wwPDB X-ray Structure Validation Summary Report (i)

Mar 18, 2024 – 11:21 AM JST

PDB ID	:	5ZAC
Title	:	Crystal structure of ConA-R2M
Authors	:	Gan, J.H.; Chen, G.S.; Hu, R.T.
Deposited on	:	2018-02-07
Resolution	:	2.59 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 (1)) were used in the production of this report:

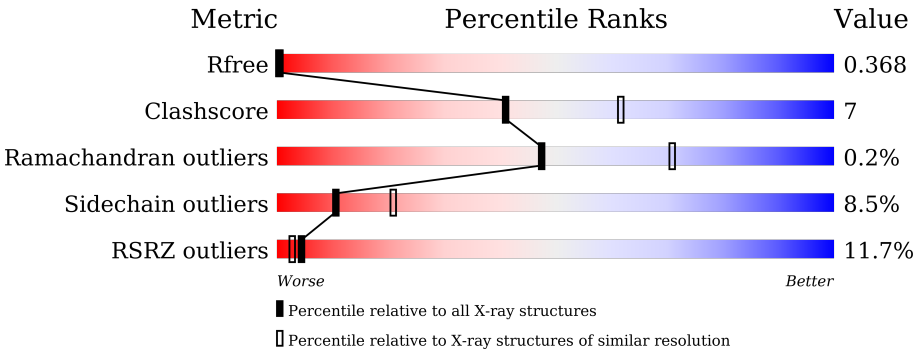
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 2.59 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)
RSRZ outliers	127900	3104 (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 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	237	11% 80% 16%
1	B	237	9% 79% 17%
1	C	237	12% 75% 21%
1	D	237	14% 77% 18%

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	TA5	D	305	-	-	-	X

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 7348 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 Concanavalin-A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	230	Total	C	N	O	S	0	0	0
			1760	1113	292	353	2			
1	B	232	Total	C	N	O	S	0	1	0
			1772	1121	294	355	2			
1	C	231	Total	C	N	O	S	0	1	0
			1763	1116	293	352	2			
1	D	231	Total	C	N	O	S	0	1	0
			1763	1116	293	352	2			

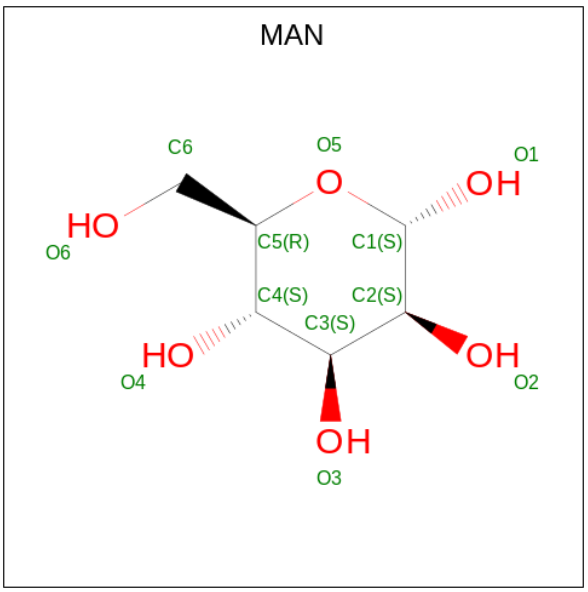
- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Mn	0	0
			1	1		
2	B	1	Total	Mn	0	0
			1	1		
2	C	1	Total	Mn	0	0
			1	1		
2	D	1	Total	Mn	0	0
			1	1		

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

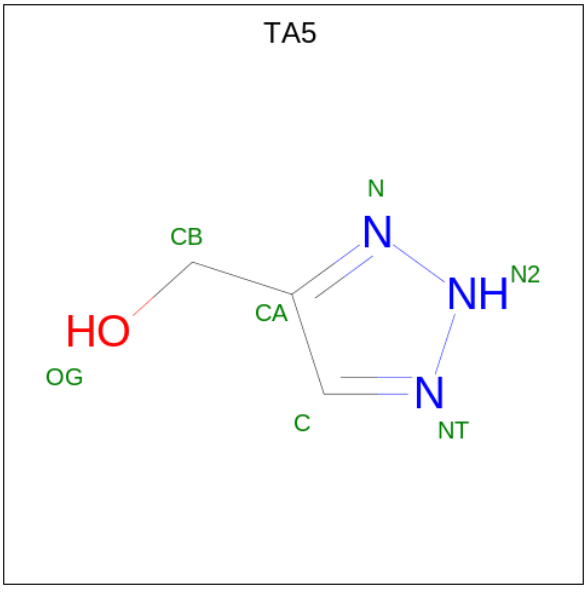
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total	Ca	0	0
			2	2		
3	B	1	Total	Ca	0	0
			1	1		
3	C	1	Total	Ca	0	0
			1	1		
3	D	2	Total	Ca	0	0
			2	2		

- Molecule 4 is alpha-D-mannopyranose (three-letter code: MAN) (formula: $C_6H_{12}O_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	C	1	Total C O 11 6 5	0	0
4	D	1	Total C O 11 6 5	0	0

- Molecule 5 is 2H-1,2,3-TRIAZOL-4-YLMETHANOL (three-letter code: TA5) (formula: $C_3H_5N_3O$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	1	Total C N O 7 3 3 1	0	0
5	D	1	Total C N O 7 3 3 1	0	0

- Molecule 6 is water.

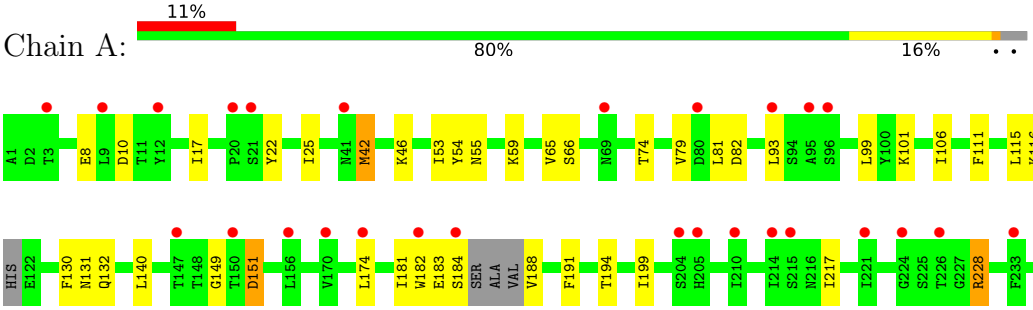
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	57	Total O 57 57	0	0
6	B	64	Total O 64 64	0	0
6	C	58	Total O 58 58	0	0
6	D	65	Total O 65 65	0	0

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.

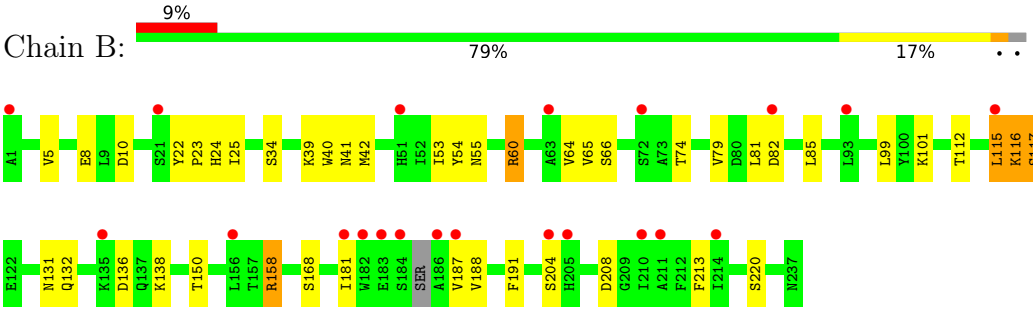
• Molecule 1: Concanavalin-A

Chain A:



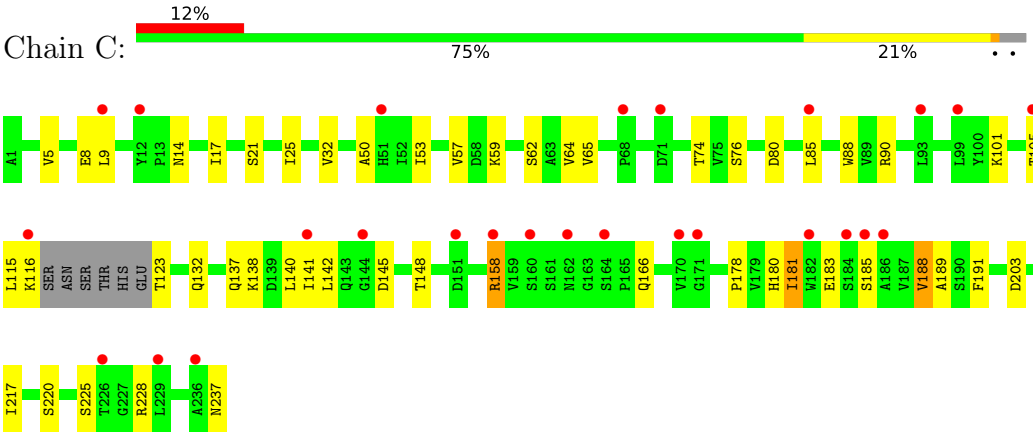
• Molecule 1: Concanavalin-A

Chain B:



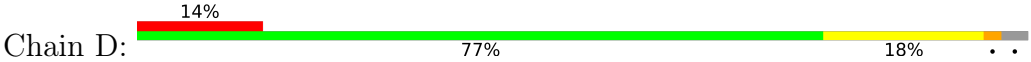
• Molecule 1: Concanavalin-A

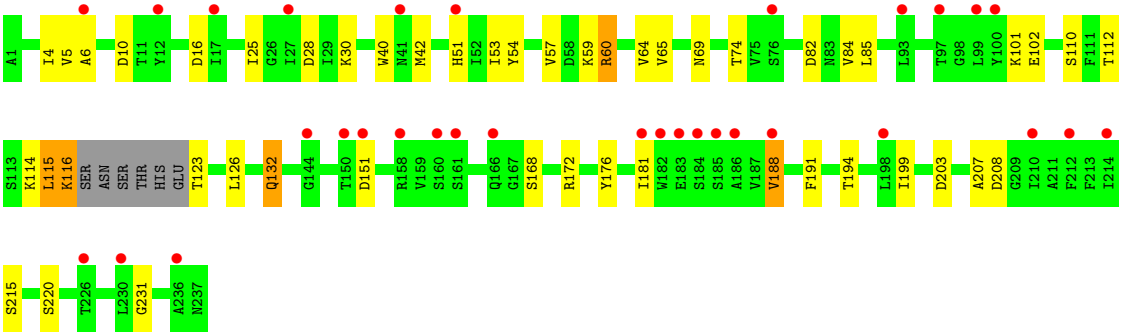
Chain C:



• Molecule 1: Concanavalin-A

Chain D:




A1
I4
V5
A6
D10
T11
Y12
D16
I17
I25
G26
I27
D28
I29
K30
W40
N41
M42
H51
I52
I53
Y54
V57
D58
K59
R60
V64
V65
M69
T74
V75
S76
D82
N83
V84
L85
L93
T97
G98
L99
Y100
K101
E102
S110
F111
T112
S113
K114
L115
K116
SER
ASN
SER
THR
HIS
GLU
T123
L126
Q132
G144
T150
D151
R158
V159
S160
S161
Q166
G167
S168
R172
Y176
I181
W182
E183
S184
S185
A186
V187
V188
F191
T194
L198
I199
D203
A207
D208
G209
I210
A211
F212
F213
I214
S215
S220
T226
L230
G231
A236
N237

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	64.23Å 119.89Å 125.21Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.59 29.83 – 2.50	Depositor EDS
% Data completeness (in resolution range)	94.2 (30.00-2.59) 94.3 (29.83-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.37 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.8.0131	Depositor
R, R_{free}	0.278 , 0.372 0.280 , 0.368	Depositor DCC
R_{free} test set	1619 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	25.3	Xtriage
Anisotropy	0.203	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 36.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.000 for -h,l,k	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	7348	wwPDB-VP
Average B, all atoms (Å^2)	27.0	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 66.37 % 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 6.1124e-06. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: CA, MAN, TA5, MN

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	0/1799	0.85	1/2447 (0.0%)
1	B	0.67	0/1811	0.84	1/2464 (0.0%)
1	C	0.67	0/1803	0.84	0/2455
1	D	0.66	0/1803	0.82	0/2455
All	All	0.67	0/7216	0.84	2/9821 (0.0%)

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	0	1
1	B	0	1
1	D	0	1
All	All	0	3

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	228	ARG	NE-CZ-NH2	5.84	123.22	120.30
1	B	79	VAL	CB-CA-C	-5.43	101.07	111.40

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	116	LYS	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	B	150	THR	Peptide
1	D	115	LEU	Peptide

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	1760	0	1709	25	0
1	B	1772	0	1723	25	0
1	C	1763	0	1718	27	0
1	D	1763	0	1718	28	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	2	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	2	0	0	0	0
4	C	11	0	10	0	0
4	D	11	0	10	0	0
5	C	7	0	2	0	0
5	D	7	0	2	0	0
6	A	57	0	0	5	0
6	B	64	0	0	8	0
6	C	58	0	0	6	0
6	D	65	0	0	7	0
All	All	7348	0	6892	100	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 7.

The worst 5 of 100 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:D:110:SER:OG	1:D:194:THR:OG1	1.80	0.98

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:5:VAL:HG23	1:D:85:LEU:HD21	1.53	0.90
1:C:88:TRP:HB3	1:C:217:ILE:HD11	1.56	0.86
1:D:116:LYS:HG3	1:D:188:VAL:O	1.82	0.80
1:B:5:VAL:HG23	1:B:85:LEU:HD21	1.62	0.80

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	224/237 (94%)	214 (96%)	10 (4%)	0	100	100
1	B	226/237 (95%)	208 (92%)	17 (8%)	1 (0%)	34	57
1	C	227/237 (96%)	212 (93%)	14 (6%)	1 (0%)	34	57
1	D	227/237 (96%)	216 (95%)	11 (5%)	0	100	100
All	All	904/948 (95%)	850 (94%)	52 (6%)	2 (0%)	47	71

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	41	ASN
1	C	225	SER

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	197/203 (97%)	184 (93%)	13 (7%)	16 33
1	B	198/203 (98%)	178 (90%)	20 (10%)	7 14
1	C	197/203 (97%)	179 (91%)	18 (9%)	9 18
1	D	197/203 (97%)	181 (92%)	16 (8%)	11 23
All	All	789/812 (97%)	722 (92%)	67 (8%)	10 21

5 of 67 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	101	LYS
1	D	115	LEU
1	D	203	ASP
1	B	117	SER
1	B	116	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	132	GLN
1	D	153	ASN
1	D	237	ASN
1	D	166	GLN
1	C	104	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 14 ligands modelled in this entry, 10 are monoatomic - leaving 4 for Mogul analysis.

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	$\#\|Z\| > 2$
4	MAN	D	304	5	11,11,12	0.55	0	15,15,17	0.99	1 (6%)
5	TA5	D	305	4	6,7,7	1.38	1 (16%)	4,8,8	2.11	2 (50%)
4	MAN	C	303	5	11,11,12	0.61	0	15,15,17	1.35	1 (6%)
5	TA5	C	304	4	6,7,7	1.41	1 (16%)	4,8,8	2.14	2 (50%)

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
4	MAN	D	304	5	-	0/2/19/22	0/1/1/1
5	TA5	D	305	4	-	0/0/2/2	0/1/1/1
4	MAN	C	303	5	-	0/2/19/22	0/1/1/1
5	TA5	C	304	4	-	0/0/2/2	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	305	TA5	N-N2	-2.56	1.31	1.34
5	C	304	TA5	N-N2	-2.49	1.31	1.34

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	C	303	MAN	C1-O5-C5	4.77	118.65	112.19
5	C	304	TA5	N-N2-NT	-3.30	107.13	111.24
5	D	305	TA5	N-N2-NT	-3.26	107.18	111.24
4	D	304	MAN	C1-O5-C5	2.76	115.93	112.19
5	D	305	TA5	C-CA-N	-2.31	107.90	111.34

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)

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	230/237 (97%)	0.98	27 (11%) 4 3	13, 25, 47, 64	0
1	B	232/237 (97%)	0.97	21 (9%) 9 6	13, 26, 48, 61	0
1	C	231/237 (97%)	0.96	28 (12%) 4 2	15, 25, 42, 56	0
1	D	231/237 (97%)	1.03	32 (13%) 2 1	14, 25, 42, 57	0
All	All	924/948 (97%)	0.99	108 (11%) 4 3	13, 25, 46, 64	0

The worst 5 of 108 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	204	SER	5.4
1	C	182	TRP	4.5
1	C	185	SER	4.4
1	D	185	SER	4.4
1	A	204	SER	4.2

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

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
5	TA5	D	305	7/7	0.58	0.40	35,43,47,48	0
4	MAN	D	304	11/12	0.76	0.30	24,28,29,30	0
5	TA5	C	304	7/7	0.77	0.30	30,40,44,44	0
3	CA	A	303	1/1	0.81	0.12	48,48,48,48	0
4	MAN	C	303	11/12	0.82	0.19	24,26,27,27	0
3	CA	A	302	1/1	0.84	0.06	32,32,32,32	0
2	MN	D	301	1/1	0.90	0.10	35,35,35,35	0
3	CA	C	302	1/1	0.90	0.11	33,33,33,33	0
3	CA	D	302	1/1	0.91	0.09	27,27,27,27	0
2	MN	B	301	1/1	0.94	0.06	39,39,39,39	0
3	CA	B	302	1/1	0.95	0.07	25,25,25,25	0
2	MN	C	301	1/1	0.96	0.06	35,35,35,35	0
3	CA	D	303	1/1	0.96	0.07	40,40,40,40	0
2	MN	A	301	1/1	0.97	0.04	43,43,43,43	0

6.5 Other polymers (i)

There are no such residues in this entry.