



# wwPDB X-ray Structure Validation Summary Report i

Feb 27, 2014 – 06:40 AM GMT

PDB ID : 3DPQ  
Title : Crystal structure of the substrate binding domain of E. coli DnaK in complex with a long pyrrolic acid-derived inhibitor peptide (form B)  
Authors : Roujeinikova, A.  
Deposited on : 2008-07-09  
Resolution : 2.60 Å (reported)

This is a wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

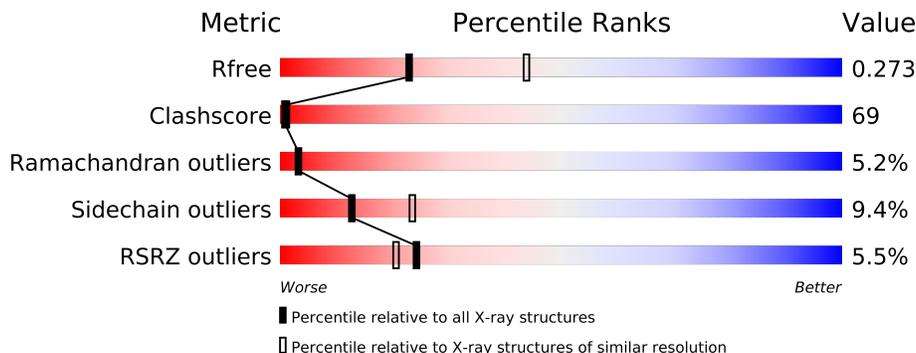
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.60 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	66092	1718 (2.60-2.60)
Clashscore	79885	2154 (2.60-2.60)
Ramachandran outliers	78287	2113 (2.60-2.60)
Sidechain outliers	78261	2113 (2.60-2.60)
RSRZ outliers	66119	1718 (2.60-2.60)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	219	
1	B	219	
1	E	219	
1	F	219	
2	C	20	
2	D	20	
2	G	20	
2	H	20	

## 2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 8427 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.

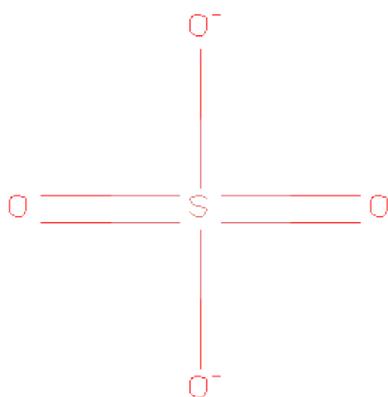
- Molecule 1 is a protein called Chaperone protein dnaK.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	212	Total 1644	C 1012	N 288	O 338	S 6	9	5	0
1	B	212	Total 1654	C 1018	N 293	O 337	S 6	9	5	0
1	E	203	Total 1626	C 999	N 292	O 329	S 6	0	9	0
1	F	214	Total 1647	C 1015	N 287	O 339	S 6	0	4	0

- Molecule 2 is a protein called inhibitor peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	C	12	Total 105	C 72	N 19	O 14	0	0	0
2	D	9	Total 80	C 56	N 13	O 11	0	0	0
2	G	12	Total 109	C 74	N 21	O 14	0	1	0
2	H	9	Total 91	C 62	N 17	O 12	11	1	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	E	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	352	Total	O	0	0
			352	352		
4	B	308	Total	O	0	0
			308	308		
4	E	334	Total	O	0	0
			334	334		
4	F	334	Total	O	0	0
			334	334		
4	C	27	Total	O	0	0
			27	27		
4	D	36	Total	O	0	0
			36	36		
4	G	50	Total	O	0	0
			50	50		
4	H	25	Total	O	0	0
			25	25		

### 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.

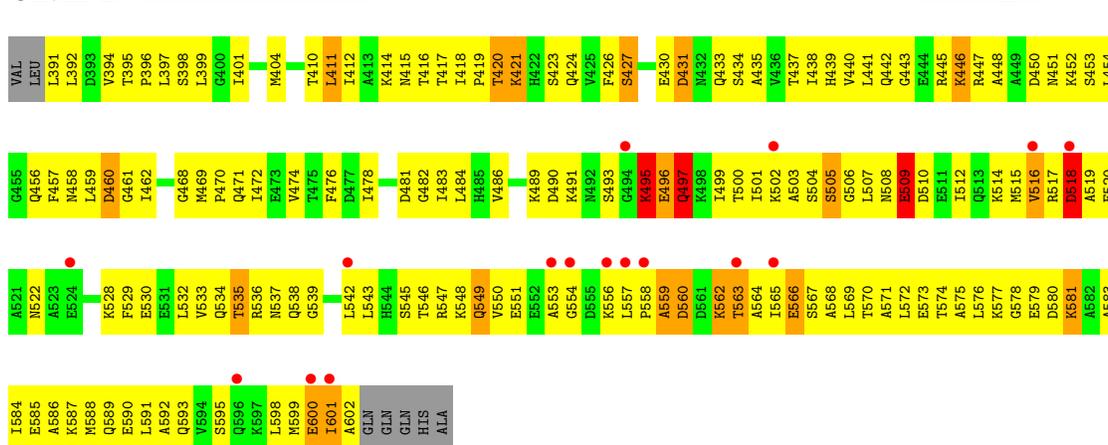
- Molecule 1: Chaperone protein dnaK

#### Chain A:



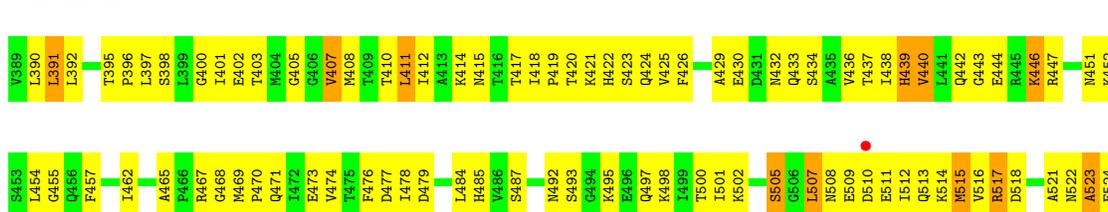
- Molecule 1: Chaperone protein dnaK

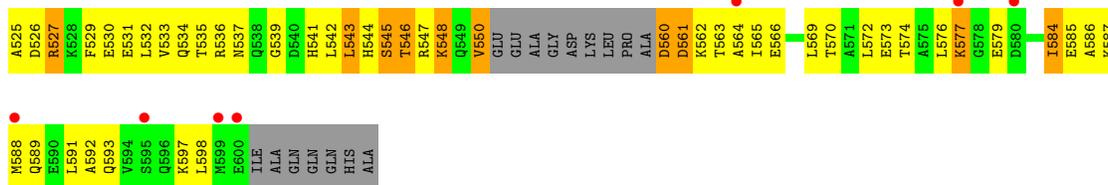
#### Chain B:



- Molecule 1: Chaperone protein dnaK

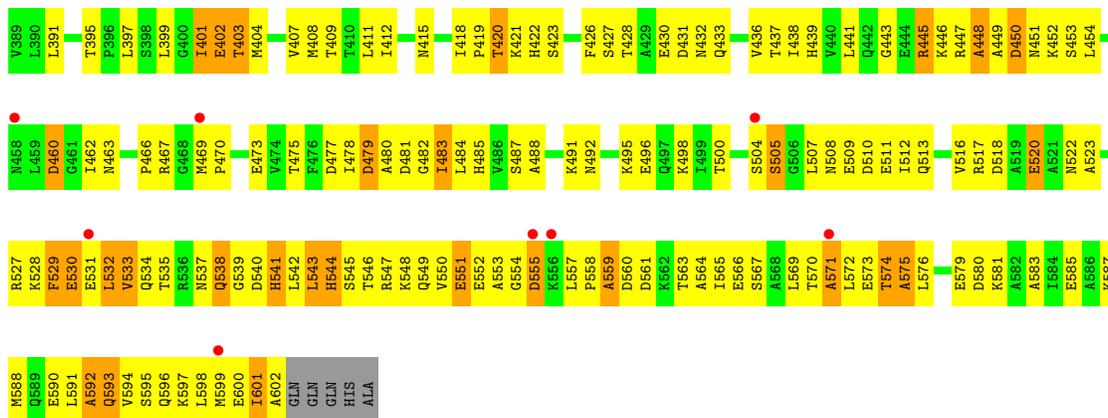
#### Chain E:





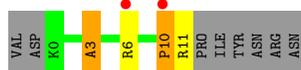
● Molecule 1: Chaperone protein dnaK

Chain F:



● Molecule 2: inhibitor peptide

Chain C:



● Molecule 2: inhibitor peptide

Chain D:



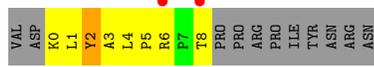
● Molecule 2: inhibitor peptide

Chain G:



● Molecule 2: inhibitor peptide

Chain H:



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	77.71Å 91.65Å 154.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 2.60 19.87 – 1.95	Depositor EDS
% Data completeness (in resolution range)	82.0 (8.00-2.60) 85.4 (19.87-1.95)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.60 (at 1.94Å)	Xtrriage
Refinement program	REFMAC 5.4.0073	Depositor
R, $R_{free}$	0.254 , 0.319 0.267 , 0.273	Depositor DCC
$R_{free}$ test set	1427 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	24.3	Xtrriage
Anisotropy	0.554	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.44 , 86.9	EDS
Estimated twinning fraction	No twinning to report.	Xtrriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Outliers	1 of 69664 reflections (0.001%)	Xtrriage
$F_o, F_c$ correlation	0.85	EDS
Total number of atoms	8427	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 22.01 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 6.2189e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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

## 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: SO4, ALC

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	0.66	0/1658	0.98	2/2236 (0.1%)
1	B	0.58	0/1668	1.08	4/2243 (0.2%)
1	E	0.56	0/1638	0.93	0/2202
1	F	0.56	0/1673	0.93	0/2253
2	C	0.49	0/97	0.80	0/131
2	D	0.65	0/70	0.87	0/93
2	G	0.50	0/104	1.08	1/142 (0.7%)
2	H	1.07	2/81 (2.5%)	1.06	0/108
All	All	0.60	2/6989 (0.0%)	0.98	7/9408 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	1	0
1	B	0	2
2	C	0	1
2	G	0	1
2	H	0	1
All	All	1	5

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	6[A]	ARG	C-N	5.91	1.45	1.34
2	H	6[B]	ARG	C-N	5.91	1.45	1.34

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	518	ASP	CB-CG-OD2	19.18	135.56	118.30
1	B	518	ASP	CB-CG-OD1	-16.14	103.77	118.30
2	G	9	PRO	C-N-CD	-7.70	103.67	120.60
1	A	547	ARG	CA-C-N	-6.15	103.68	117.20
1	B	496	GLU	N-CA-C	5.47	125.77	111.00

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	548	LYS	CA

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	495	LYS	Peptide
1	B	496	GLU	Peptide
2	C	3	ALC	Mainchain
2	G	2	TYR	Mainchain
2	H	2	TYR	Mainchain

## 5.2 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 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	1644	0	1651	207	0
1	B	1654	0	1679	254	1
1	E	1626	0	1649	242	1
1	F	1647	0	1673	243	0
2	C	105	0	121	7	0
2	D	80	0	94	8	0
2	G	109	0	123	15	0
2	H	91	0	106	6	0
3	E	5	0	0	0	0
4	A	352	0	0	10	0
4	B	308	0	0	12	0
4	C	27	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	36	0	0	2	0
4	E	334	0	0	12	0
4	F	334	0	0	6	0
4	G	50	0	0	1	0
4	H	25	0	0	0	0
All	All	8427	0	7096	964	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 69.

The worst 5 of 964 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:E:438[B]:ILE:HG21	1:E:457:PHE:CE1	1.32	1.63
1:E:438[B]:ILE:CG2	1:E:457:PHE:CE1	1.94	1.45
1:E:438[B]:ILE:HG21	1:E:457:PHE:CZ	1.53	1.41
1:E:434:SER:O	1:E:462:ILE:HG22	1.26	1.30
1:B:536:ARG:NH2	1:B:578:GLY:O	1.73	1.20

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:B:503:ALA:CB	1:E:390:LEU:CD1[3_545]	2.14	0.06

## 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	215/219 (98%)	183 (85%)	23 (11%)	9 (4%)	4	5
1	B	216/219 (99%)	177 (82%)	29 (13%)	10 (5%)	4	4
1	E	208/219 (95%)	186 (89%)	18 (9%)	4 (2%)	12	23
1	F	216/219 (99%)	173 (80%)	23 (11%)	20 (9%)	1	1
2	C	9/20 (45%)	7 (78%)	1 (11%)	1 (11%)	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	6/20 (30%)	6 (100%)	0	0	100	100
2	G	10/20 (50%)	7 (70%)	3 (30%)	0	100	100
2	H	7/20 (35%)	5 (71%)	1 (14%)	1 (14%)	0	0
All	All	887/956 (93%)	744 (84%)	98 (11%)	45 (5%)	3	3

5 of 45 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	548	LYS
1	A	549	GLN
1	A	596	GLN
1	B	600	GLU
1	B	601	ILE

### 5.3.2 Protein sidechains ❶

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	179/181 (99%)	165 (92%)	14 (8%)	18	34
1	B	180/181 (99%)	160 (89%)	20 (11%)	9	16
1	E	179/181 (99%)	157 (88%)	22 (12%)	7	13
1	F	181/181 (100%)	166 (92%)	15 (8%)	16	30
2	C	11/19 (58%)	10 (91%)	1 (9%)	14	25
2	D	8/19 (42%)	8 (100%)	0	100	100
2	G	11/19 (58%)	11 (100%)	0	100	100
2	H	9/19 (47%)	8 (89%)	1 (11%)	9	16
All	All	758/800 (95%)	685 (90%)	73 (10%)	13	22

5 of 73 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	566	GLU
1	E	446[A]	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	F	520	GLU
1	E	411[B]	LEU
1	E	471	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	549	GLN
1	E	439	HIS
1	F	534	GLN
1	E	433	GLN
1	E	458	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains 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
2	ALC	C	3	2	11,11,12	4.83	1 (9%)	11,13,15	2.10	3 (27%)
2	ALC	D	3	2	11,11,12	5.37	2 (18%)	11,13,15	1.18	2 (18%)
2	ALC	G	3	2	11,11,12	4.56	1 (9%)	11,13,15	2.68	2 (18%)
2	ALC	H	3	2	11,11,12	4.76	1 (9%)	11,13,15	1.74	2 (18%)

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	ALC	C	3	2	-	0/4/14/16	0/1/1/1
2	ALC	D	3	2	-	0/4/14/16	0/1/1/1
2	ALC	G	3	2	-	0/4/14/16	0/1/1/1
2	ALC	H	3	2	-	0/4/14/16	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	3	ALC	O-C	17.61	1.23	1.11
2	C	3	ALC	O-C	15.85	1.22	1.11
2	H	3	ALC	O-C	15.64	1.22	1.11
2	G	3	ALC	O-C	15.03	1.21	1.11
2	D	3	ALC	CA-C	2.26	1.52	1.48

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	G	3	ALC	C-CA-N	-7.95	105.89	113.83
2	C	3	ALC	C-CA-N	-4.99	108.85	113.83
2	H	3	ALC	C-CA-N	-3.61	110.22	113.83
2	H	3	ALC	CG-CB-CA	-3.38	109.98	115.12
2	C	3	ALC	CE2-CD2-CG	-3.25	106.89	112.23

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected

value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
3	SO4	E	1001	-	4,4,4	0.15	0	6,6,6	0.44	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	SO4	E	1001	-	-	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 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	A	212/219 (96%)	0.04	9 (4%) 35 31	4, 17, 38, 53	7 (3%)
1	B	212/219 (96%)	0.39	16 (7%) 14 11	5, 27, 44, 52	7 (3%)
1	E	203/219 (92%)	0.22	8 (3%) 37 33	10, 22, 40, 62	7 (3%)
1	F	214/219 (97%)	0.20	8 (3%) 39 35	6, 25, 44, 51	5 (2%)
2	C	12/20 (60%)	1.44	2 (16%) 2 1	18, 37, 47, 54	0
2	D	9/20 (45%)	0.24	1 (11%) 6 4	6, 16, 22, 24	1 (11%)
2	G	12/20 (60%)	1.20	3 (25%) 1 1	20, 31, 54, 61	0
2	H	9/20 (45%)	0.86	2 (22%) 1 1	17, 21, 35, 38	1 (11%)
All	All	883/956 (92%)	0.25	49 (5%) 24 20	4, 24, 43, 62	28 (3%)

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	G	9	PRO	5.3
2	C	10	PRO	5.1
1	B	557	LEU	4.7
1	E	599	MET	4.3
1	B	494	GLY	4.0

### 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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	ALC	G	3	11/12	0.24	2.67	26,28,35,36	0
2	ALC	C	3	11/12	0.19	0.21	20,25,29,30	0
2	ALC	H	3	11/12	0.18	0.14	19,27,31,32	0
2	ALC	D	3	11/12	0.15	-0.24	2,3,9,12	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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	SO4	E	1001	5/5	0.15	-1.34	51,53,54,55	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.