



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 04:46 pm GMT

PDB ID : 2JLE
Title : NOVEL INDAZOLE NNRTIS CREATED USING MOLECULAR TEMPLATE HYBRIDIZATION BASED ON CRYSTALLOGRAPHIC OVERLAYS
Authors : Jones, L.H.; Allan, G.; Barba, O.; Burt, C.; Corbau, R.; Dupont, T.; Irving, S.; Mowbray, C.E.; Phillips, C.; Swain, N.A.; Webster, R.; Westby, M.
Deposited on : 2008-09-08
Resolution : 2.90 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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : trunk28620
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
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) : recalc28949

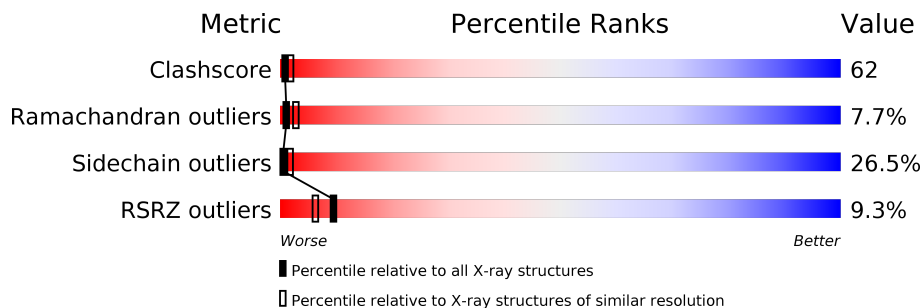
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

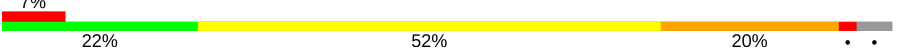

The reported resolution of this entry is 2.90 Å.

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))
Clashscore	112137	1807 (2.90-2.90)
Ramachandran outliers	110173	1768 (2.90-2.90)
Sidechain outliers	110143	1770 (2.90-2.90)
RSRZ outliers	101464	1596 (2.90-2.90)

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	566	
1	B	566	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 8120 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.

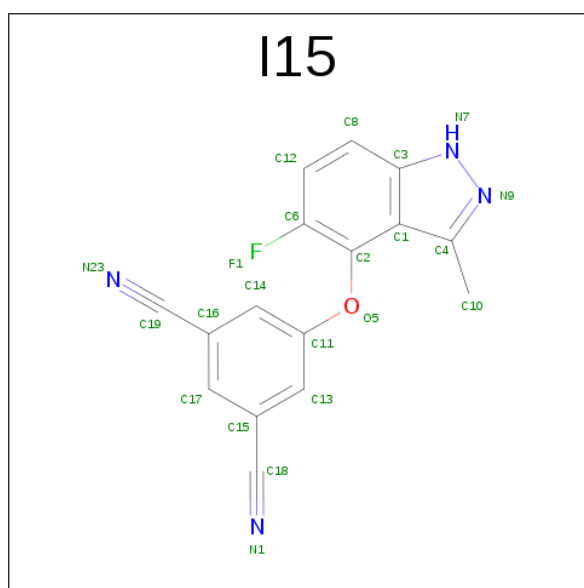
- Molecule 1 is a protein called REVERSE TRANSCRIPTASE/RNASEH.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	545	Total	C	N	O	S	0	0	0
			4445	2875	742	820	8			
1	B	416	Total	C	N	O	S	0	0	1
			3414	2218	569	620	7			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	103	LYS	ARG	CONFLICT	UNP Q72547
B	103	LYS	ARG	CONFLICT	UNP Q72547
A	350	LYS	ARG	CONFLICT	UNP Q72547
B	350	LYS	ARG	CONFLICT	UNP Q72547

- Molecule 2 is 5-[(5-FLUORO-3-METHYL-1H-INDAZOL-4-YL)OXY]BENZENE-1,3-DICARBONITRILE (three-letter code: I15) (formula: C₁₆H₉FN₄O).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	F	N	O	0	0
			22	16	1	4	1		

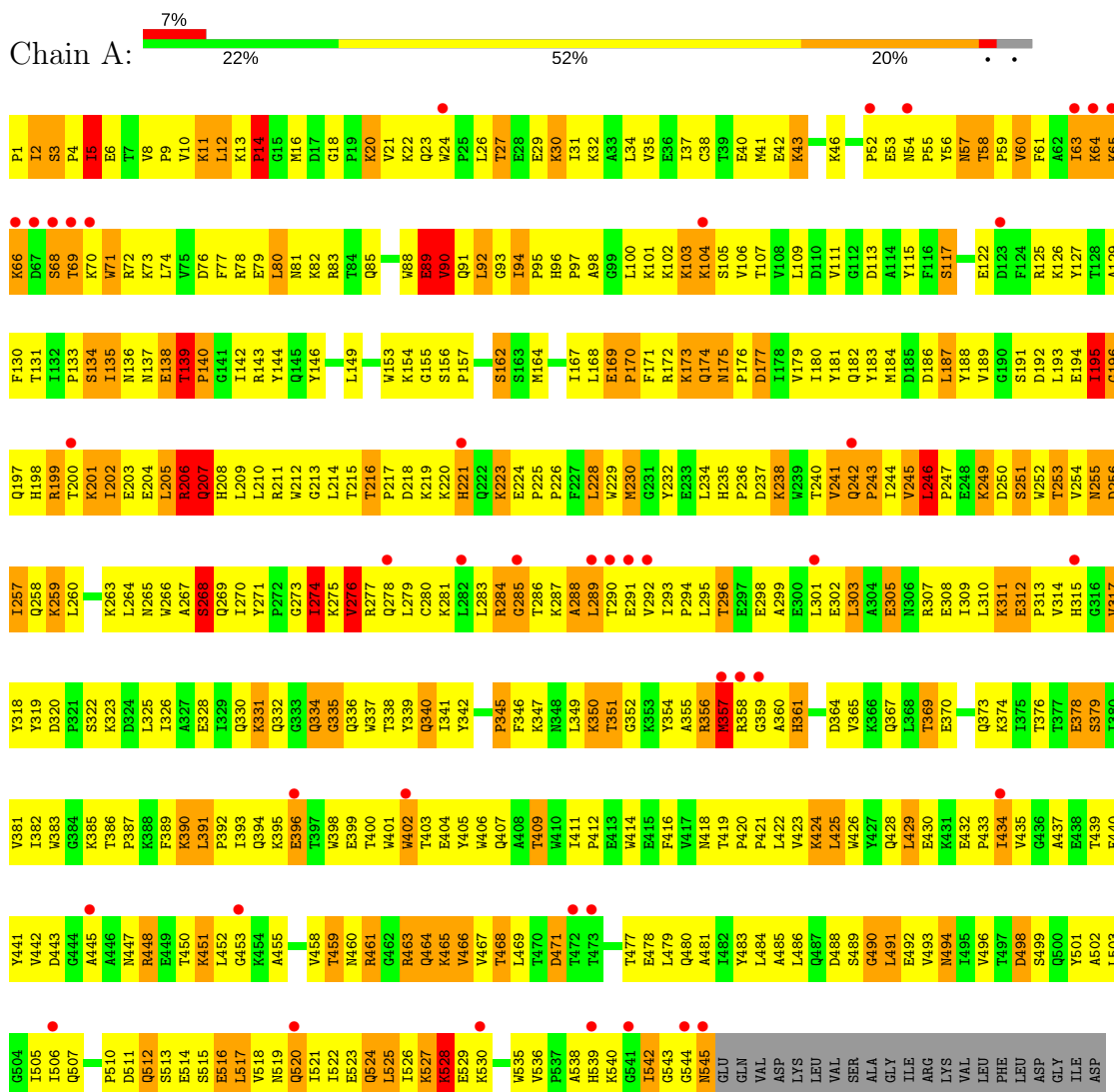
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	152	Total	O	0	0
			152	152		
3	B	87	Total	O	0	0
			87	87		

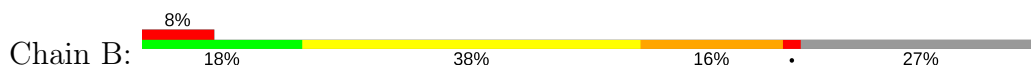
3 Residue-property plots

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 ($RSRZ > 2$). 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: REVERSE TRANSCRIPTASE/RNASEH



• Molecule 1: REVERSE TRANSCRIPTASE/RNASEH



ASP	ILE	ALA	Y318	T257	D192	A129	K65	PR0
GLN	ASN	ASN	Y319	Q258	L193	I132	K66	ILE
ALA	ARG	GLU	D320	K259	E194	P133	S3	Y
GLN	THR	THR	P321	L260	I195	S68	P4	Y
ASP	LYS	LYS	S322	V261	G196	S196	I5	Y
GLN	LEU	LEU	D324		Q197	I135	E6	Y
SER	GLY	LEU	P387		H198	H136	T7	Y
GLU	LYS	LYS	L325	L264	R199	M137	V8	Y
GLU	LYS	LYS	W266	W265	T200	K73		Y
GLU	ALA	ALA	E327	A267	K201	R72	P9	Y
GLY	GLY	GLY	A328	S268	I202	T139	V10	Y
LEU	TYR	TYR	E328	Q269	G141	D76	K11	Y
VAL	VAL	VAL	K331	E204	I142	F77	L12	Y
ASN	THR	THR	Q332	L205	R143	R78	K13	Y
ASN	ASN	ASN	G333	R206	Y144	E79	P14	Y
ILE	ARG	ARG	Q334	Q207	Q145	L80	G15	Y
ILE	GLY	GLY	G335	H208	V148	N81	D17	Y
GLY	ARG	ARG	Q336	L274	L149	K82	G18	Y
GLN	GLN	GLN	Y337	V276	P150	R83	P19	Y
LYS	LYS	LYS	T338	R277	T210	T84	K20	Y
VAL	VAL	VAL	Y339	Q278	Q151	Q85	P21	Y
LYS	VAL	VAL	Q340	L279	G152	F87	L26	Y
LYS	THR	THR	Y341	C280	W153	N87	T27	Y
LEU	LEU	LEU	Y342	K281	K154	W88	T27	Y
LYS	THR	THR	Q343	L282	G155	E89	E28	Y
ASP	ASP	ASP	E344	L283	S156	V90	E29	Y
THR	THR	THR	P345	R284	P157	Q91	K30	Y
LEU	THR	THR	F346	G285	A158	L92	I31	Y
ALA	ASN	ASN	K347	T286	I159	G93	I32	Y
GLN	GLN	GLN	N348	K287	F160	I94	K33	Y
LYS	LYS	LYS	I349	A288	Q161	P95	A33	Y
THR	THR	THR	K350	L289	S162	H96	L34	Y
ALA	GLU	GLU	T351	T290	M164	P97	V35	Y
LYS	LEU	LEU	G352	E291	T165	L100	I37	Y
GLN	GLN	GLN	K353	V292	K166	K101	E36	Y
LYS	LYS	LYS	Y354	I293	I167	K102	C38	Y
ALA	ALA	ALA	A355	P294	L168	K103	T39	Y
ILE	ILE	ILE	R356	L295	E169	E40	E40	Y
TYR	TYR	TYR	K357	T296	P170	M41	M41	Y
LEU	LEU	LEU	R358	E297	F171	L109	K43	Y
ALA	ALA	ALA	G359	E298	R172	D110	E44	Y
GLN	GLN	GLN	A360	A299	K173	G111	G45	Y
GLN	GLN	GLN	H361	E300	Q174	G112	I47	Y
ASP	ASP	ASP	T362	L301	M175	D113	S48	Y
SER	SER	SER	K363	E302	I178	A114	K49	Y
GLY	GLY	GLY	D364	E303	V179	Y115	I50	Y
LEU	LEU	LEU	V365	A304	I180	S117	G51	Y
VAL	VAL	VAL	K366	E305	Y181	V118	P52	Y
VAL	VAL	VAL	Q367	N306	Q182	P119	E53	Y
ILE	ILE	ILE	I368	R307	Y183	L120	N54	Y
VAL	VAL	VAL	T369	E308	M184	D121	P55	Y
ARG	ARG	ARG	E370	I309	D185	E122	T58	Y
ALA	ALA	ALA	A371	L310	D186	F124	P59	Y
LYS	LYS	LYS	THR	K311	D250	R125	A62	Y
VAL	VAL	VAL	PHE	E312	W251	F126	I63	Y
LEU	LEU	LEU	TYR	P313	T253	Y127	I63	Y
PHE	PHE	PHE	VAL	V314	V254		I63	Y
LEU	LEU	LEU	T376	H315	N255		I63	Y
ASP	ASP	ASP	T377	G316	G190		I63	Y
GLY	GLY	GLY	S379	V317	D256		I63	Y
ILE	ILE	ILE					I63	Y

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	117.20Å 154.60Å 155.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	14.13 – 2.90 29.94 – 2.90	Depositor EDS
% Data completeness (in resolution range)	96.7 (14.13-2.90) 90.4 (29.94-2.90)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.24 (at 2.90Å)	Xtriage
Refinement program	BUSTER/TNT	Depositor
R, R_{free}	0.261 , 0.351 0.276 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	66.3	Xtriage
Anisotropy	0.213	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 84.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.58$, $\langle L^2 \rangle = 0.44$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	8120	wwPDB-VP
Average B, all atoms (Å ²)	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.04% 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 ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: I15

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	0.90	3/4562 (0.1%)	1.08	8/6199 (0.1%)
1	B	0.90	0/3510	1.07	8/4772 (0.2%)
All	All	0.90	3/8072 (0.0%)	1.08	16/10971 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	516	GLU	CG-CD	6.18	1.61	1.51
1	A	516	GLU	CB-CG	5.39	1.62	1.52
1	A	432	GLU	CG-CD	5.20	1.59	1.51

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	420	PRO	C-N-CD	-9.69	99.27	120.60
1	A	139	THR	C-N-CD	-9.58	99.53	120.60
1	B	312	GLU	C-N-CD	-9.18	100.40	120.60
1	B	344	GLU	C-N-CD	-7.69	103.68	120.60
1	B	132	ILE	C-N-CD	-7.13	104.90	120.60

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	4445	0	4493	620	0
1	B	3414	0	3443	399	0
2	A	22	0	9	1	0
3	A	152	0	0	25	0
3	B	87	0	0	10	0
All	All	8120	0	7945	989	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 62.

The worst 5 of 989 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:A:104:LYS:HB3	1:A:192:ASP:HA	1.20	1.18
1:B:103:LYS:HE3	1:B:179:VAL:HG21	1.24	1.15
1:A:64:LYS:HE3	1:A:69:THR:HA	1.22	1.10
1:A:174:GLN:HA	1:A:174:GLN:HE21	1.14	1.07
1:A:288:ALA:HB3	1:A:291:GLU:HB2	1.33	1.06

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	543/566 (96%)	415 (76%)	86 (16%)	42 (8%)	1	3
1	B	412/566 (73%)	338 (82%)	42 (10%)	32 (8%)	1	3
All	All	955/1132 (84%)	753 (79%)	128 (13%)	74 (8%)	1	3

5 of 74 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2	ILE
1	A	14	PRO
1	A	90	VAL
1	A	135	ILE
1	A	139	THR

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	487/505 (96%)	356 (73%)	131 (27%)	0	2
1	B	375/505 (74%)	278 (74%)	97 (26%)	0	2
All	All	862/1010 (85%)	634 (74%)	228 (26%)	0	2

5 of 228 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	428	GLN
1	A	520	GLN
1	B	353	LYS
1	A	439	THR
1	A	468	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	367	GLN
1	A	494	ASN
1	B	255	ASN
1	A	480	GLN
1	A	507	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](#)

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](#)

1 ligand is 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	I15	A	1546	-	21,24,24	2.41	7 (33%)	25,34,34	2.11	10 (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	I15	A	1546	-	-	0/8/8/8	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1546	I15	C10-C4	-7.26	1.45	1.50
2	A	1546	I15	C8-C3	-2.85	1.36	1.41
2	A	1546	I15	C13-C15	-2.78	1.34	1.39
2	A	1546	I15	C4-N9	2.51	1.37	1.33
2	A	1546	I15	C12-C6	3.14	1.44	1.37

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1546	I15	C12-C6-C2	-3.36	115.23	122.81
2	A	1546	I15	C16-C14-C11	-3.26	115.21	119.24
2	A	1546	I15	C13-C15-C18	-2.17	116.90	119.53
2	A	1546	I15	C15-C13-C11	2.27	122.06	119.24
2	A	1546	I15	C11-O5-C2	2.35	122.50	118.54

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1546	I15	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	545/566 (96%)	0.40	42 (7%) 14 10	34, 59, 86, 107	0
1	B	416/566 (73%)	0.58	47 (11%) 6 4	36, 58, 95, 119	0
All	All	961/1132 (84%)	0.48	89 (9%) 9 6	34, 59, 90, 119	0

The worst 5 of 89 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	357	MET	8.9
1	B	284	ARG	5.8
1	A	285	GLY	4.8
1	A	291	GLU	4.6
1	A	290	THR	4.5

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
2	I15	A	1546	22/22	0.95	0.21	-0.51	42,51,55,57	0

6.5 Other polymers [i](#)

There are no such residues in this entry.