



wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 10:25 pm BST

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 X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at 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	:	2.11
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)	:	2.11

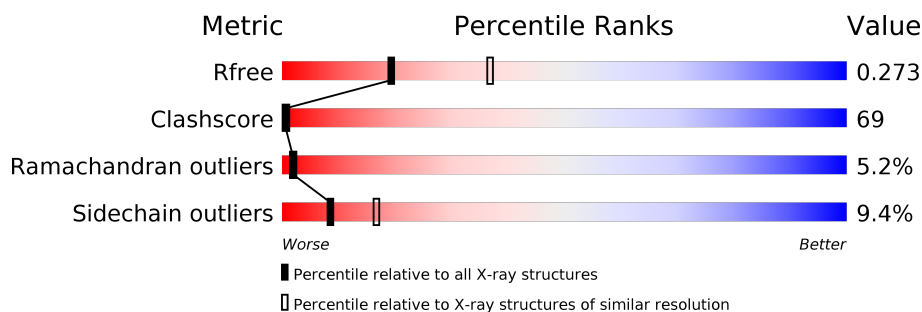
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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}	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (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 ≥ 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	219	41% (green), 48% (yellow), 8% (orange), 3% (red), 2% (grey)
1	B	219	25% (green), 61% (yellow), 9% (orange), 3% (red), 2% (grey)
1	E	219	30% (green), 53% (yellow), 10% (orange), 7% (red), 2% (grey)
1	F	219	29% (green), 55% (yellow), 13% (orange), 3% (red), 2% (grey)
2	C	20	40% (green), 10% (yellow), 10% (orange), 40% (grey)
2	D	20	30% (green), 15% (yellow), 55% (grey)
2	G	20	15% (green), 35% (yellow), 10% (orange), 40% (grey)

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	20	 A horizontal bar chart showing the quality of chain H. The bar is divided into four segments: a green segment (5%), a yellow segment (35%), an orange segment (5%), and a grey segment (55%). The percentages are labeled below the bar.

2 Entry composition

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

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

- Molecule 2 is a protein called inhibitor peptide.

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

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



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

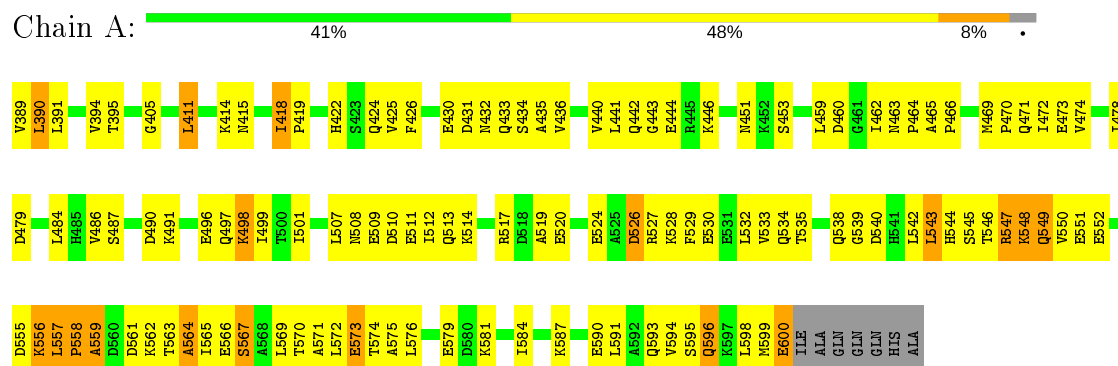
- Molecule 4 is water.

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		

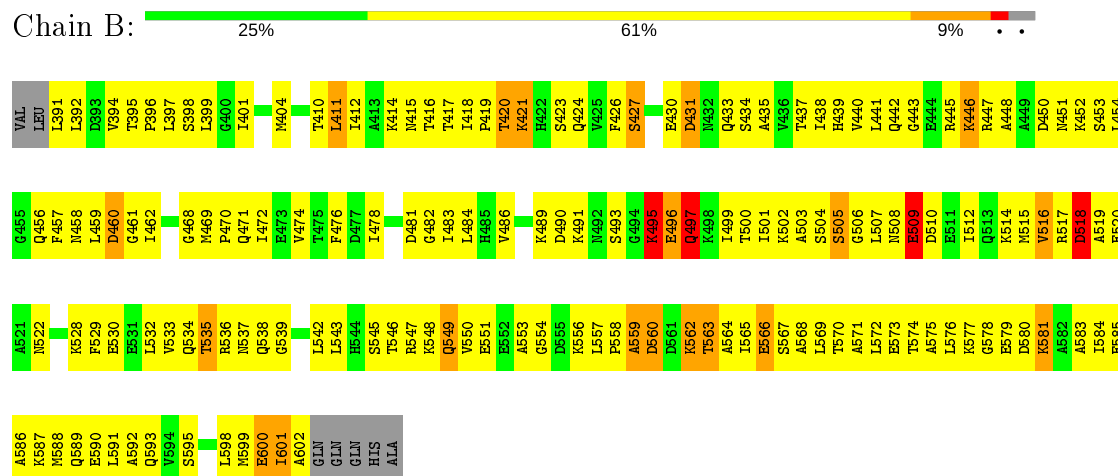
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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.

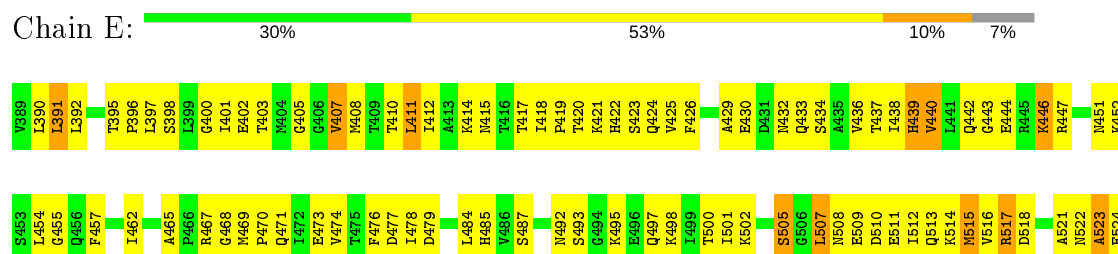
• Molecule 1: Chaperone protein dnaK

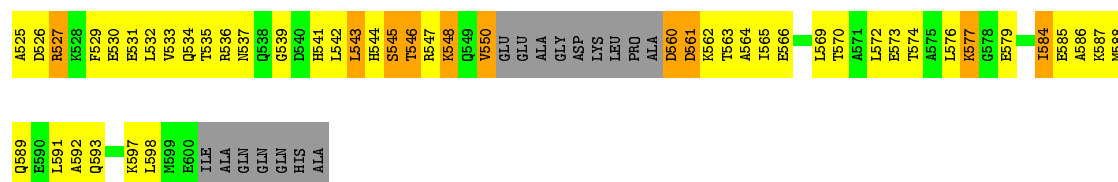


• Molecule 1: Chaperone protein dnaK



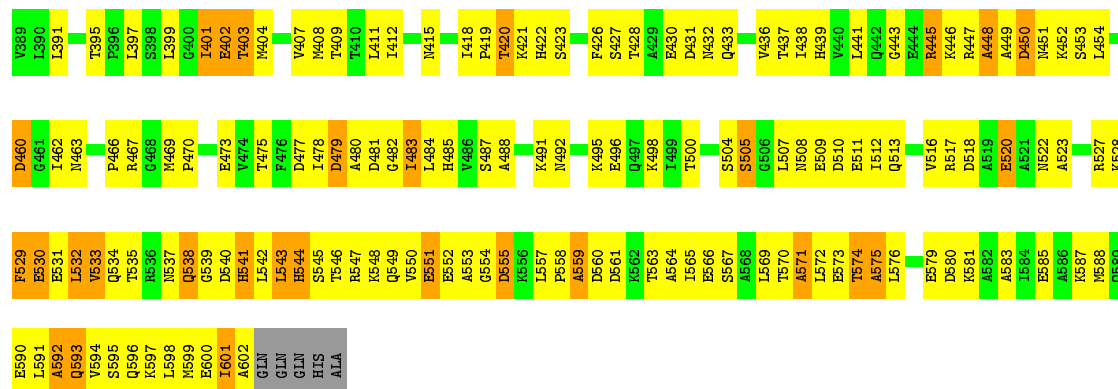
• Molecule 1: Chaperone protein dnaK





• Molecule 1: Chaperone protein dnaK

Chain F: 29% 55% 13% •



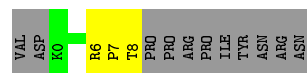
• Molecule 2: inhibitor peptide

Chain C: 40% 10% 10% 40%



• Molecule 2: inhibitor peptide

Chain D: 30% 15% 55%



• Molecule 2: inhibitor peptide

Chain G: 15% 35% 10% 40%



• Molecule 2: inhibitor peptide

Chain H: 5% 35% 5% 55%



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	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
$\langle I/\sigma(I) \rangle$ ¹	3.60 (at 1.94Å)	Xtriage
Refinement program	REFMAC 5.4.0073	Depositor
R, R_{free}	0.254 , 0.319 0.266 , 0.273	Depositor DCC
R_{free} test set	3491 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	24.3	Xtriage
Anisotropy	0.554	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 93.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.85	EDS
Total number of atoms	8427	wwPDB-VP
Average B, all atoms (Å ²)	24.0	wwPDB-VP

Xtriage'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.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: 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 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	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
4	D	36	0	0	2	0
4	E	334	0	0	12	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
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

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 69.

The worst 5 of 964 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: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	Interatomic distance (Å)	Clash overlap (Å)
1:B:503:ALA:CB	1:E:390:LEU:CD1[3_545]	2.14	0.06

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	215/219 (98%)	183 (85%)	23 (11%)	9 (4%)	3	3
1	B	216/219 (99%)	177 (82%)	29 (13%)	10 (5%)	2	3
1	E	208/219 (95%)	186 (89%)	18 (9%)	4 (2%)	8	15

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	216/219 (99%)	173 (80%)	23 (11%)	20 (9%)	0	0
2	C	9/20 (45%)	7 (78%)	1 (11%)	1 (11%)	0	0
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%)	2	2

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 ⓘ

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%)	12	25
1	B	180/181 (99%)	160 (89%)	20 (11%)	6	11
1	E	179/181 (99%)	157 (88%)	22 (12%)	4	9
1	F	181/181 (100%)	166 (92%)	15 (8%)	11	22
2	C	11/19 (58%)	10 (91%)	1 (9%)	9	18
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%)	6	11
All	All	758/800 (95%)	685 (90%)	73 (10%)	8	16

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
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 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	$\# Z > 2$
2	ALC	D	3	2	9,11,12	0.72	0	10,13,15	1.04	1 (10%)
2	ALC	G	3	2	9,11,12	0.55	0	10,13,15	1.21	1 (10%)
2	ALC	H	3	2	9,11,12	0.52	0	10,13,15	1.41	1 (10%)
2	ALC	C	3	2	9,11,12	0.57	0	10,13,15	1.40	1 (10%)

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	D	3	2	-	0/5/14/16	0/1/1/1
2	ALC	G	3	2	-	1/5/14/16	0/1/1/1
2	ALC	H	3	2	-	1/5/14/16	0/1/1/1
2	ALC	C	3	2	-	0/5/14/16	0/1/1/1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	3	ALC	CG-CB-CA	-3.38	109.98	114.52
2	C	3	ALC	CE2-CD2-CG	-2.78	106.89	112.15
2	D	3	ALC	CB-CG-CD2	-2.38	106.57	111.73
2	G	3	ALC	CE1-CD1-CG	-2.02	108.33	112.15

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	G	3	ALC	CA-CB-CG-CD1
2	H	3	ALC	N-CA-CB-CG

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	G	3	ALC	2	0
2	H	3	ALC	1	0
2	C	3	ALC	4	0

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	0.19	0	6,6,6	0.28	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.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	H	1
2	G	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	G	3:ALC	C	4:LEU	N	1.63
1	H	3:ALC	C	4:LEU	N	1.13

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.