wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 05:20 PM GMT

PDB ID	:	4HYP
Title	:	Pyrrolopyrimidine inhibitors of dna gyrase b and topoisomerase iv, part i: structure guided discovery and optimization of dual targeting agents with potent, broad-spectrum enzymatic activity.
Authors	:	Bensen, D.C.; Creighton, C.J.; Tari, L.W.
Deposited on	:	2012-11-13
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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

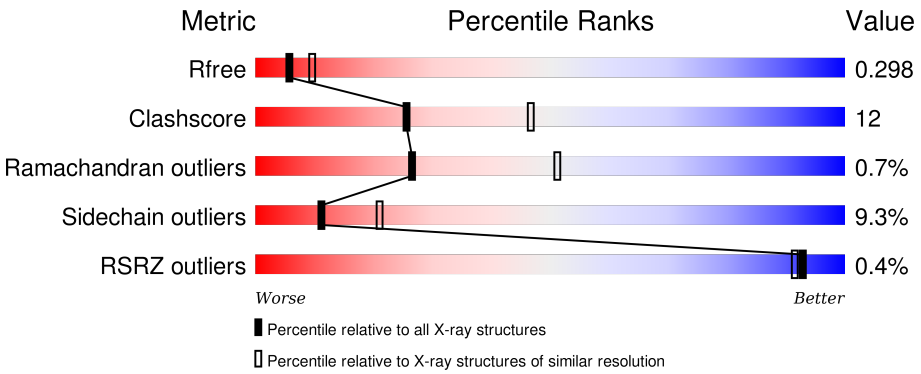
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)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

1 Overall quality at a glance (i)

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}	91344	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (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. 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	215	60% 28% • 10%
1	B	215	65% 23% • 10%
1	C	215	64% 23% • 10%
1	D	215	63% 23% • 10%

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
3	MG	A	302	-	-	-	X
3	MG	C	302	-	-	-	X
3	MG	D	302	-	-	-	X

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 6475 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 DNA gyrase subunit B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	194	1498	941	268	284	5	0	0	0
1	B	194	1498	941	268	284	5	0	0	0
1	C	194	1498	941	268	284	5	0	0	0
1	D	194	1498	941	268	284	5	0	0	0

There are 36 discrepancies between the modelled and reference sequences:

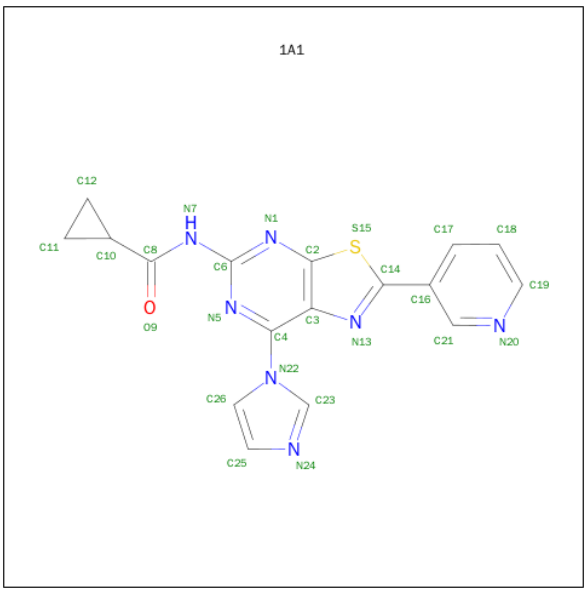
Chain	Residue	Modelled	Actual	Comment	Reference
A	14	MET	-	EXPRESSION TAG	UNP P0AES6
A	225	LEU	-	EXPRESSION TAG	UNP P0AES6
A	226	GLU	-	EXPRESSION TAG	UNP P0AES6
A	227	HIS	-	EXPRESSION TAG	UNP P0AES6
A	228	HIS	-	EXPRESSION TAG	UNP P0AES6
A	229	HIS	-	EXPRESSION TAG	UNP P0AES6
A	230	HIS	-	EXPRESSION TAG	UNP P0AES6
A	231	HIS	-	EXPRESSION TAG	UNP P0AES6
A	232	HIS	-	EXPRESSION TAG	UNP P0AES6
B	14	MET	-	EXPRESSION TAG	UNP P0AES6
B	225	LEU	-	EXPRESSION TAG	UNP P0AES6
B	226	GLU	-	EXPRESSION TAG	UNP P0AES6
B	227	HIS	-	EXPRESSION TAG	UNP P0AES6
B	228	HIS	-	EXPRESSION TAG	UNP P0AES6
B	229	HIS	-	EXPRESSION TAG	UNP P0AES6
B	230	HIS	-	EXPRESSION TAG	UNP P0AES6
B	231	HIS	-	EXPRESSION TAG	UNP P0AES6
B	232	HIS	-	EXPRESSION TAG	UNP P0AES6
C	14	MET	-	EXPRESSION TAG	UNP P0AES6
C	225	LEU	-	EXPRESSION TAG	UNP P0AES6
C	226	GLU	-	EXPRESSION TAG	UNP P0AES6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	227	HIS	-	EXPRESSION TAG	UNP P0AES6
C	228	HIS	-	EXPRESSION TAG	UNP P0AES6
C	229	HIS	-	EXPRESSION TAG	UNP P0AES6
C	230	HIS	-	EXPRESSION TAG	UNP P0AES6
C	231	HIS	-	EXPRESSION TAG	UNP P0AES6
C	232	HIS	-	EXPRESSION TAG	UNP P0AES6
D	14	MET	-	EXPRESSION TAG	UNP P0AES6
D	225	LEU	-	EXPRESSION TAG	UNP P0AES6
D	226	GLU	-	EXPRESSION TAG	UNP P0AES6
D	227	HIS	-	EXPRESSION TAG	UNP P0AES6
D	228	HIS	-	EXPRESSION TAG	UNP P0AES6
D	229	HIS	-	EXPRESSION TAG	UNP P0AES6
D	230	HIS	-	EXPRESSION TAG	UNP P0AES6
D	231	HIS	-	EXPRESSION TAG	UNP P0AES6
D	232	HIS	-	EXPRESSION TAG	UNP P0AES6

- Molecule 2 is N-[7-(1H-IMIDAZOL-1-YL)-2-(PYRIDIN-3-YL)[1,3]THIAZOLO[5,4-D]PYRIMIDIN-5-YL]CYCLOPROPANECARBOXAMIDE (three-letter code: 1A1) (formula: $C_{17}H_{13}N_7OS$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
2	A	1	26	17	7	1	1	0	0
			Total	C	N	O	S		
2	B	1	26	17	7	1	1	0	0
			Total	C	N	O	S		
2	C	1	26	17	7	1	1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
2	D	1	26	17	7	1	1	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 4 is water.

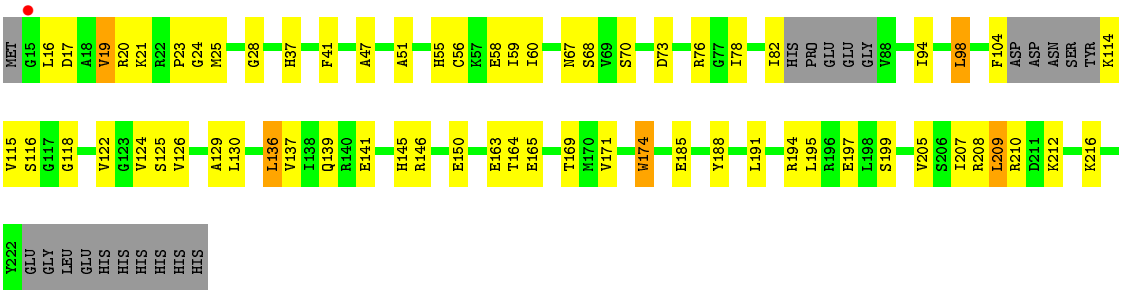
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	90	90	90	0	0
4	B	96	96	96	0	0
4	C	101	101	101	0	0
4	D	88	88	88	0	0

3 Residue-property plots (i)

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.

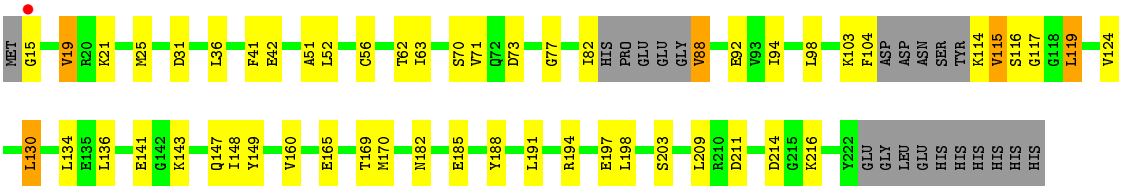
- Molecule 1: DNA gyrase subunit B

Chain A:



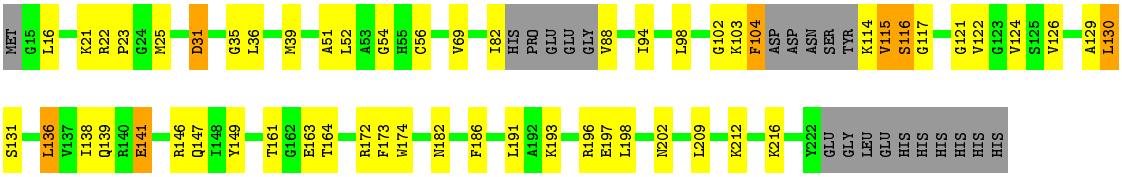
- Molecule 1: DNA gyrase subunit B

Chain B:



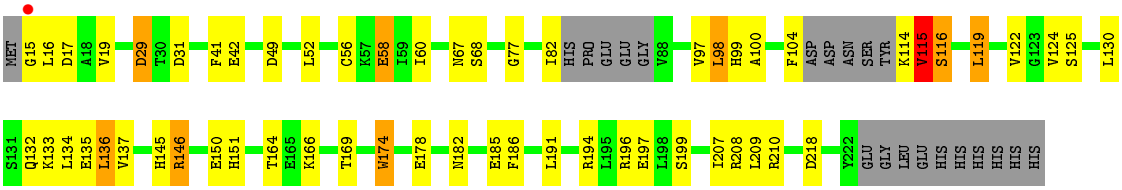
- Molecule 1: DNA gyrase subunit B

Chain C:



- Molecule 1: DNA gyrase subunit B

Chain D:


MET
G15
L16
D17
A18
V19
D29
T30
D31
F41
E42
D49
L52
C56
K57
E58
I59
I60
N67
S68
G77
I82
HIS
PRO
GLU
GLU
GLY
V88
V97
I98
H99
A100
F104
ASP
ASP
ASN
SER
TYR
K114
V115
S116
L119
V122
G123
V124
S125
L130
S131
Q132
K133
L134
E135
L136
V137
H145
R146
E150
H151
T164
E165
K166
T169
W174
E178
N182
E185
F186
L191
R194
L195
R196
E197
L198
S199
I207
R208
L209
R210
D218
Y222
GLU
GLY
GLY
LEU
GLU
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	48.41Å 101.31Å 82.28Å 90.00° 101.60° 90.00°	Depositor
Resolution (Å)	42.89 – 2.60 42.89 – 2.60	Depositor EDS
% Data completeness (in resolution range)	96.3 (42.89-2.60) 96.3 (42.89-2.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.87 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.195 , 0.294 0.200 , 0.298	Depositor DCC
R_{free} test set	1230 reflections (5.61%)	DCC
Wilson B-factor (Å^2)	15.1	Xtriage
Anisotropy	0.115	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 45.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	1 of 23127 reflections (0.004%)	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6475	wwPDB-VP
Average B, all atoms (Å^2)	30.0	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 71.34 % 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 2.6731e-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, 0.375 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: MG, 1A1

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	1/1523 (0.1%)	0.79	2/2054 (0.1%)
1	B	0.67	0/1523	0.81	0/2054
1	C	0.67	1/1523 (0.1%)	0.81	0/2054
1	D	0.66	0/1523	0.82	1/2054 (0.0%)
All	All	0.67	2/6092 (0.0%)	0.81	3/8216 (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	D	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	174	TRP	CD2-CE2	5.88	1.48	1.41
1	C	174	TRP	CD2-CE2	5.38	1.47	1.41

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	136	LEU	CA-CB-CG	6.25	129.68	115.30
1	A	136	LEU	CA-CB-CG	6.06	129.24	115.30
1	A	116	SER	N-CA-C	5.37	125.49	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	116	SER	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	1498	0	1487	38	0
1	B	1498	0	1487	36	0
1	C	1498	0	1487	43	0
1	D	1498	0	1487	42	0
2	A	26	0	13	2	0
2	B	26	0	13	2	0
2	C	26	0	13	2	0
2	D	26	0	13	2	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	90	0	0	4	0
4	B	96	0	0	1	0
4	C	101	0	0	9	0
4	D	88	0	0	9	0
All	All	6475	0	6000	151	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 12.

The worst 5 of 151 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:52:LEU:HD13	1:C:115:VAL:HG21	1.46	0.98
1:D:31:ASP:HB2	4:D:450:HOH:O	1.67	0.93
1:C:115:VAL:O	1:C:116:SER:HB3	1.68	0.91
1:D:99:HIS:HD2	4:D:480:HOH:O	1.55	0.87
1:A:137:VAL:HG22	1:A:146:ARG:HG3	1.54	0.87

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	188/215 (87%)	175 (93%)	13 (7%)	0	100	100
1	B	188/215 (87%)	181 (96%)	5 (3%)	2 (1%)	17	36
1	C	188/215 (87%)	180 (96%)	6 (3%)	2 (1%)	17	36
1	D	188/215 (87%)	178 (95%)	9 (5%)	1 (0%)	34	60
All	All	752/860 (87%)	714 (95%)	33 (4%)	5 (1%)	26	51

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	116	SER
1	B	115	VAL
1	C	115	VAL
1	D	115	VAL
1	B	116	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	161/180 (89%)	145 (90%)	16 (10%)	10	18
1	B	161/180 (89%)	147 (91%)	14 (9%)	13	24
1	C	161/180 (89%)	150 (93%)	11 (7%)	20	39
1	D	161/180 (89%)	142 (88%)	19 (12%)	6	12
All	All	644/720 (89%)	584 (91%)	60 (9%)	11	21

5 of 60 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	165	GLU
1	C	104	PHE
1	D	174	TRP
1	C	21	LYS
1	C	130	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	182	ASN
1	C	145	HIS
1	D	72	GLN
1	B	147	GLN
1	D	67	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)

Of 8 ligands modelled in this entry, 4 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	$\#\|Z\| > 2$
2	1A1	A	301	-	25,30,30	2.28	4 (16%)	29,43,43	2.17	10 (34%)
2	1A1	B	301	-	25,30,30	2.24	3 (12%)	29,43,43	2.81	11 (37%)
2	1A1	C	301	-	25,30,30	2.44	3 (12%)	29,43,43	2.76	12 (41%)
2	1A1	D	301	-	25,30,30	2.31	5 (20%)	29,43,43	2.33	12 (41%)

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	1A1	A	301	-	-	0/13/18/18	0/4/5/5
2	1A1	B	301	-	-	0/13/18/18	0/4/5/5
2	1A1	C	301	-	-	0/13/18/18	0/4/5/5
2	1A1	D	301	-	-	0/13/18/18	0/4/5/5

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	301	1A1	C4-N22	-10.02	1.33	1.46
2	B	301	1A1	C4-N22	-9.23	1.34	1.46
2	D	301	1A1	C4-N22	-9.10	1.34	1.46
2	A	301	1A1	C4-N22	-8.81	1.34	1.46
2	C	301	1A1	C26-N22	-5.04	1.31	1.39

The worst 5 of 45 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	301	1A1	C11-C10-C8	-9.23	108.44	117.26
2	C	301	1A1	C11-C10-C8	-7.82	109.80	117.26
2	D	301	1A1	C12-C10-C8	-7.22	110.36	117.26
2	B	301	1A1	C12-C10-C8	-4.78	112.69	117.26
2	C	301	1A1	N1-C6-N5	-4.74	118.95	126.22

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	301	1A1	2	0
2	B	301	1A1	2	0
2	C	301	1A1	2	0
2	D	301	1A1	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	194/215 (90%)	-0.32	1 (0%)	91	90	21, 30, 43, 57	0
1	B	194/215 (90%)	-0.41	1 (0%)	91	90	21, 29, 43, 54	0
1	C	194/215 (90%)	-0.41	0	100	100	20, 29, 44, 52	0
1	D	194/215 (90%)	-0.32	1 (0%)	91	90	18, 30, 45, 56	0
All	All	776/860 (90%)	-0.36	3 (0%)	93	91	18, 29, 44, 57	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	15	GLY	4.5
1	A	15	GLY	4.2
1	B	15	GLY	2.3

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(Å^2)	Q<0.9
3	MG	A	302	1/1	0.95	0.19	5.32	25,25,25,25	0
3	MG	C	302	1/1	0.88	0.22	3.79	39,39,39,39	0
3	MG	D	302	1/1	0.86	0.18	3.74	27,27,27,27	0
2	1A1	C	301	26/26	0.96	0.16	1.15	21,27,29,32	0
2	1A1	A	301	26/26	0.97	0.13	-0.63	23,25,28,29	0
2	1A1	D	301	26/26	0.96	0.13	-0.79	26,29,31,33	0
2	1A1	B	301	26/26	0.97	0.12	-1.22	24,28,30,32	0
3	MG	B	302	1/1	0.91	0.07	-2.51	26,26,26,26	0

6.5 Other polymers (i)

There are no such residues in this entry.