

wwPDB X-ray Structure Validation Summary Report (i)

Oct 21, 2024 – 04:14 PM EDT

PDB ID : 2RGU

Title : Crystal structure of complex of human DPP4 and inhibitor

Authors: Nar, H.; Himmelsbach, F.; Eckhardt, M.

Deposited on : 2007-10-05

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
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 : 2022.3.0, CSD as543be (2022)

Xtriage (Phenix) : 1.20.1

EDS : 3.0

buster-report : 1.1.7 (2018)

Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)

CCP4 : 9.0.003 (Gargrove)

Density-Fitness : 1.0.11

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

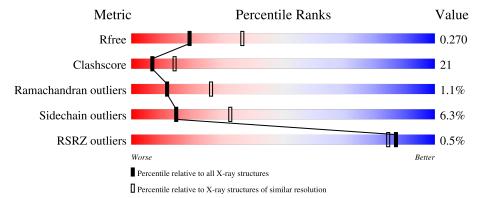
Validation Pipeline (wwPDB-VP) : 2.39

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	Similar resolution
Metric	$(\# ext{Entries})$	$(\# ext{Entries}, ext{ resolution range}(ext{Å}))$
R_{free}	164625	3775 (2.60-2.60)
Clashscore	180529	4181 (2.60-2.60)
Ramachandran outliers	177936	4129 (2.60-2.60)
Sidechain outliers	177891	4129 (2.60-2.60)
RSRZ outliers	164620	3775 (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 chair	n	
1	A	734	60%	35%	. .
1	В	734	64%	32%	

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
2	NAG	A	795	X	-	-	-
2	NAG	A	796	X	-	-	-
2	NAG	В	793	X	-	-	-
2	NAG	В	794	X	-	-	-
2	NAG	В	796	X	-	-	-



2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 12323 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 Dipeptidyl peptidase 4.

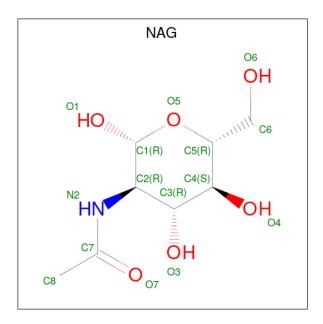
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	A	728	Total 5963	C 3827	N 982	O 1128	S 26	0	0	0
1	В	728	Total 5963	C 3827	N 982	O 1128	S 26	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	767	HIS	=	expression tag	UNP P27487
A	768	HIS	-	expression tag	UNP P27487
A	769	HIS	-	expression tag	UNP P27487
A	770	HIS	ı	expression tag	UNP P27487
A	771	HIS	-	expression tag	UNP P27487
A	772	HIS	-	expression tag	UNP P27487
В	767	HIS	-	expression tag	UNP P27487
В	768	HIS	-	expression tag	UNP P27487
В	769	HIS	ı	expression tag	UNP P27487
В	770	HIS	-	expression tag	UNP P27487
В	771	HIS	-	expression tag	UNP P27487
В	772	HIS	-	expression tag	UNP P27487

• Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

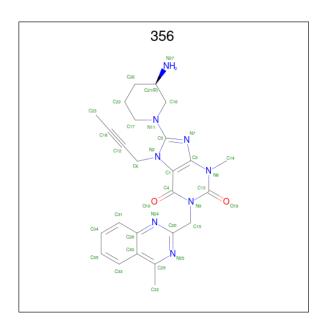




Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C N O 15 8 1 6	0	0
2	A	1	Total C N O 15 8 1 6	0	0
2	A	1	Total C N O 15 8 1 6	0	0
2	A	1	Total C N O 15 8 1 6	0	0
2	В	1	Total C N O 15 8 1 6	0	0
2	В	1	Total C N O 15 8 1 6	0	0
2	В	1	Total C N O 15 8 1 6	0	0
2	В	1	Total C N O 15 8 1 6	0	0

• Molecule 3 is 8-[(3R)-3-Aminopiperidin-1-yl]-7-but-2-yn-1-yl-3-methyl-1-[(4-methylqui nazolin-2-yl)methyl]-3,7-dihydro-1H-purine-2,6-d ione (three-letter code: 356) (formula: $C_{25}H_{28}N_8O_2$).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 35			0	0
3	В	1	Total 35			0	0

• Molecule 4 is water.

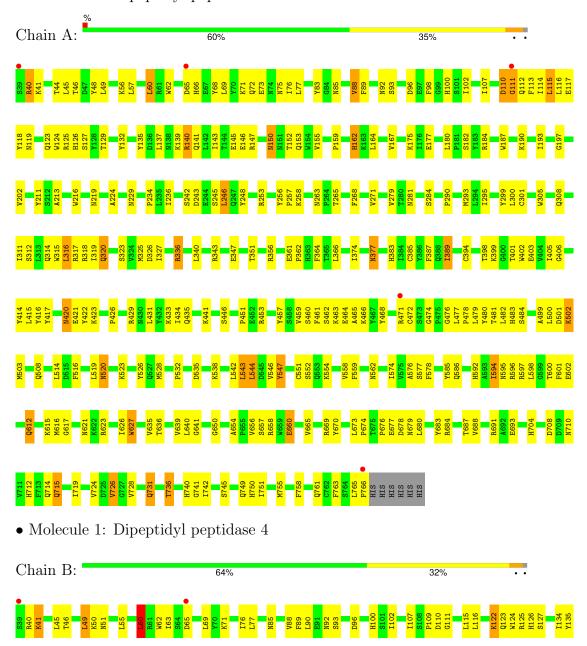
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	96	Total O 96 96	0	0
4	В	111	Total O 111 111	0	0



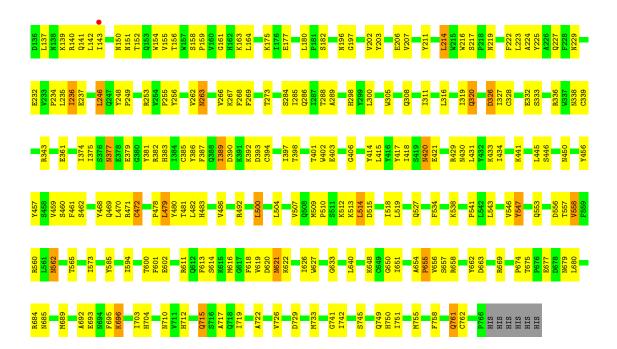
3 Residue-property plots (i)

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: Dipeptidyl peptidase 4









4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants	65.30Å 67.10Å 419.90Å	Depositor
a, b, c, α , β , γ	90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 - 2.60	Depositor
resolution (A)	40.00 - 2.60	EDS
% Data completeness	97.5 (40.00-2.60)	Depositor
(in resolution range)	97.5 (40.00-2.60)	EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	2.58 (at 2.61Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.217 , 0.276	Depositor
it, it free	0.213 , 0.270	DCC
R_{free} test set	2874 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å ²)	53.5	Xtriage
Anisotropy	0.654	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.32, 35.7	EDS
L-test for twinning ²	$< L > = 0.48, < L^2> = 0.31$	Xtriage
Estimated twinning fraction	0.034 for k,h,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	12323	wwPDB-VP
Average B, all atoms (\mathring{A}^2)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 2.67% of the height of the origin peak. No significant pseudotranslation is detected.

²Theoretical 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.



¹Intensities estimated from amplitudes.

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: 356, NAG

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).

Mal	Mol Chain		lengths	Bond angles		
IVIOI	Chain	RMSZ	# Z > 5	RMSZ	# Z > 5	
1	A	0.39	0/6135	0.65	0/8344	
1	В	0.41	0/6135	0.67	1/8344 (0.0%)	
All	All	0.40	0/12270	0.66	$1/16688 \; (0.0\%)$	

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
1	В	60	LEU	CA-CB-CG	5.47	127.89	115.30

There are no chirality outliers.

There are no planarity outliers.

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	5963	0	5685	269	0
1	В	5963	0	5685	231	0
2	A	60	0	60	21	0
2	В	60	0	60	14	0
3	A	35	0	28	0	0
3	В	35	0	28	1	0
4	A	96	0	0	28	0
4	В	111	0	0	18	0



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	12323	0	11546	492	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 21.

The worst 5 of 492 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$\begin{array}{c} {\rm Interatomic} \\ {\rm distance} \ ({\rm \AA}) \end{array}$	Clash overlap (Å)
1:B:85:ASN:ND2	2:B:794:NAG:H1	1.30	1.42
1:B:85:ASN:HD21	2:B:794:NAG:C1	1.43	1.31
1:A:229:ASN:HD21	2:A:796:NAG:C1	1.72	1.01
1:B:403:GLU:H	1:B:420:ASN:HD21	1.09	0.98
1:A:85:ASN:ND2	2:A:794:NAG:H1	1.79	0.96

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	nalysed Favoured Allowed		Outliers	Percentiles
1	A	726/734~(99%)	657 (90%)	59 (8%)	10 (1%)	9 19
1	В	726/734~(99%)	662 (91%)	58 (8%)	6 (1%)	16 34
All	All	1452/1468~(99%)	1319 (91%)	117 (8%)	16 (1%)	12 26

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	320	GLN
1	В	140	ARG
1	В	333	SER
1	В	393	ASP



Continued from previous page...

Mol	Chain	Res	Type
1	A	111	GLY

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	653/659 (99%)	611 (94%)	42 (6%)	14 32		
1	В	653/659 (99%)	613 (94%)	40 (6%)	15 34		
All	All	1306/1318 (99%)	1224 (94%)	82 (6%)	15 32		

5 of 82 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	В	256	TYR
1	В	547	TYR
1	В	300	LEU
1	В	472	CYS
1	В	655	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 48 such sidechains are listed below:

Mol	Chain	Res	\mathbf{Type}
1	В	123	GLN
1	В	383	HIS
1	В	126	HIS
1	В	263	ASN
1	В	483	HIS

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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

10 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	l Type Chain Re		Res	Link	Вс	ond leng	$_{ m ths}$	Bond angles		
IVIOI	туре	Chain	nes	Lilik	Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	NAG	В	793	-	15,15,15	0.49	0	21,21,21	0.64	0
2	NAG	В	794	-	15,15,15	0.35	0	21,21,21	0.67	0
2	NAG	A	794	-	15,15,15	0.46	0	21,21,21	0.85	1 (4%)
3	356	В	902	-	35,39,39	1.71	9 (25%)	37,57,57	2.90	12 (32%)
2	NAG	A	796	-	15,15,15	0.43	0	21,21,21	0.72	1 (4%)
2	NAG	В	796	-	15,15,15	0.34	0	21,21,21	0.92	1 (4%)
2	NAG	В	797	-	15,15,15	0.48	0	21,21,21	0.71	0
2	NAG	A	793	-	15,15,15	0.30	0	21,21,21	0.65	0
2	NAG	A	795	-	15,15,15	0.42	0	21,21,21	0.57	0
3	356	A	901	_	35,39,39	1.59	9 (25%)	37,57,57	2.82	13 (35%)

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	NAG	В	793	-	1/1/6/7	4/6/26/26	0/1/1/1
2	NAG	В	794	-	1/1/6/7	2/6/26/26	0/1/1/1
2	NAG	A	794	_	-	2/6/26/26	0/1/1/1



Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	356	В	902	-	-	1/6/22/22	0/5/5/5
2	NAG	A	796	-	1/1/6/7	2/6/26/26	0/1/1/1
2	NAG	В	796	-	1/1/6/7	2/6/26/26	0/1/1/1
2	NAG	В	797	-	-	2/6/26/26	0/1/1/1
2	NAG	A	793	-	-	2/6/26/26	0/1/1/1
2	NAG	A	795	-	1/1/6/7	2/6/26/26	0/1/1/1
3	356	A	901	-	-	2/6/22/22	0/5/5/5

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(A)	$Ideal(\AA)$
3	В	902	356	C5-N7	-3.84	1.28	1.35
3	В	902	356	C30-C28	3.45	1.47	1.42
3	A	901	356	C5-N7	-3.11	1.30	1.35
3	В	902	356	C22-C26	3.01	1.60	1.53
3	A	901	356	C15-C20	2.89	1.53	1.50

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
3	В	902	356	C1-C4-N9	7.63	119.83	113.58
3	A	901	356	C1-C4-N9	7.17	119.45	113.58
3	В	902	356	C20-N25-C29	6.53	122.47	117.22
3	В	902	356	C4-N9-C13	-6.44	120.62	125.33
3	A	901	356	C30-C29-N25	-6.37	118.08	122.24

All (5) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	795	NAG	C1
2	A	796	NAG	C1
2	В	793	NAG	C1
2	В	794	NAG	C1
2	В	796	NAG	C1

5 of 21 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	901	356	C18-C12-C6-N2
3	В	902	356	N9-C15-C20-N25



Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	A	793	NAG	C4-C5-C6-O6
2	A	793	NAG	O5-C5-C6-O6
2	В	793	NAG	C4-C5-C6-O6

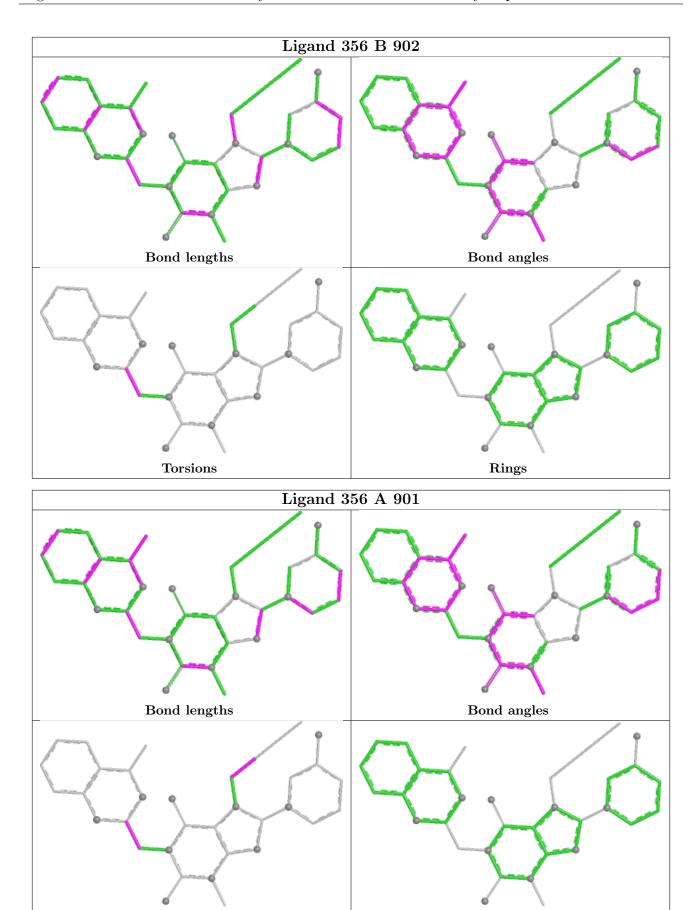
There are no ring outliers.

9 monomers are involved in 36 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	В	793	NAG	2	0
2	В	794	NAG	6	0
2	A	794	NAG	6	0
3	В	902	356	1	0
2	A	796	NAG	6	0
2	В	796	NAG	5	0
2	В	797	NAG	1	0
2	A	793	NAG	3	0
2	A	795	NAG	6	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.







Torsions

Rings

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, 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>	$\#\mathrm{RSRZ}{>}2$	$\mathbf{OWAB}(\mathbf{\mathring{A}}^2)$	Q<0.9
1	A	728/734 (99%)	-0.15	5 (0%) 84 81	36, 55, 82, 98	0
1	В	728/734 (99%)	-0.34	3 (0%) 89 86	34, 48, 72, 86	0
All	All	$1456/1468 \; (99\%)$	-0.25	8 (0%) 87 84	34, 51, 77, 98	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	В	39	SER	2.7
1	В	65	ASP	2.4
1	A	471	ARG	2.3
1	A	766	PRO	2.3
1	A	111	GLY	2.2

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 monosaccharides in this entry.

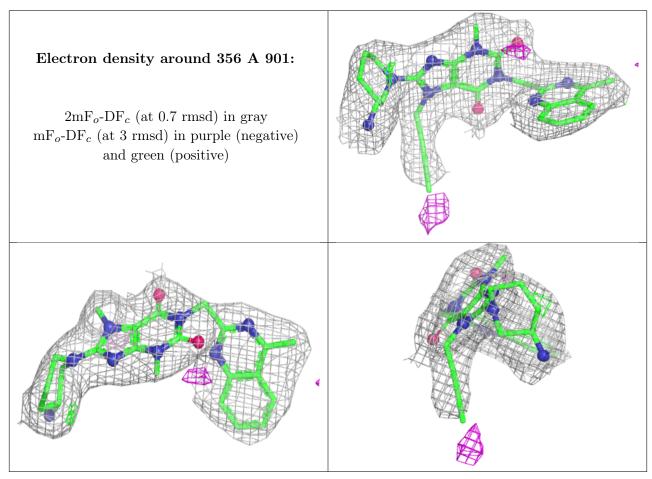
6.4 Ligands (i)

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.

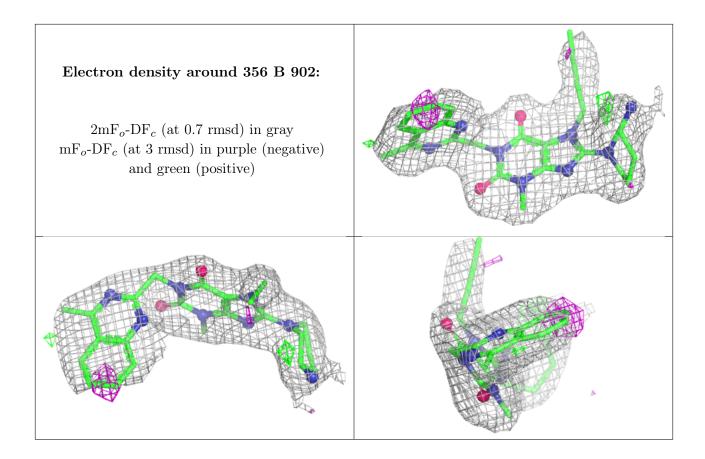


Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\operatorname{B-factors}(\mathring{\mathbf{A}}^2)$	Q<0.9
2	NAG	В	793	15/15	0.72	0.15	98,99,100,100	0
2	NAG	A	795	15/15	0.79	0.12	99,99,100,100	0
2	NAG	В	797	15/15	0.80	0.13	92,93,93,93	0
2	NAG	A	793	15/15	0.81	0.13	95,96,96,97	0
2	NAG	A	796	15/15	0.88	0.12	79,80,81,82	0
2	NAG	В	796	15/15	0.89	0.10	65,65,66,67	0
2	NAG	В	794	15/15	0.89	0.10	50,52,54,57	0
2	NAG	A	794	15/15	0.90	0.08	68,68,69,69	0
3	356	A	901	35/35	0.91	0.10	40,42,45,46	0
3	356	В	902	35/35	0.93	0.10	34,36,45,48	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.







6.5 Other polymers (i)

There are no such residues in this entry.

