



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 16, 2024 – 12:23 AM EDT

PDB ID : 2GBI
Title : rat DPP-IV with xanthine inhibitor 4
Authors : Longenecker, K.L.; Jakob, C.G.; Fry, E.H.; Wilk, S.
Deposited on : 2006-03-10
Resolution : 3.30 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 ⓘ 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 ⓘ](#)) 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	:	2.37.1
buster-report	:	1.1.7 (2018)
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)	:	2.37.1

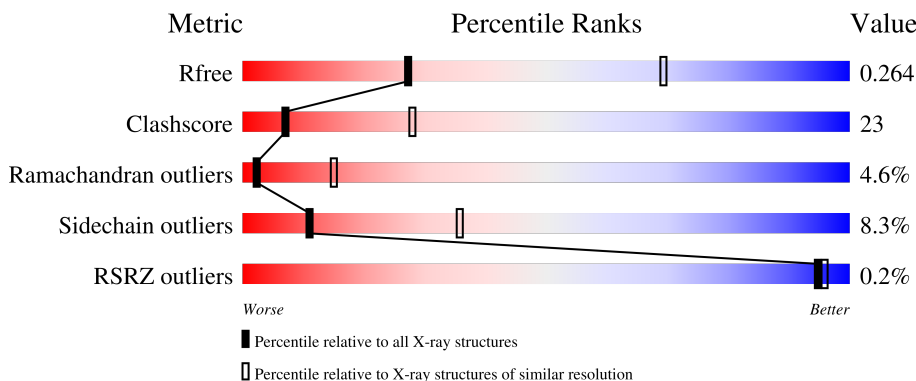
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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 $\leq 5\%$. 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	730	<div> <div>55%</div> <div>37%</div> <div>7%</div> </div>
1	B	730	<div> <div>58%</div> <div>34%</div> <div>8%</div> </div>

2 Entry composition [i](#)

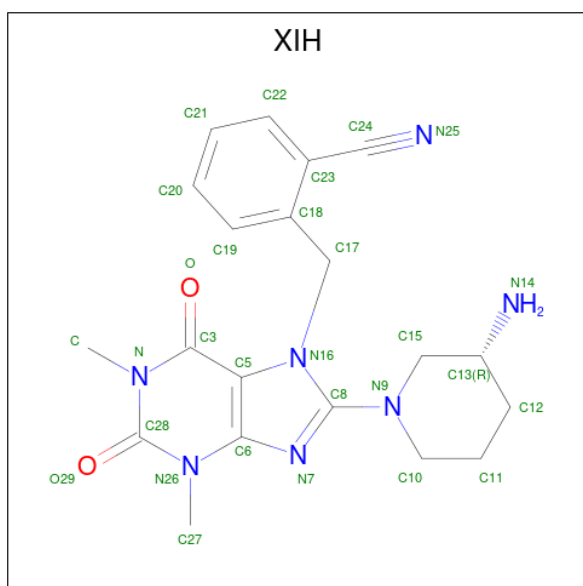
There are 2 unique types of molecules in this entry. The entry contains 11869 atoms, of which 0 are hydrogens and 0 are deuteriums.

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	730	Total	C	N	O	S	0	0	0
			5920	3789	981	1124	26			
1	B	730	Total	C	N	O	S	0	0	0
			5920	3789	981	1124	26			

- Molecule 2 is 2-({8-[(3R)-3-AMINOPIPERIDIN-1-YL]-1,3-DIMETHYL-2,6-DIOXO-1,2,3,6-TETRAHYDRO-7H-PURIN-7-YL}METHYL)BENZONITRILE (three-letter code: XIH) (formula: C₂₀H₂₃N₇O₂).

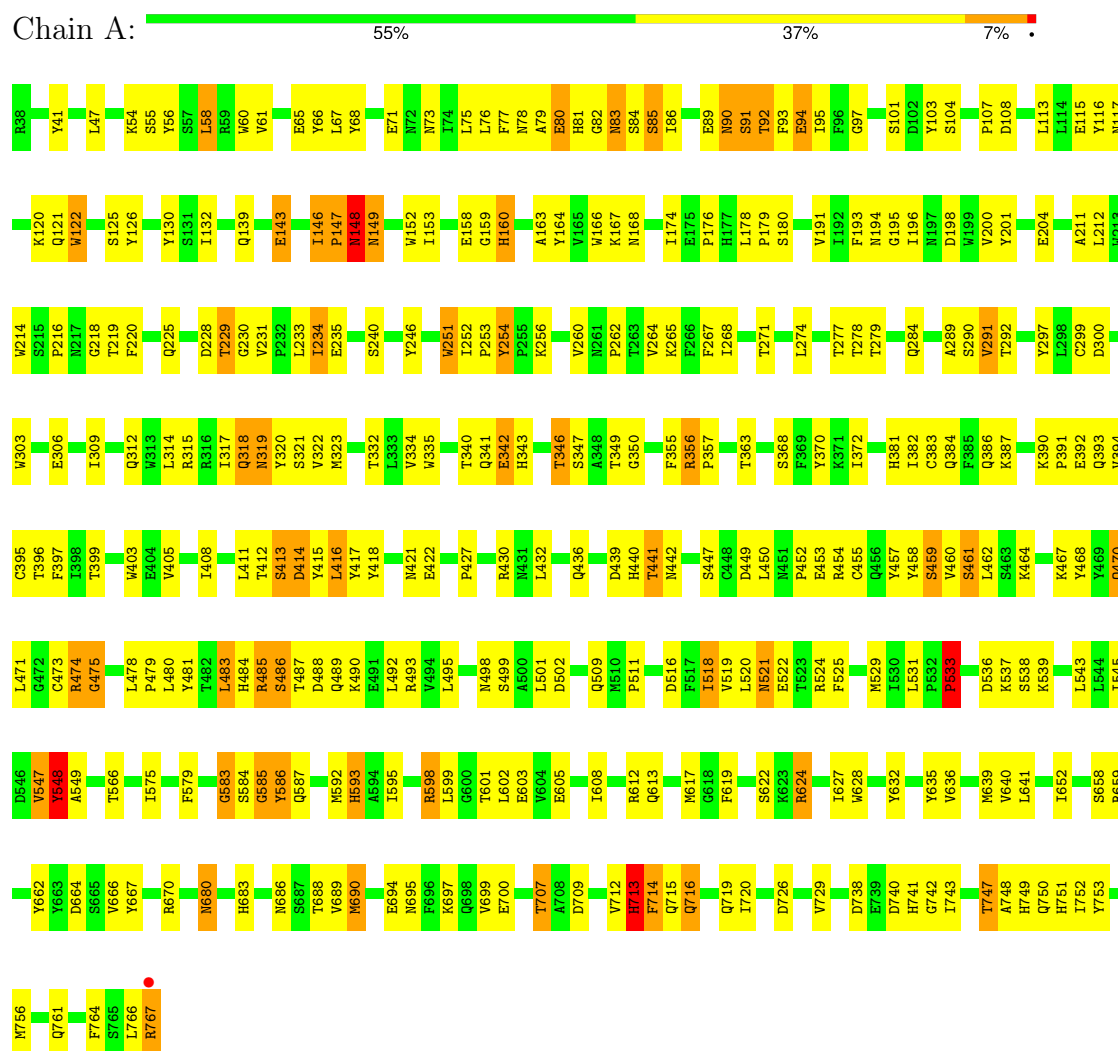


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			29	20	7	2		

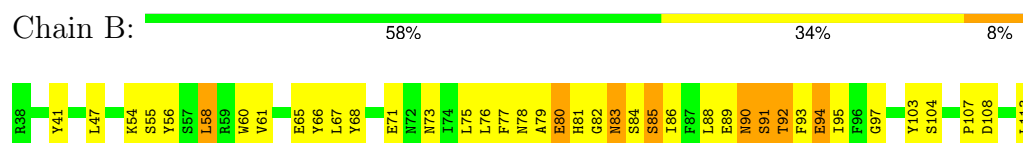
3 Residue-property plots

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



• Molecule 1: Dipeptidyl peptidase 4



T707	A708	D709	H713	F714	Q715	Q716	Q719	I720	W725	D726	W729	D738	E739	D740	H741	G742	I743	A744	S745	S746	T747	A748	H749	Q750	H751	I752	Y753	W756	Q761	F764	S765	L766	R767																				
I595	R598	L599	G600	T601	L602	E603	R612	F619	S622	K623	R624	I627	W628	Y632	Y635	V636	M639	V640	L641	I652	S658	R659	Y662	Y663	D664	S665	W666	R670	I679	N680	H683	N686	S687	T688	W689	E694	K697	W698	E699	W700													
E491	L492	R493	W494	L495	M498	L501	Q509	W510	P511	D516	F517	I518	W519	L520	N521	E522	T523	R524	F525	M529	T530	L531	P532	P533	D536	K537	S538	K539	L543	L544	I545	D546	V547	Y548	A549	T556	I575	F579	G583	S584	G585	Y586	Q587	N592	H593	A594							
N319	Y320	S321	V322	M323	T332	L333	V334	W335	T339	T340	Q341	E342	H343	T346	S347	A348	T349	G350	D348	D449	L450	N451	P452	E453	R454	C455	N456	Y457	Y458	S459	V460	S461	L462	S463	K464	K467	Q470	L471	C472	C473	R474	G475	P479	L480	Y481	T482	L483	H484	R485	S486	T487	D488	Q489
F220	Q225	D228	T229	G230	V231	P232	L233	I234	E235	W251	I252	P253	Y254	P255	K256	A257	V260	N261	P262	T263	V264	K265	F266	P267	I268	T271	L274	T277	T278	T279	Q284	A289	S290	V291	T292	Y297	L298	C299	D300	W303	E306	I309	L314	I317	G218	T219							
Q121	W122	S125	Y126	I132	Q139	E143	I146	P147	N148	N149	W152	I153	E158	G159	H160	A163	Y164	V165	W166	Y171	I174	E175	P176	H177	L178	P179	S180	V191	I192	F193	N194	G195	I196	N197	D198	W199	V200	Y201	A211	L212	W213	W214	S215	P216	R217	G218	T219						

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, α , β , γ	208.55Å 208.55Å 208.55Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.30 29.79 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.9 (20.00-3.30) 99.9 (29.79-3.20)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.59 (at 3.18Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.237 , 0.277 0.230 , 0.264	Depositor DCC
R_{free} test set	2537 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å ²)	58.4	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 47.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.36$	Xtriage
Estimated twinning fraction	0.000 for l,-k,h	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	11869	wwPDB-VP
Average B, all atoms (Å ²)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.82% 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: XIH

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	0.45	0/6088	0.71	2/8278 (0.0%)
1	B	0.43	0/6088	0.71	3/8278 (0.0%)
All	All	0.44	0/12176	0.71	5/16556 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	713	HIS	N-CA-C	6.95	129.77	111.00
1	B	713	HIS	N-CA-C	6.87	129.54	111.00
1	B	548	TYR	N-CA-C	-5.90	95.06	111.00
1	A	548	TYR	N-CA-C	-5.87	95.14	111.00
1	B	383	CYS	CA-CB-SG	5.29	123.53	114.00

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	5920	0	5632	267	0
1	B	5920	0	5632	266	0
2	A	29	0	23	3	0
All	All	11869	0	11287	527	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 23.

The worst 5 of 527 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:174:ILE:HD12	1:A:174:ILE:H	1.20	1.03
1:B:174:ILE:HD12	1:B:174:ILE:H	1.22	1.01
1:A:595:ILE:HD11	1:A:603:GLU:HB3	1.47	0.96
1:B:595:ILE:HD11	1:B:603:GLU:HB3	1.47	0.94
1:A:412:THR:HG22	1:A:414:ASP:H	1.28	0.93

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	728/730 (100%)	614 (84%)	80 (11%)	34 (5%)	2	14
1	B	728/730 (100%)	619 (85%)	76 (10%)	33 (4%)	2	15
All	All	1456/1460 (100%)	1233 (85%)	156 (11%)	67 (5%)	2	15

5 of 67 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	91	SER
1	A	92	THR
1	A	274	LEU
1	A	450	LEU
1	A	475	GLY

5.3.2 Protein sidechains ⓘ

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	643/651 (99%)	590 (92%)	53 (8%)	11	36
1	B	643/651 (99%)	589 (92%)	54 (8%)	11	35
All	All	1286/1302 (99%)	1179 (92%)	107 (8%)	11	36

5 of 107 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	122	TRP
1	B	340	THR
1	B	680	ASN
1	B	143	GLU
1	B	229	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	698	GLN
1	B	713	HIS
1	B	719	GLN
1	A	713	HIS
1	A	698	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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](#)

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	XIH	A	901	-	25,32,32	3.88	6 (24%)	26,47,47	2.38	5 (19%)

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	XIH	A	901	-	-	2/6/20/20	0/4/4/4

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	901	XIH	C6-N7	14.94	1.47	1.33
2	A	901	XIH	C8-N9	8.68	1.46	1.35
2	A	901	XIH	C5-C3	5.95	1.51	1.41
2	A	901	XIH	C23-C18	3.47	1.46	1.40
2	A	901	XIH	C5-C6	-2.36	1.35	1.39

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	901	XIH	C17-C18-C23	7.41	131.40	120.76
2	A	901	XIH	C18-C23-C24	6.41	126.13	120.16
2	A	901	XIH	C17-C18-C19	-4.22	111.73	121.04
2	A	901	XIH	C22-C23-C24	-3.36	113.54	119.36

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	901	XIH	C18-C17-N16	2.89	118.00	113.23

There are no chirality outliers.

All (2) torsion outliers are listed below:

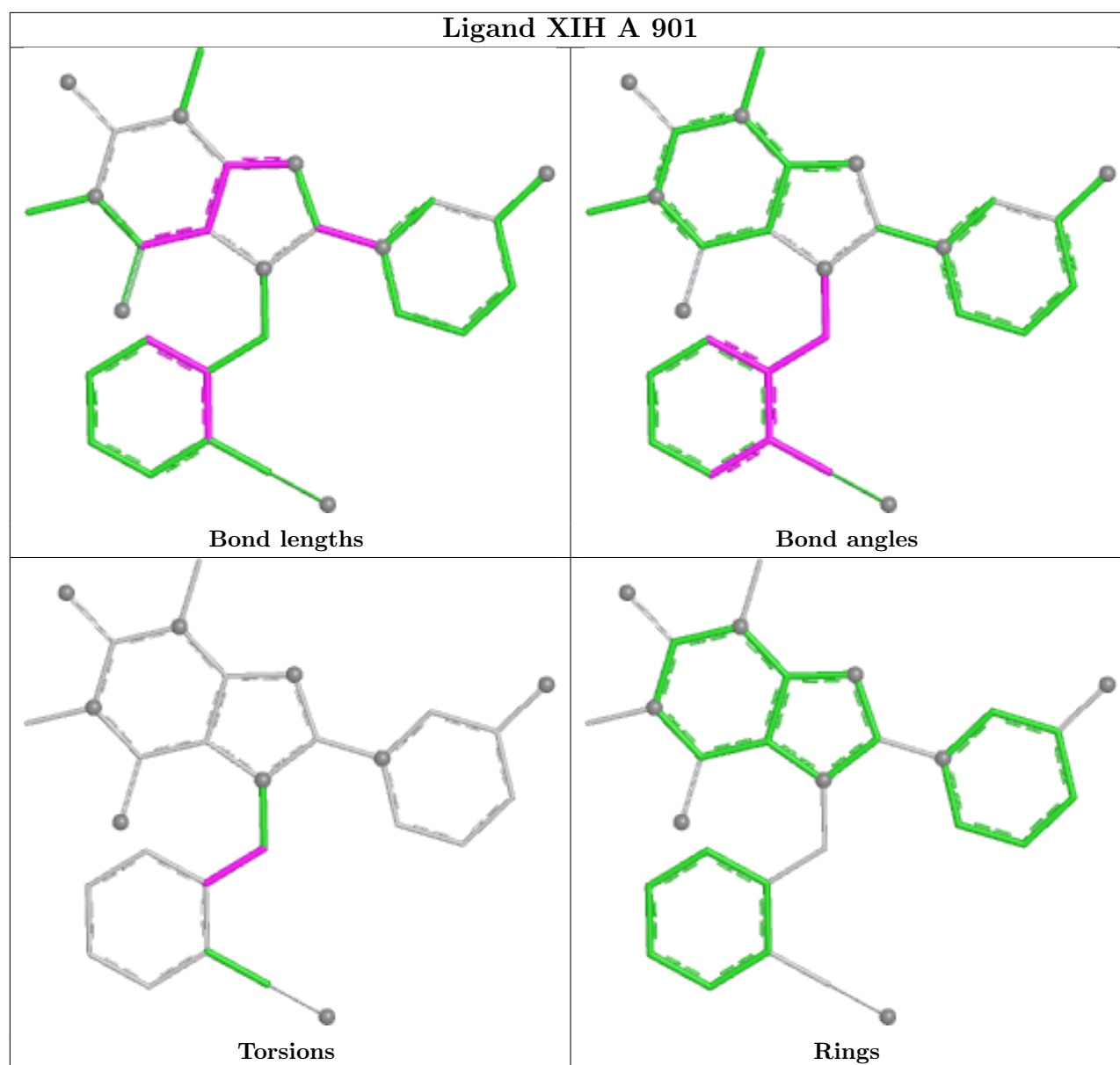
Mol	Chain	Res	Type	Atoms
2	A	901	XIH	N16-C17-C18-C19
2	A	901	XIH	N16-C17-C18-C23

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	901	XIH	3	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 than 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.



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	730/730 (100%)	-0.38	1 (0%) 95 97	4, 31, 79, 187	0
1	B	730/730 (100%)	-0.34	2 (0%) 94 94	5, 37, 92, 161	0
All	All	1460/1460 (100%)	-0.36	3 (0%) 95 96	4, 33, 88, 187	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	767	ARG	3.6
1	B	392	GLU	3.3
1	B	393	GLN	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](#)

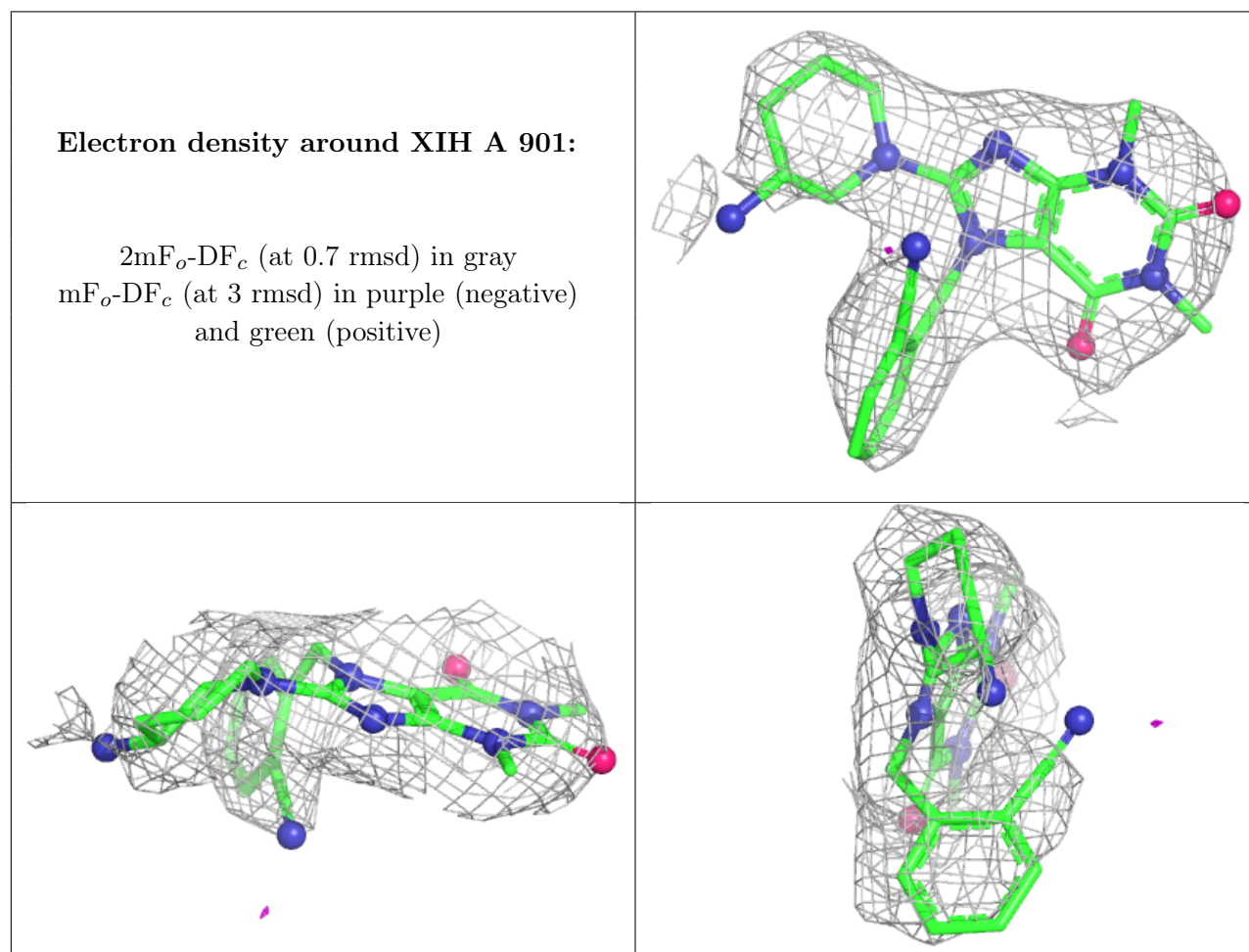
There are no monosaccharides in this entry.

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, 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	B-factors(Å ²)	Q<0.9
2	XIH	A	901	29/29	0.92	0.25	48,48,48,48	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.