



wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 10:02 pm BST

PDB ID : 1JCH
Title : Crystal Structure of Colicin E3 in Complex with its Immunity Protein
Authors : Soelaiman, S.; Jakes, K.; Wu, N.; Li, C.; Shoham, M.
Deposited on : 2001-06-09
Resolution : 3.02 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

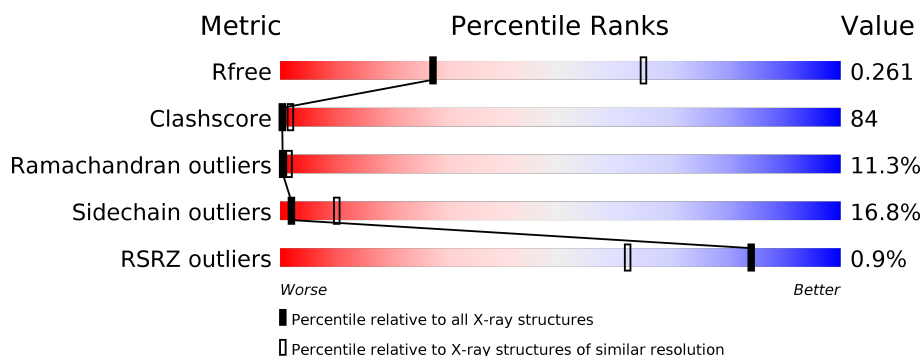
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.02 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	2399 (3.04-3.00)
Clashscore	141614	2734 (3.04-3.00)
Ramachandran outliers	138981	2640 (3.04-3.00)
Sidechain outliers	138945	2643 (3.04-3.00)
RSRZ outliers	127900	2287 (3.04-3.00)

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 $\leq 5\%$. 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	551	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; font-size: 0.8em;"> 19% 48% 15% • 15% </div> </div>
1	C	551	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; font-size: 0.8em;"> 19% 48% 15% • 15% </div> </div>
2	B	84	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green, yellow, orange);"></div> <div style="display: flex; justify-content: space-between; font-size: 0.8em;"> 32% 61% 7% </div> </div>
2	D	84	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green, yellow, orange);"></div> <div style="display: flex; justify-content: space-between; font-size: 0.8em;"> 35% 58% 7% </div> </div>

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 8998 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 COLICIN E3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	468	Total	C	N	O	S	0	0	0
			3570	2206	659	697	8			
1	C	468	Total	C	N	O	S	0	0	0
			3570	2206	659	697	8			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	260	PHE	GLY	CONFLICT	UNP P00646
A	261	GLY	PHE	CONFLICT	UNP P00646
C	260	PHE	GLY	CONFLICT	UNP P00646
C	261	GLY	PHE	CONFLICT	UNP P00646

- Molecule 2 is a protein called COLICIN E3 IMMUNITY PROTEIN.

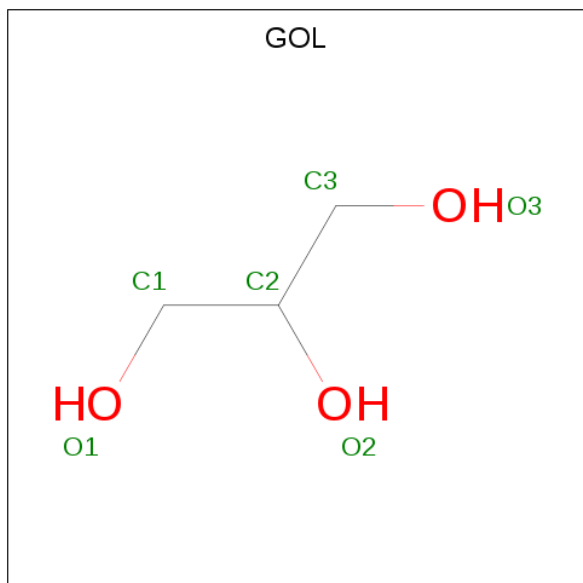
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	84	Total	C	N	O	S	0	0	0
			693	445	104	142	2			
2	D	84	Total	C	N	O	S	0	0	0
			693	445	104	142	2			

- Molecule 3 is CITRIC ACID (three-letter code: CIT) (formula: C₆H₈O₇).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			13	6	7		
3	A	1	Total	C	O	0	0
			13	6	7		
3	C	1	Total	C	O	0	0
			13	6	7		
3	C	1	Total	C	O	0	0
			13	6	7		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



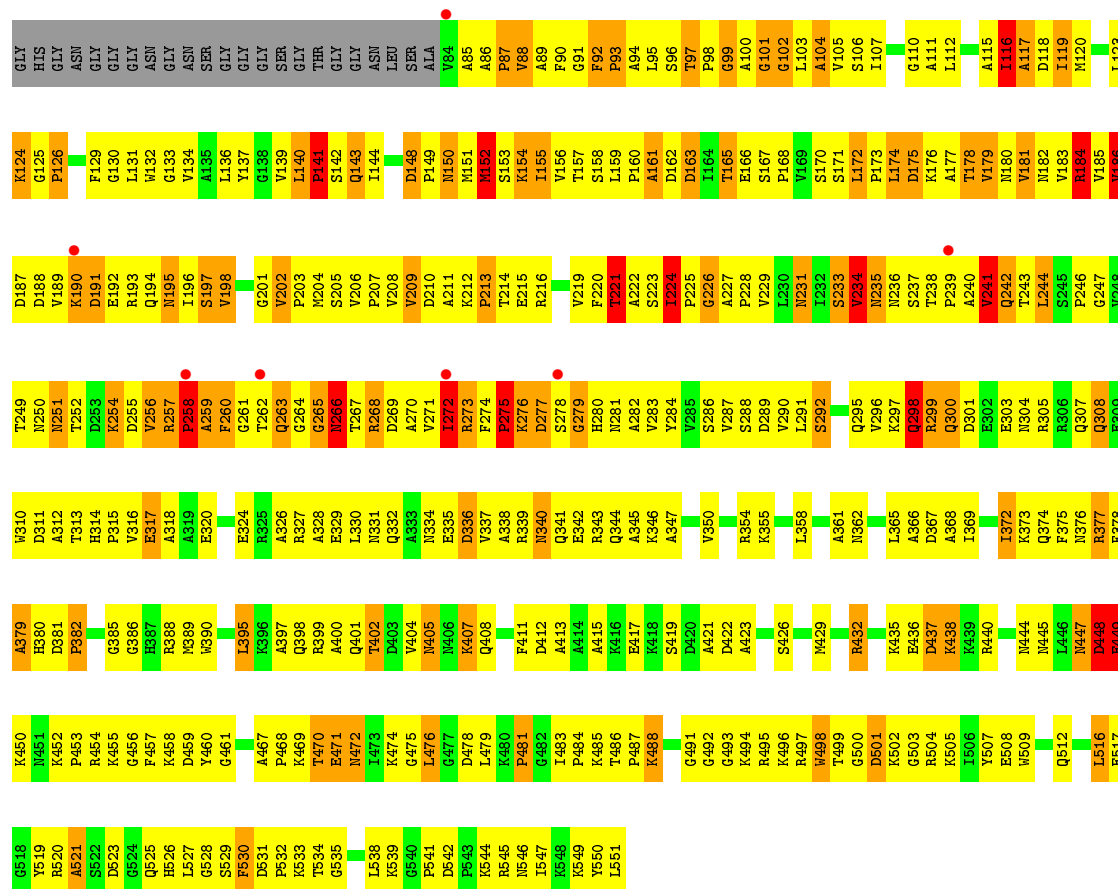
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 6	C 3	O 3	0	0
4	B	1	Total 6	C 3	O 3	0	0
4	C	1	Total 6	C 3	O 3	0	0
4	D	1	Total 6	C 3	O 3	0	0

- Molecule 5 is water.

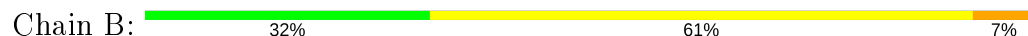
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	169	Total 169	O 169	0	0
5	B	28	Total 28	O 28	0	0
5	C	170	Total 170	O 170	0	0
5	D	29	Total 29	O 29	0	0

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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 ($\text{RSRZ} > 2$). 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.

[illegible][illegible]



• Molecule 2: COLICIN E3 IMMUNITY PROTEIN



• Molecule 2: COLICIN E3 IMMUNITY PROTEIN



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	66.99Å 195.73Å 85.12Å 90.00° 113.22° 90.00°	Depositor
Resolution (Å)	20.17 – 3.02 20.17 – 3.02	Depositor EDS
% Data completeness (in resolution range)	94.1 (20.17-3.02) 94.1 (20.17-3.02)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.26 (at 3.04Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.237 , 0.283 0.221 , 0.261	Depositor DCC
R_{free} test set	1127 reflections (3.04%)	wwPDB-VP
Wilson B-factor (Å ²)	72.3	Xtriage
Anisotropy	0.423	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 70.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.450 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	8998	wwPDB-VP
Average B, all atoms (Å ²)	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.91% 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, 0.333 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: GOL, CIT

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	0.67	0/3639	0.88	8/4922 (0.2%)
1	C	0.67	0/3639	0.88	8/4922 (0.2%)
2	B	0.68	0/714	0.80	0/967
2	D	0.68	0/714	0.80	0/967
All	All	0.67	0/8706	0.87	16/11778 (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
2	B	0	1
2	D	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	260	PHE	CB-CG-CD2	-11.78	112.56	120.80
1	C	260	PHE	CB-CG-CD2	-11.78	112.56	120.80
1	C	260	PHE	CB-CG-CD1	10.12	127.89	120.80
1	A	260	PHE	CB-CG-CD1	10.05	127.84	120.80
1	A	260	PHE	CB-CA-C	6.01	122.42	110.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	61	TYR	Sidechain
2	D	61	TYR	Sidechain

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	3570	0	3532	653	8
1	C	3570	0	3532	655	9
2	B	693	0	617	75	0
2	D	693	0	617	73	0
3	A	26	0	10	1	0
3	C	26	0	10	1	0
4	A	6	0	8	0	0
4	B	6	0	8	2	0
4	C	6	0	8	0	0
4	D	6	0	8	1	0
5	A	169	0	0	42	1
5	B	28	0	0	2	0
5	C	170	0	0	43	0
5	D	29	0	0	2	0
All	All	8998	0	8350	1418	9

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 84.

The worst 5 of 1418 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:C:257:ARG:HB3	1:C:258:PRO:CD	1.41	1.39
1:A:257:ARG:HB3	1:A:258:PRO:CD	1.41	1.35
1:C:179:VAL:HB	1:C:286:SER:HB2	1.22	1.17
1:C:272:ILE:HB	1:C:282:ALA:HA	1.22	1.16
1:A:272:ILE:HB	1:A:282:ALA:HA	1.22	1.16

The worst 5 of 9 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:A:234:VAL:CG2	1:C:324:GLU:OE2[2_655]	1.62	0.58
1:A:324:GLU:OE2	1:C:234:VAL:CG2[2_555]	1.86	0.34
1:A:126:PRO:O	1:C:239:PRO:CD[2_555]	1.88	0.32
1:A:239:PRO:CD	1:C:126:PRO:O[2_655]	1.94	0.26
1:A:127:PHE:CD1	1:C:236:ASN:O[2_555]	2.00	0.20

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	466/551 (85%)	294 (63%)	115 (25%)	57 (12%)	0	1
1	C	466/551 (85%)	294 (63%)	115 (25%)	57 (12%)	0	1
2	B	82/84 (98%)	63 (77%)	14 (17%)	5 (6%)	1	8
2	D	82/84 (98%)	63 (77%)	14 (17%)	5 (6%)	1	8
All	All	1096/1270 (86%)	714 (65%)	258 (24%)	124 (11%)	0	1

5 of 124 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	101	GLY
1	A	102	GLY
1	A	142	SER
1	A	152	MET
1	A	155	ILE

5.3.2 Protein sidechains [i](#)

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	377/422 (89%)	305 (81%)	72 (19%)	1	7
1	C	377/422 (89%)	305 (81%)	72 (19%)	1	7
2	B	76/76 (100%)	72 (95%)	4 (5%)	22	57
2	D	76/76 (100%)	72 (95%)	4 (5%)	22	57
All	All	906/996 (91%)	754 (83%)	152 (17%)	2	10

5 of 152 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	501	ASP
1	C	152	MET
1	C	471	GLU
1	A	529	SER
1	C	96	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	525	GLN
1	C	195	ASN
1	C	525	GLN
2	B	59	GLN
2	B	65	GLN

5.3.3 RNA ⓘ

There are no RNA molecules 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

8 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	CIT	A	601	-	3,12,12	3.16	2 (66%)	3,17,17	2.14	2 (66%)
4	GOL	D	5702	-	5,5,5	1.01	0	5,5,5	0.24	0
3	CIT	A	602	-	3,12,12	3.24	2 (66%)	3,17,17	2.14	1 (33%)
4	GOL	A	5701	-	5,5,5	1.14	1 (20%)	5,5,5	0.57	0
4	GOL	C	701	-	5,5,5	1.14	1 (20%)	5,5,5	0.57	0
3	CIT	C	5601	-	3,12,12	3.14	2 (66%)	3,17,17	2.14	2 (66%)
4	GOL	B	702	-	5,5,5	1.01	0	5,5,5	0.25	0
3	CIT	C	5602	-	3,12,12	3.24	2 (66%)	3,17,17	2.15	1 (33%)

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	CIT	A	601	-	-	0/6/16/16	-
4	GOL	D	5702	-	-	4/4/4/4	-
3	CIT	A	602	-	-	0/6/16/16	-
4	GOL	A	5701	-	-	2/4/4/4	-
4	GOL	C	701	-	-	2/4/4/4	-
3	CIT	C	5601	-	-	0/6/16/16	-
4	GOL	B	702	-	-	4/4/4/4	-
3	CIT	C	5602	-	-	0/6/16/16	-

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	602	CIT	C2-C3	4.21	1.60	1.54
3	C	5602	CIT	C2-C3	4.19	1.60	1.54

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	5601	CIT	C2-C3	3.90	1.60	1.54
3	A	601	CIT	C2-C3	3.87	1.60	1.54
3	A	601	CIT	C4-C3	3.86	1.60	1.54

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	602	CIT	C3-C2-C1	3.35	120.34	114.98
3	C	5602	CIT	C3-C2-C1	3.34	120.34	114.98
3	A	601	CIT	C3-C2-C1	3.05	119.86	114.98
3	C	5601	CIT	C3-C2-C1	3.04	119.86	114.98
3	A	601	CIT	C3-C4-C5	-2.12	111.59	114.98

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	D	5702	GOL	O1-C1-C2-C3
4	D	5702	GOL	C1-C2-C3-O3
4	A	5701	GOL	O1-C1-C2-C3
4	C	701	GOL	O1-C1-C2-C3
4	B	702	GOL	O1-C1-C2-C3

There are no ring outliers.

4 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	601	CIT	1	0
4	D	5702	GOL	1	0
3	C	5601	CIT	1	0
4	B	702	GOL	2	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	468/551 (84%)	-0.30	3 (0%) 89 72	25, 73, 122, 147	0
1	C	468/551 (84%)	-0.31	7 (1%) 73 46	25, 73, 122, 147	0
2	B	84/84 (100%)	-0.34	0 100 100	26, 58, 87, 95	0
2	D	84/84 (100%)	-0.32	0 100 100	26, 58, 87, 95	0
All	All	1104/1270 (86%)	-0.31	10 (0%) 84 62	25, 70, 120, 147	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	84	VAL	3.4
1	A	272	ILE	3.1
1	C	272	ILE	2.9
1	A	84	VAL	2.9
1	A	190	LYS	2.6

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. 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	B-factors(\AA^2)	Q<0.9
3	CIT	A	602	13/13	0.81	0.30	103,143,168,170	0
4	GOL	C	701	6/6	0.81	0.28	63,76,79,81	0
4	GOL	A	5701	6/6	0.82	0.21	63,76,79,81	0
3	CIT	C	5601	13/13	0.85	0.17	67,122,144,151	0
4	GOL	B	702	6/6	0.86	0.24	76,101,102,106	0
3	CIT	A	601	13/13	0.88	0.16	67,122,144,151	0
3	CIT	C	5602	13/13	0.89	0.24	103,143,168,170	0
4	GOL	D	5702	6/6	0.93	0.15	76,101,102,106	0

6.5 Other polymers [i](#)

There are no such residues in this entry.