



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 01:09 PM GMT

PDB ID : 3T39
Title : Crystal structure of the complex of camel peptidoglycan recognition protein(CPGRP-S) with a mycobacterium metabolite shikimate at 2.7 Å resolution
Authors : Sharma, P.; Dube, D.; Sinha, M.; Kaur, P.; Sharma, S.; Singh, T.P.
Deposited on : 2011-07-25
Resolution : 2.70 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
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

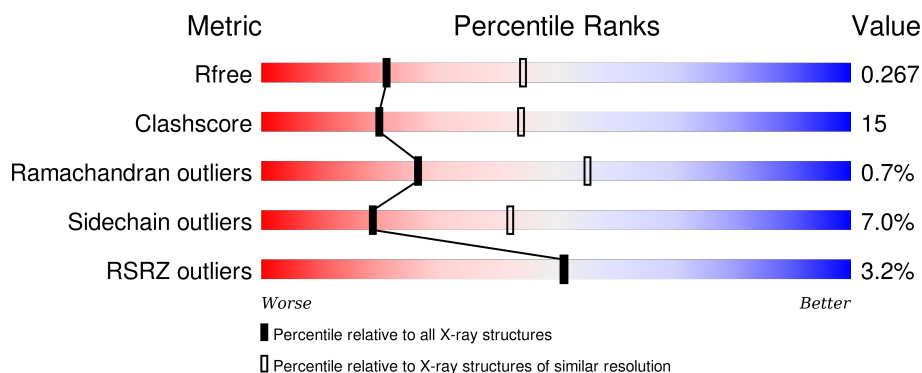
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 Å.

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



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (2.70-2.70)

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. 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	171	<div> <div>4%</div> <div>76%</div> <div>22%</div> <div>.</div> </div>
1	B	171	<div> <div>5%</div> <div>63%</div> <div>33%</div> <div>5%</div> </div>
1	C	171	<div> <div>3%</div> <div>75%</div> <div>24%</div> <div>..</div> </div>
1	D	171	<div> <div>2%</div> <div>68%</div> <div>30%</div> <div>.</div> </div>

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

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

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	GOL	D	503	-	-	-	X
3	SRT	D	172	-	-	-	X

2 Entry composition [i](#)

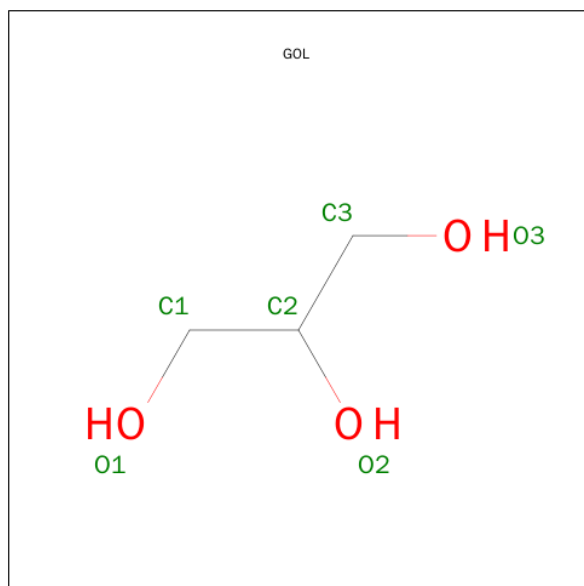
There are 5 unique types of molecules in this entry. The entry contains 5663 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Peptidoglycan recognition protein 1.

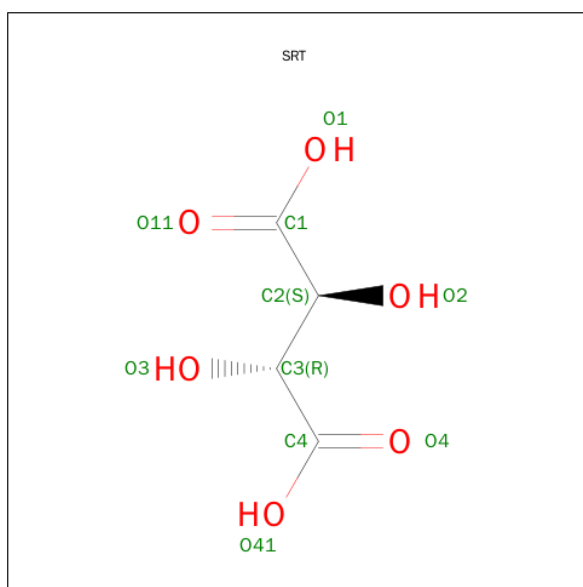
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	171	Total	C	N	O	S	0	0	0
			1336	834	254	240	8			
1	B	171	Total	C	N	O	S	0	0	0
			1336	834	254	240	8			
1	C	171	Total	C	N	O	S	0	0	0
			1336	834	254	240	8			
1	D	171	Total	C	N	O	S	0	0	0
			1336	834	254	240	8			

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



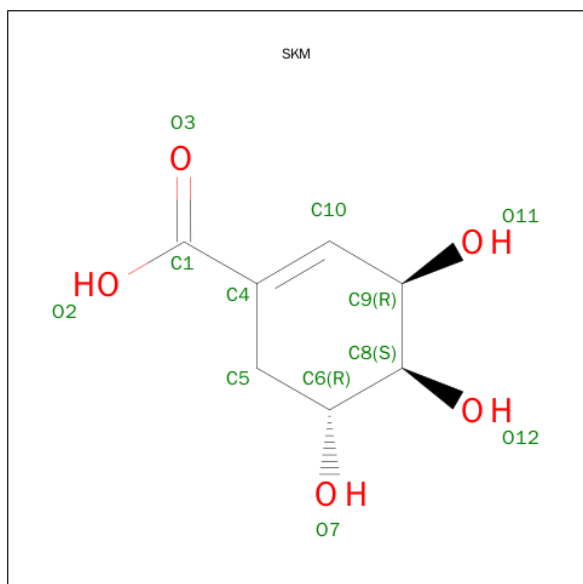
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 3 is S,R MESO-TARTARIC ACID (three-letter code: SRT) (formula: C₄H₆O₆).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	D	1	Total	C	O	0	0
			10	4	6		

- Molecule 4 is (3R,4S,5R)-3,4,5-TRIHIDROXYCYCLOHEX-1-ENE-1-CARBOXYLIC ACID (three-letter code: SKM) (formula: C₇H₁₀O₅).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	D	1	Total	C	O	0	0
			12	7	5		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	85	Total 85	O 85	0	0
5	B	74	Total 74	O 74	0	0
5	C	65	Total 65	O 65	0	0
5	D	67	Total 67	O 67	0	0

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.

Chain A:

Sequence logo for Chain A. The y-axis represents information content in bits (0.00 to 0.25). The x-axis lists amino acids. A horizontal bar at the top shows the percentage of positions where a specific amino acid is conserved: 76% for green (conserved), 22% for yellow (partially conserved), and 4% for red (not conserved).

Amino Acid	Conservation Status	Percentage of Positions
E1	Conserved (Green)	76%
D2	Conserved (Green)	
P3	Conserved (Green)	
P4	Conserved (Green)	
A5	Conserved (Green)	
S8	Conserved (Green)	22%
R12	Conserved (Green)	
V15	Conserved (Green)	
A16	Conserved (Green)	
A17	Conserved (Green)	
V35	Conserved (Green)	
S36	Conserved (Green)	
H37	Conserved (Green)	
C43	Conserved (Green)	
H60	Conserved (Green)	
V61	Conserved (Green)	4%
R62	Conserved (Green)	
G70	Conserved (Green)	
Y71	Conserved (Green)	
L74	Conserved (Green)	
L75	Conserved (Green)	
G76	Conserved (Green)	
E77	Conserved (Green)	
V81	Conserved (Green)	
V82	Conserved (Green)	
E83	Conserved (Green)	
G84	Conserved (Green)	
R85	Conserved (Green)	
G91	Conserved (Green)	
G95	Conserved (Green)	
V98	Conserved (Green)	
I99	Conserved (Green)	
P100	Conserved (Green)	
I105	Conserved (Green)	
S106	Conserved (Green)	
M112	Conserved (Green)	
M113	Conserved (Green)	
L121	Conserved (Green)	
Q125	Conserved (Green)	
M126	Conserved (Green)	
E142	Conserved (Green)	
V140	Conserved (Green)	

Chain B:

63% 33% 5%

E1 D2 P3 P4 A5 G6 G7 S8 R28 P29 V30 R31 Y32 S36 H37 C43 S48 Y59 M63 V69 L74 I75 G76 E77 V81 Y82 R85 G86 H87 H88 L89 T97 P98 I99 P100 I101 S102 I103 Y111 M112 R113 A114 P117 R122 H126

L127 G130 G131 V132 A133 L134 G135 A136 L137 R138 S139 M140 Y141 E142 H146 R147 D148 V149 T152 L153 S154 G156 D157 R158 I163 Q164 T165 H168 Y169 R170 A171

Chain C:

75% 24% 3%

••

Chain C items (174 total):

- V115 (Red)
- P116 (Orange)
- R119 (Orange)
- R138 (Orange)
- L153 (Orange)
- D157 (Orange)
- R158 (Orange)
- E161 (Orange)
- I162 (Orange)
- A174 (Orange)
- E1 (Green)
- D2 (Green)
- P3 (Orange)
- P4 (Orange)
- A5 (Green)
- C6 (Green)
- G7 (Green)
- S8 (Green)
- I9 (Green)
- R12 (Yellow)
- A17 (Yellow)
- L18 (Yellow)
- E24 (Yellow)
- T27 (Yellow)
- R31 (Yellow)
- Y32 (Yellow)
- E37 (Yellow)
- C43 (Yellow)
- Y59 (Yellow)
- H60 (Yellow)
- M63 (Yellow)
- M66 (Yellow)
- Y71 (Yellow)
- I75 (Green)
- G76 (Green)
- E77 (Yellow)
- V81 (Yellow)
- Y82 (Green)
- E83 (Green)
- G84 (Green)
- R85 (Yellow)
- K90 (Yellow)
- G91 (Yellow)
- W98 (Yellow)
- N99 (Yellow)
- P100 (Yellow)
- I101 (Yellow)
- S102 (Green)
- I103 (Green)
- S106 (Yellow)
- M112 (Yellow)
- N113 (Green)
- R114 (Yellow)

Chain D:

Category	Percentage
E1	2%
D2	2%
P3	2%
D3	2%
P4	2%
A5	2%
A6	2%
C6	2%
G7	2%
W5	2%
R46	2%
A17	2%
E21	2%
Y32	2%
H37	2%
T38	2%
A39	2%
C43	2%
D44	2%
T45	2%
S48	2%
Q51	2%
Q52	2%
A53	2%
Q57	2%
S58	2%
Y59	2%
H60	2%
V61	2%
R62	2%
M63	2%
D68	2%
V69	2%
G70	2%
Y71	2%
L74	2%
I75	2%
G76	2%
E77	2%
D78	2%
Y82	2%
R85	2%
K90	2%
W98	2%
I101	2%
I105	2%

S106	F107	M108	G109	M112	N113	P117	L121	A124	Q125	N126	L127	C130	N140	K144	L153	S154	P155	G156	L159	I162	T169	R170	A171
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	89.07Å 101.14Å 163.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.02 – 2.70 48.02 – 2.70	Depositor EDS
% Data completeness (in resolution range)	98.8 (48.02-2.70) 98.9 (48.02-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.40 (at 2.69Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.219 , 0.244 0.214 , 0.267	Depositor DCC
R_{free} test set	1045 reflections (5.12%)	DCC
Wilson B-factor (Å ²)	50.7	Xtriage
Anisotropy	0.730	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 46.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 20479 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	5663	wwPDB-VP
Average B, all atoms (Å ²)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.42% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SKM, GOL, SRT

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	0.43	0/1373	0.80	0/1871
1	B	0.55	0/1373	0.89	2/1871 (0.1%)
1	C	0.40	0/1373	0.81	2/1871 (0.1%)
1	D	0.43	0/1373	0.84	2/1871 (0.1%)
All	All	0.46	0/5492	0.84	6/7484 (0.1%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	2	ASP	N-CA-C	-7.94	89.55	111.00
1	B	5	ALA	N-CA-C	6.64	128.92	111.00
1	B	5	ALA	N-CA-CB	-6.32	101.25	110.10
1	C	18	LEU	N-CA-C	-5.48	96.20	111.00
1	C	115	VAL	N-CA-C	-5.29	96.71	111.00

There are no chirality outliers.

There are no planarity outliers.

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	1336	0	1288	29	0
1	B	1336	0	1288	52	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1336	0	1288	31	0
1	D	1336	0	1288	44	0
2	D	6	0	8	1	0
3	D	10	0	3	1	0
4	D	12	0	9	0	0
5	A	85	0	0	0	1
5	B	74	0	0	1	1
5	C	65	0	0	1	0
5	D	67	0	0	0	0
All	All	5663	0	5172	156	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 15.

The worst 5 of 156 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:134:LEU:N	1:B:134:LEU:HD23	1.71	1.05
1:B:112:MET:HE1	1:B:153:LEU:HG	1.52	0.90
1:D:112:MET:HE1	1:D:153:LEU:HB3	1.56	0.86
1:B:133:ALA:C	1:B:134:LEU:HD23	2.00	0.82
1:C:75:ILE:HD13	1:C:81:VAL:HG22	1.60	0.82

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)
5:A:202:HOH:O	5:B:232:HOH:O[7_454]	1.46	0.74

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	169/171 (99%)	155 (92%)	13 (8%)	1 (1%)	30	59
1	B	169/171 (99%)	149 (88%)	20 (12%)	0	100	100
1	C	169/171 (99%)	159 (94%)	8 (5%)	2 (1%)	16	39
1	D	169/171 (99%)	155 (92%)	12 (7%)	2 (1%)	16	39
All	All	676/684 (99%)	618 (91%)	53 (8%)	5 (1%)	26	55

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	7	GLY
1	D	4	PRO
1	D	7	GLY
1	A	3	PRO
1	C	3	PRO

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	139/139 (100%)	130 (94%)	9 (6%)	21	46
1	B	139/139 (100%)	122 (88%)	17 (12%)	6	14
1	C	139/139 (100%)	133 (96%)	6 (4%)	35	66
1	D	139/139 (100%)	132 (95%)	7 (5%)	30	60
All	All	556/556 (100%)	517 (93%)	39 (7%)	19	42

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	130	CYS
1	B	152	THR
1	D	113	ASN
1	B	134	LEU
1	B	140	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	140	ASN
1	D	140	ASN
1	C	52	GLN
1	B	37	HIS
1	D	63	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

3 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	$\# Z > 2$
3	SRT	D	172	-	3,9,9	4.12	2 (66%)	6,12,12	2.55	2 (33%)
2	GOL	D	503	-	5,5,5	0.29	0	5,5,5	0.30	0
4	SKM	D	6873	-	10,12,12	2.87	5 (50%)	9,17,17	1.37	1 (11%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means

no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	SRT	D	172	-	-	0/4/12/12	0/0/0/0
2	GOL	D	503	-	-	0/4/4/4	0/0/0/0
4	SKM	D	6873	-	-	0/0/20/20	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	172	SRT	O3-C3	-3.53	1.35	1.42
4	D	6873	SKM	C8-C9	2.01	1.57	1.52
4	D	6873	SKM	O7-C6	2.29	1.48	1.43
4	D	6873	SKM	O11-C9	2.75	1.49	1.43
4	D	6873	SKM	C6-C8	5.11	1.59	1.52

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	172	SRT	C1-C2-C3	-5.73	101.59	113.35
3	D	172	SRT	O2-C2-C3	2.06	114.41	108.61
4	D	6873	SKM	O12-C8-C6	2.90	115.23	110.00

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	172	SRT	1	0
2	D	503	GOL	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 [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, 95th 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 ²)	Q<0.9
1	A	171/171 (100%)	-0.04	6 (3%)	48	48	30, 47, 63, 100	0
1	B	171/171 (100%)	0.15	8 (4%)	35	34	33, 56, 83, 102	0
1	C	171/171 (100%)	-0.14	5 (2%)	55	55	30, 43, 61, 99	0
1	D	171/171 (100%)	-0.02	3 (1%)	71	72	31, 46, 65, 100	0
All	All	684/684 (100%)	-0.01	22 (3%)	51	51	30, 47, 76, 102	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	2	ASP	11.7
1	A	2	ASP	10.1
1	A	3	PRO	9.6
1	B	1	GLU	8.8
1	A	4	PRO	7.8

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 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, 95th percentile and maximum values of B factors

of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	SRT	D	172	10/10	0.84	0.33	8.10	72,75,76,77	0
2	GOL	D	503	6/6	0.85	0.22	4.05	57,58,59,59	0
4	SKM	D	6873	12/12	0.75	0.25	0.18	96,96,97,97	0

6.5 Other polymers [i](#)

There are no such residues in this entry.