wwPDB X-ray Structure Validation Summary Report (i)

Dec 6, 2023 – 05:42 am GMT

PDB ID	:	1H0G
Title	:	Complex of a chitinase with the natural product cyclopentapeptide argadin from Clonostachys
Authors	:	Houston, D.; Shiomi, K.; Arai, N.; Omura, S.; Peter, M.G.; Turberg, A.; Synstad, B.; Eijsink, V.G.H.; Aalten, D.M.F.
Deposited on	:	2002-06-19
Resolution	:	2.00 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

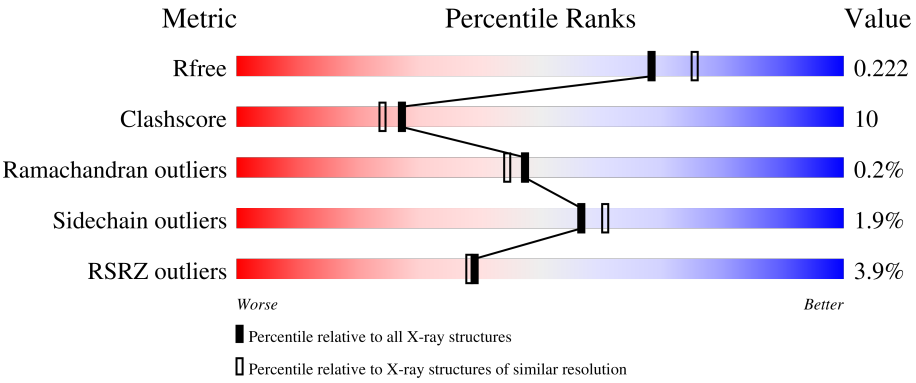
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 of 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 <=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	499	5% 80% 19%
1	B	499	3% 81% 18%
2	C	5	40% 20% 40%
2	D	5	40% 20% 40%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 8542 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 CHITINASE B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	497	Total	C	N	O	S	0	1	1
			3903	2496	659	734	14			
1	B	497	Total	C	N	O	S	0	1	0
			3899	2494	654	737	14			

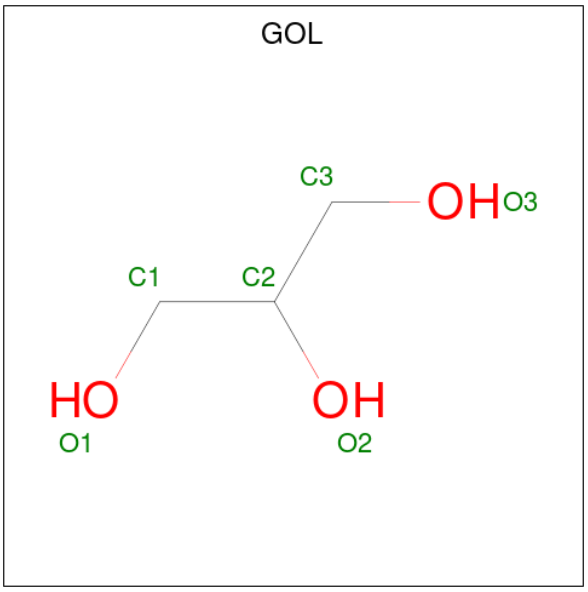
There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	119	SER	ALA	conflict	UNP P11797
A	171	THR	ALA	conflict	UNP P11797
A	223	VAL	ILE	conflict	UNP P11797
A	320	SER	ASN	conflict	UNP P11797
A	321	THR	ALA	conflict	UNP P11797
A	329	GLU	ASP	conflict	UNP P11797
A	498	VAL	LEU	conflict	UNP P11797
B	119	SER	ALA	conflict	UNP P11797
B	171	THR	ALA	conflict	UNP P11797
B	223	VAL	ILE	conflict	UNP P11797
B	320	SER	ASN	conflict	UNP P11797
B	321	THR	ALA	conflict	UNP P11797
B	329	GLU	ASP	conflict	UNP P11797
B	498	VAL	LEU	conflict	UNP P11797

- Molecule 2 is a protein called Argadin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	5	Total	C	N	O	0	0	0
			48	29	10	9			
2	D	5	Total	C	N	O	0	0	0
			48	29	10	9			

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 4 is water.

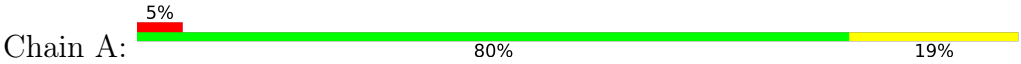
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	310	Total O 310 310	0	0
4	B	298	Total O 298 298	0	0
4	C	2	Total O 2 2	0	0
4	D	4	Total O 4 4	0	0

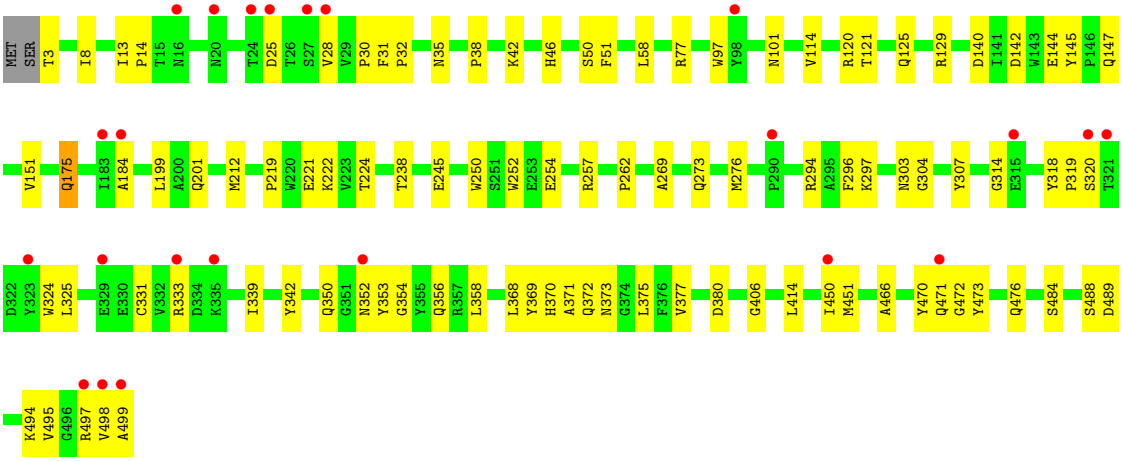
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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 (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: CHITINASE B

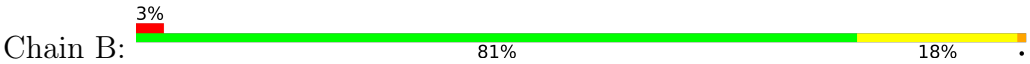
Chain A:

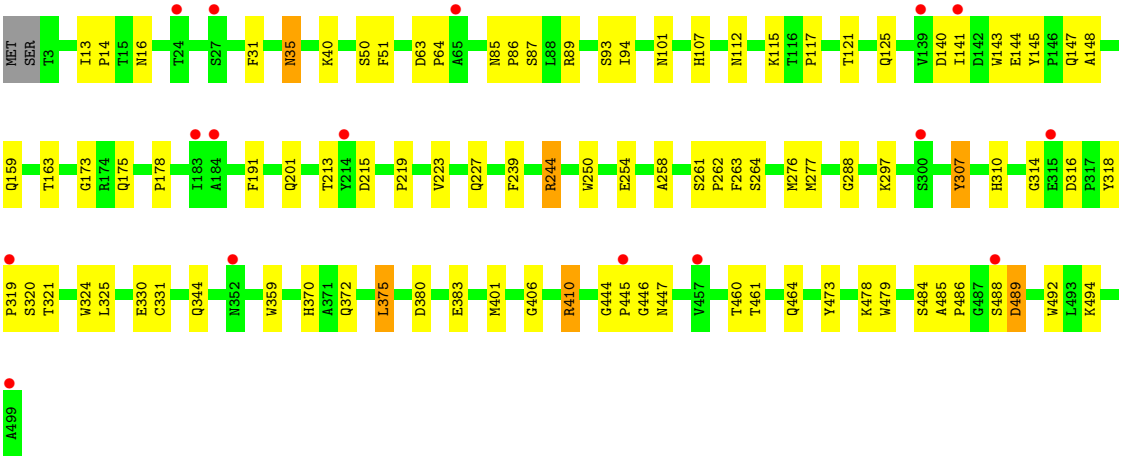




• Molecule 1: CHITINASE B

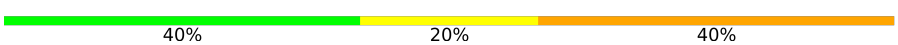
Chain B:





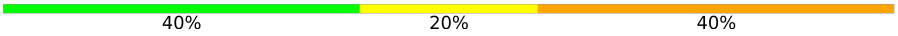
• Molecule 2: Argadin

Chain C:



• Molecule 2: Argadin

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	55.06Å 102.93Å 185.78Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.65 – 2.00 29.65 – 2.00	Depositor EDS
% Data completeness (in resolution range)	97.0 (29.65-2.00) 97.2 (29.65-2.00)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.24 (at 2.00Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.203 , 0.231 0.195 , 0.222	Depositor DCC
R_{free} test set	1030 reflections (1.47%)	wwPDB-VP
Wilson B-factor (Å^2)	28.6	Xtriage
Anisotropy	0.257	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 65.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	8542	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.02% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 0AR, DPR, HSE, GOL, UN1

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	0.48	0/4018	0.67	0/5478
1	B	0.50	0/4014	0.67	0/5473
2	C	1.32	0/10	1.06	0/12
2	D	1.25	0/10	1.02	0/12
All	All	0.49	0/8052	0.67	0/10975

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-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 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	3903	0	3729	74	0
1	B	3899	0	3712	76	0
2	C	48	0	31	4	0
2	D	48	0	31	5	0
3	A	6	0	8	0	0
3	B	18	0	24	6	0
3	C	6	0	8	2	0
4	A	310	0	0	15	0
4	B	298	0	0	9	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	C	2	0	0	0	0
4	D	4	0	0	0	0
All	All	8542	0	7543	156	0

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 10.

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:16:ASN:HB3	4:B:2012:HOH:O	1.75	0.86
2:C:1:0AR:N	2:C:5:UN1:H6C1	1.94	0.82
2:D:1:0AR:N	2:D:5:UN1:H6C1	1.93	0.81
1:A:352:ASN:O	1:A:354:GLY:N	2.14	0.79
1:B:219:PRO:HB2	1:B:314:GLY:HA2	1.64	0.78

There are no symmetry-related clashes.

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	496/499 (99%)	485 (98%)	10 (2%)	1 (0%)	47 44
1	B	496/499 (99%)	481 (97%)	14 (3%)	1 (0%)	47 44
2	C	1/5 (20%)	1 (100%)	0	0	100 100
2	D	1/5 (20%)	1 (100%)	0	0	100 100
All	All	994/1008 (99%)	968 (97%)	24 (2%)	2 (0%)	47 44

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	353	TYR
1	B	489	ASP

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	403/406 (99%)	397 (98%)	6 (2%)	65	69
1	B	401/406 (99%)	392 (98%)	9 (2%)	52	55
2	C	1/1 (100%)	1 (100%)	0	100	100
2	D	1/1 (100%)	1 (100%)	0	100	100
All	All	806/814 (99%)	791 (98%)	15 (2%)	57	61

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	87	SER
1	B	380	ASP
1	B	213	THR
1	B	410	ARG
1	B	316	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	107	HIS
1	B	167	GLN
1	B	431	GLN
1	B	147	GLN
1	B	227	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

8 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	0AR	D	1	2	10,13,14	1.00	0	10,15,17	1.03	1 (10%)
2	HSE	D	3	2	5,6,7	0.87	0	2,6,8	3.78	1 (50%)
2	UN1	D	5	2	8,9,10	0.84	0	5,10,12	1.54	1 (20%)
2	HSE	C	3	2	5,6,7	0.85	0	2,6,8	3.71	1 (50%)
2	0AR	C	1	2	10,13,14	0.90	0	10,15,17	1.16	1 (10%)
2	UN1	C	5	2	8,9,10	0.88	0	5,10,12	1.84	2 (40%)

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	0AR	D	1	2	-	0/12/13/15	-
2	HSE	D	3	2	-	3/4/5/7	-
2	UN1	D	5	2	-	2/7/8/10	-
2	HSE	C	3	2	-	3/4/5/7	-
2	0AR	C	1	2	-	0/12/13/15	-
2	UN1	C	5	2	-	1/7/8/10	-

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	3	HSE	C4-C3-CA	-5.33	104.84	113.08
2	C	3	HSE	C4-C3-CA	-5.21	105.02	113.08
2	C	1	0AR	C0B-NH2-CZ	-2.74	121.85	127.56
2	C	5	UN1	C6-C1-C1'	-2.71	107.63	114.47
2	D	5	UN1	C6-C1-C1'	-2.34	108.59	114.47

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	3	HSE	O-C-CA-C3
2	C	3	HSE	C4-C3-CA-C
2	C	3	HSE	CA-C3-C4-O3
2	D	3	HSE	O-C-CA-C3
2	D	3	HSE	C4-C3-CA-C

There are no ring outliers.

4 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	1	0AR	5	0
2	D	5	UN1	4	0
2	C	1	0AR	4	0
2	C	5	UN1	3	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

5 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	GOL	A	1500	-	5,5,5	0.31	0	5,5,5	0.35	0
3	GOL	B	1501	-	5,5,5	0.54	0	5,5,5	0.40	0
3	GOL	B	1504	-	5,5,5	0.27	0	5,5,5	0.26	0
3	GOL	C	1502	-	5,5,5	0.50	0	5,5,5	0.18	0
3	GOL	B	1503	-	5,5,5	0.35	0	5,5,5	0.23	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	GOL	A	1500	-	-	0/4/4/4	-
3	GOL	B	1501	-	-	0/4/4/4	-
3	GOL	B	1504	-	-	0/4/4/4	-
3	GOL	C	1502	-	-	0/4/4/4	-
3	GOL	B	1503	-	-	0/4/4/4	-

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.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1501	GOL	3	0
3	B	1504	GOL	1	0
3	C	1502	GOL	2	0
3	B	1503	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

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^{th} 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	497/499 (99%)	0.16	23 (4%) 32 31	20, 34, 57, 75	0
1	B	497/499 (99%)	0.06	16 (3%) 47 46	19, 34, 51, 63	0
2	C	1/5 (20%)	0.38	0 100 100	28, 28, 28, 28	0
2	D	1/5 (20%)	1.24	0 100 100	33, 33, 33, 33	0
All	All	996/1008 (98%)	0.11	39 (3%) 39 38	19, 34, 55, 75	0

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	28	VAL	6.0
1	A	323	TYR	4.0
1	B	488	SER	3.8
1	A	450	ILE	3.8
1	A	321	THR	3.4

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

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, 95^{th} 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(Å^2)	Q<0.9
2	UN1	C	5	10/11	0.89	0.22	31,38,44,46	0
2	0AR	D	1	14/15	0.90	0.14	39,41,47,47	0
2	UN1	D	5	10/11	0.90	0.20	34,39,46,49	0
2	DPR	C	2	7/8	0.93	0.13	36,37,38,41	0
2	HSE	D	3	7/8	0.94	0.19	33,34,37,39	0
2	0AR	C	1	14/15	0.94	0.17	31,36,39,42	0
2	DPR	D	2	7/8	0.94	0.14	41,42,44,44	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	HSE	C	3	7/8	0.96	0.17	26,29,29,33	0

6.3 Carbohydrates (i)

There are no monosaccharides 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, 95^{th} 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(Å^2)	Q<0.9
3	GOL	C	1502	6/6	0.70	0.18	61,62,62,63	0
3	GOL	B	1503	6/6	0.85	0.15	51,52,53,54	0
3	GOL	A	1500	6/6	0.85	0.18	58,59,60,61	0
3	GOL	B	1504	6/6	0.88	0.27	51,54,55,57	0
3	GOL	B	1501	6/6	0.91	0.17	32,40,42,43	0

6.5 Other polymers (i)

There are no such residues in this entry.