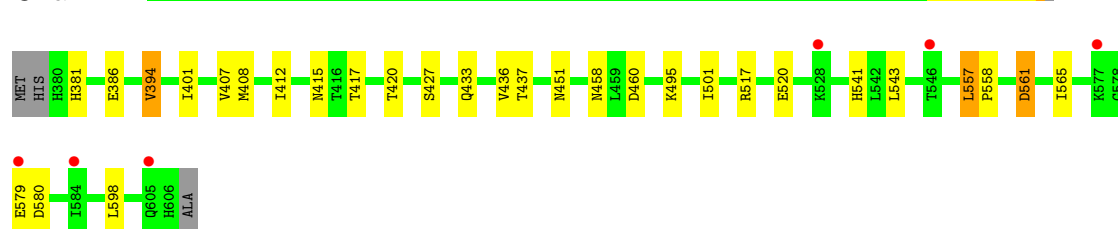
#### • Molecule 1: Chaperone protein DnaK

Chain C:



#### • Molecule 1: Chaperone protein DnaK

Chain D:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	92.31Å 99.91Å 133.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.90 – 2.97 43.93 – 2.96	Depositor EDS
% Data completeness (in resolution range)	99.9 (41.90-2.97) 99.6 (43.93-2.96)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.63 (at 2.96Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, $R_{free}$	0.205 , 0.243 0.205 , 0.241	Depositor DCC
$R_{free}$ test set	1337 reflections (5.12%)	DCC
Wilson B-factor (Å <sup>2</sup> )	53.8	Xtriage
Anisotropy	0.613	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 63.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 26200 reflections	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	6651	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.82% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.



## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: PO4, CA, 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	0.25	0/1676	0.45	0/2265
1	B	0.27	0/1602	0.59	2/2179 (0.1%)
1	C	0.26	0/1642	0.47	0/2228
1	D	0.25	0/1713	0.46	0/2314
All	All	0.26	0/6633	0.49	2/8986 (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 (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	569	LEU	CA-CB-CG	6.00	129.11	115.30
1	B	580	ASP	N-CA-CB	-5.02	101.57	110.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	576	LEU	Peptide

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1660	0	1631	21	1
1	B	1583	0	1474	34	0
1	C	1624	0	1553	25	1
1	D	1696	0	1659	17	0
2	A	15	0	0	1	0
2	B	15	0	0	0	0
2	C	10	0	0	0	0
2	D	10	0	0	0	0
3	A	10	0	0	0	0
3	C	5	0	0	0	0
4	C	1	0	0	0	0
5	A	10	0	0	1	0
5	B	8	0	0	1	0
5	C	1	0	0	0	0
5	D	3	0	0	0	0
All	All	6651	0	6317	87	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 7.

All (87) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:511:GLU:HA	1:A:514:LYS:HB3	1.57	0.84
1:B:581:LYS:HG3	1:B:582:ALA:H	1.43	0.83
1:C:547:ARG:HG2	1:C:569:LEU:HD13	1.67	0.75
1:D:394:VAL:HG13	1:D:415:ASN:HA	1.74	0.70
1:B:580:ASP:CG	1:B:581:LYS:HB2	2.14	0.68
1:B:509:GLU:HA	1:B:512:ILE:HD13	1.75	0.67
1:C:386:GLU:OE2	1:D:541:HIS:NE2	2.29	0.66
1:D:460:ASP:OD1	1:D:495:LYS:NZ	2.29	0.66
1:C:543:LEU:HD21	1:C:573:GLU:HG2	1.80	0.63
1:A:520:GLU:OE2	1:D:517:ARG:NH1	2.31	0.62
1:A:558:PRO:HB2	1:A:561:ASP:HB2	1.83	0.61
1:C:384:HIS:O	1:D:433:GLN:NE2	2.33	0.60

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:581:LYS:HG3	1:B:582:ALA:N	2.14	0.59
1:C:404:MET:HE1	1:D:386:GLU:HG3	1.85	0.59
1:C:518:ASP:O	1:C:522:ASN:ND2	2.30	0.59
1:B:397:LEU:HD12	1:B:443:GLY:HA2	1.85	0.58
1:C:394:VAL:HG13	1:C:415:ASN:HA	1.84	0.58
1:B:446:LYS:NZ	5:B:808:HOH:O	2.36	0.58
1:C:508:ASN:N	1:C:511:GLU:OE1	2.27	0.58
1:B:581:LYS:CG	1:B:582:ALA:H	2.14	0.58
1:A:394:VAL:HG13	1:A:415:ASN:HA	1.86	0.58
1:B:581:LYS:HB3	1:B:583:ALA:H	1.69	0.57
1:A:386:GLU:OE2	1:B:467:ARG:NH1	2.38	0.56
1:A:447:ARG:O	1:A:451:ASN:ND2	2.39	0.56
1:A:427:SER:N	1:B:386:GLU:O	2.26	0.56
1:B:512:ILE:HD12	1:B:512:ILE:H	1.70	0.55
1:A:555:ASP:N	1:A:555:ASP:OD1	2.38	0.54
1:C:408:MET:SD	1:C:451:ASN:ND2	2.81	0.54
1:B:507:LEU:HB3	1:B:512:ILE:HD11	1.89	0.53
1:A:395:THR:HG23	1:A:418:ILE:HD11	1.91	0.53
1:A:403:THR:HG22	1:B:385:ILE:HG22	1.93	0.51
1:B:408:MET:SD	1:B:451:ASN:ND2	2.83	0.51
1:C:559:ALA:O	1:C:563:THR:HG23	2.10	0.51
1:A:434:SER:OG	2:A:703:SO4:O2	2.21	0.51
1:B:580:ASP:OD2	1:B:581:LYS:HB2	2.11	0.50
1:B:582:ALA:HA	1:B:585:GLU:HB2	1.94	0.50
1:C:560:ASP:O	1:C:563:THR:OG1	2.30	0.49
1:D:437:THR:OG1	1:D:458:ASN:OD1	2.28	0.49
1:A:396:PRO:HB2	1:A:397:LEU:HD22	1.93	0.49
1:A:397:LEU:HD21	1:A:512:ILE:HG12	1.95	0.49
1:B:507:LEU:CB	1:B:512:ILE:HD11	2.42	0.49
1:C:412:ILE:HD13	1:C:420:THR:HG23	1.94	0.48
1:C:389:VAL:HG22	1:D:427:SER:HB2	1.96	0.48
1:B:581:LYS:CG	1:B:582:ALA:N	2.75	0.48
1:B:575:ALA:HB1	1:B:584:ILE:HD12	1.96	0.48
1:B:463:ASN:ND2	1:B:492:ASN:OD1	2.47	0.48
1:D:412:ILE:HD13	1:D:420:THR:HG23	1.96	0.48
1:C:397:LEU:HD13	1:C:512:ILE:HG23	1.96	0.47
1:C:397:LEU:HD23	1:C:442:GLN:HG2	1.96	0.47
1:C:550:VAL:HG11	1:C:565:ILE:HG21	1.97	0.46
1:C:385:ILE:HD11	1:D:401:ILE:HD12	1.96	0.46
1:D:408:MET:SD	1:D:451:ASN:ND2	2.89	0.46
1:B:430:GLU:HA	1:B:467:ARG:HB2	1.97	0.46
1:B:412:ILE:HD13	1:B:420:THR:HG23	1.98	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:483:ILE:HD12	1:A:502:LYS:HD3	1.97	0.45
1:B:511:GLU:O	1:B:515:MET:HG3	2.17	0.45
1:A:412:ILE:HD13	1:A:420:THR:HG23	1.98	0.45
1:C:546:THR:O	1:C:550:VAL:HG13	2.15	0.45
1:C:477:ASP:HB3	1:C:485:HIS:HB2	1.98	0.45
1:B:536:ARG:HG3	1:B:576:LEU:HD22	1.99	0.45
1:C:511:GLU:O	1:C:515:MET:HG3	2.17	0.45
1:A:508:ASN:OD1	1:A:508:ASN:N	2.50	0.44
1:A:543:LEU:HD12	1:A:569:LEU:HD22	2.00	0.44
1:B:502:LYS:HE3	1:B:505:SER:HB2	1.98	0.44
1:A:557:LEU:HD12	1:A:558:PRO:HD2	2.00	0.44
1:A:561:ASP:OD2	1:A:597:LYS:HD2	2.18	0.44
1:B:536:ARG:NH1	1:B:578:GLY:O	2.50	0.43
1:C:385:ILE:HG22	1:D:436:VAL:HB	2.01	0.43
1:D:579:GLU:HG2	1:D:580:ASP:H	1.83	0.43
1:C:418:ILE:HG22	1:C:420:THR:HG22	1.99	0.43
1:C:565:ILE:O	1:C:569:LEU:HG	2.19	0.43
1:B:594:VAL:HG23	1:B:595:SER:H	1.83	0.43
1:D:565:ILE:HD13	1:D:598:LEU:HD22	2.00	0.43
1:B:507:LEU:HA	1:B:507:LEU:HD12	1.91	0.43
1:C:576:LEU:HD12	1:C:584:ILE:HD13	2.01	0.43
1:D:558:PRO:HG2	1:D:561:ASP:HB2	2.00	0.43
1:C:471:GLN:HB2	1:C:491:LYS:HB2	2.00	0.42
1:B:584:ILE:O	1:B:588:MET:HG3	2.19	0.42
5:A:809:HOH:O	1:B:379:HIS:HE1	2.01	0.42
1:B:389:VAL:CG1	1:B:417:THR:HG21	2.49	0.42
1:B:545:SER:OG	1:B:546:THR:N	2.52	0.42
1:B:566:GLU:O	1:B:569:LEU:HD22	2.20	0.42
1:A:497:GLN:OE1	1:A:498:LYS:N	2.52	0.41
1:D:557:LEU:HA	1:D:558:PRO:HD2	1.95	0.41
1:A:390:LEU:HG	1:A:390:LEU:H	1.45	0.40
1:B:565:ILE:O	1:B:569:LEU:HD13	2.21	0.40
1:D:557:LEU:HD23	1:D:558:PRO:HD2	2.04	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	Distance(Å)	Clash(Å)
1:A:600:GLU:OE1	1:C:388:ARG:NH2[2_764]	2.09	0.11

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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	218/230 (95%)	213 (98%)	5 (2%)	0	100	100
1	B	224/230 (97%)	206 (92%)	17 (8%)	1 (0%)	43	86
1	C	221/230 (96%)	211 (96%)	10 (4%)	0	100	100
1	D	225/230 (98%)	214 (95%)	11 (5%)	0	100	100
All	All	888/920 (96%)	844 (95%)	43 (5%)	1 (0%)	59	93

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	509	GLU

### 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	174/191 (91%)	168 (97%)	6 (3%)	49	87
1	B	149/191 (78%)	136 (91%)	13 (9%)	15	48
1	C	164/191 (86%)	153 (93%)	11 (7%)	23	63
1	D	179/191 (94%)	170 (95%)	9 (5%)	34	76
All	All	666/764 (87%)	627 (94%)	39 (6%)	28	69

All (39) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	389	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	390	LEU
1	A	501	ILE
1	A	561	ASP
1	A	604	GLN
1	A	605	GLN
1	B	379	HIS
1	B	381	HIS
1	B	383	HIS
1	B	389	VAL
1	B	390	LEU
1	B	423	SER
1	B	467	ARG
1	B	507	LEU
1	B	536	ARG
1	B	569	LEU
1	B	576	LEU
1	B	590	GLU
1	B	594	VAL
1	C	388	ARG
1	C	397	LEU
1	C	518	ASP
1	C	546	THR
1	C	547	ARG
1	C	550	VAL
1	C	561	ASP
1	C	572	LEU
1	C	576	LEU
1	C	594	VAL
1	C	595	SER
1	D	381	HIS
1	D	394	VAL
1	D	407	VAL
1	D	417	THR
1	D	501	ILE
1	D	520	GLU
1	D	543	LEU
1	D	557	LEU
1	D	561	ASP

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

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 14 ligands modelled in this entry, 1 is monoatomic - leaving 13 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	$\# Z  > 2$
2	SO4	A	701	-	4,4,4	0.26	0	6,6,6	0.08	0
2	SO4	A	702	-	4,4,4	0.23	0	6,6,6	0.09	0
2	SO4	A	703	-	4,4,4	0.26	0	6,6,6	0.11	0
3	PO4	A	704	-	4,4,4	0.33	0	6,6,6	0.29	0
3	PO4	A	705	-	4,4,4	0.34	0	6,6,6	0.28	0
2	SO4	B	701	-	4,4,4	0.23	0	6,6,6	0.09	0
2	SO4	B	702	-	4,4,4	0.25	0	6,6,6	0.08	0
2	SO4	B	703	-	4,4,4	0.25	0	6,6,6	0.05	0
2	SO4	C	701	-	4,4,4	0.26	0	6,6,6	0.08	0
2	SO4	C	702	-	4,4,4	0.27	0	6,6,6	0.09	0
3	PO4	C	703	-	4,4,4	0.34	0	6,6,6	0.28	0
2	SO4	D	701	-	4,4,4	0.28	0	6,6,6	0.12	0
2	SO4	D	702	-	4,4,4	0.26	0	6,6,6	0.10	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
2	SO4	A	701	-	-	0/0/0/0	0/0/0/0
2	SO4	A	702	-	-	0/0/0/0	0/0/0/0
2	SO4	A	703	-	-	0/0/0/0	0/0/0/0
3	PO4	A	704	-	-	0/0/0/0	0/0/0/0
3	PO4	A	705	-	-	0/0/0/0	0/0/0/0
2	SO4	B	701	-	-	0/0/0/0	0/0/0/0
2	SO4	B	702	-	-	0/0/0/0	0/0/0/0
2	SO4	B	703	-	-	0/0/0/0	0/0/0/0
2	SO4	C	701	-	-	0/0/0/0	0/0/0/0
2	SO4	C	702	-	-	0/0/0/0	0/0/0/0
3	PO4	C	703	-	-	0/0/0/0	0/0/0/0
2	SO4	D	701	-	-	0/0/0/0	0/0/0/0
2	SO4	D	702	-	-	0/0/0/0	0/0/0/0

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.

## 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	222/230 (96%)	0.14	7 (3%) 45 18	49, 73, 99, 115	0
1	B	226/230 (98%)	0.82	42 (18%) 2 1	51, 90, 144, 172	0
1	C	223/230 (96%)	0.78	39 (17%) 2 1	54, 86, 134, 151	0
1	D	227/230 (98%)	0.18	6 (2%) 53 21	48, 74, 103, 135	0
All	All	898/920 (97%)	0.48	94 (10%) 7 4	48, 77, 135, 172	0

All (94) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	553	ALA	11.4
1	B	592	ALA	6.7
1	B	552	GLU	6.5
1	C	568	ALA	6.2
1	C	564	ALA	6.0
1	B	588	MET	5.6
1	B	589	GLN	5.6
1	C	601	ILE	5.6
1	C	576	LEU	5.3
1	B	560	ASP	5.0
1	B	558	PRO	4.9
1	B	555	ASP	4.7
1	B	561	ASP	4.7
1	B	595	SER	4.6
1	A	381	HIS	4.5
1	B	570	THR	4.4
1	C	565	ILE	4.3
1	C	570	THR	4.2
1	B	554	GLY	4.1
1	C	595	SER	4.1
1	B	564	ALA	4.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	601	ILE	4.1
1	B	557	LEU	4.0
1	C	561	ASP	4.0
1	B	591	LEU	3.8
1	B	567	SER	3.8
1	B	586	ALA	3.8
1	C	592	ALA	3.8
1	B	593	GLN	3.8
1	C	578	GLY	3.7
1	B	578	GLY	3.7
1	C	554	GLY	3.7
1	C	591	LEU	3.7
1	C	596	GLN	3.7
1	C	575	ALA	3.6
1	C	567	SER	3.5
1	B	563	THR	3.4
1	B	550	VAL	3.3
1	B	573	GLU	3.3
1	C	553	ALA	3.3
1	C	580	ASP	3.2
1	B	569	LEU	3.2
1	B	559	ALA	3.2
1	B	574	THR	3.2
1	C	584	ILE	3.2
1	C	579	GLU	3.0
1	C	572	LEU	3.0
1	C	600	GLU	3.0
1	B	602	ALA	2.9
1	C	563	THR	2.9
1	D	605	GLN	2.9
1	C	551	GLU	2.9
1	C	583	ALA	2.8
1	B	575	ALA	2.8
1	B	556	LYS	2.7
1	C	573	GLU	2.6
1	A	503	ALA	2.6
1	B	585	GLU	2.6
1	C	582	ALA	2.6
1	C	550	VAL	2.6
1	C	569	LEU	2.6
1	B	505	SER	2.5
1	D	577	LYS	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	590	GLU	2.5
1	A	391	LEU	2.5
1	C	555	ASP	2.5
1	B	568	ALA	2.5
1	C	557	LEU	2.4
1	C	566	GLU	2.4
1	B	596	GLN	2.4
1	B	579	GLU	2.4
1	C	577	LYS	2.4
1	B	576	LEU	2.4
1	C	556	LYS	2.4
1	C	571	ALA	2.4
1	B	461	GLY	2.3
1	C	517	ARG	2.3
1	B	507	LEU	2.3
1	B	562	LYS	2.3
1	C	462	ILE	2.3
1	B	600	GLU	2.2
1	B	504	SER	2.2
1	A	507	LEU	2.2
1	C	542	LEU	2.2
1	D	528	LYS	2.2
1	D	579	GLU	2.2
1	A	578	GLY	2.1
1	A	382	HIS	2.1
1	D	546	THR	2.1
1	B	571	ALA	2.1
1	C	536	ARG	2.1
1	B	580	ASP	2.0
1	D	584	ILE	2.0
1	A	383	HIS	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	SO4	A	702	5/5	0.37	10.75	115,120,131,149	0
2	SO4	B	703	5/5	0.35	7.13	127,157,171,177	0
2	SO4	B	701	5/5	0.28	3.04	107,109,129,153	0
3	PO4	A	704	5/5	0.23	1.83	80,86,104,134	0
3	PO4	C	703	5/5	0.23	1.16	85,89,109,129	0
2	SO4	D	702	5/5	0.30	1.04	130,135,146,148	0
3	PO4	A	705	5/5	0.26	-0.12	113,114,147,149	0
2	SO4	A	701	5/5	0.23	-0.48	84,98,118,120	0
2	SO4	A	703	5/5	0.17	-1.13	108,120,141,153	0
2	SO4	C	701	5/5	0.26	-1.38	120,125,141,141	0
2	SO4	B	702	5/5	0.23	-1.57	116,126,143,149	0
2	SO4	C	702	5/5	0.14	-1.69	107,115,129,138	0
4	CA	C	704	1/1	0.11	-2.00	114,114,114,114	0
2	SO4	D	701	5/5	0.26	-3.52	93,105,115,117	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.