



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2026 – 03:10 AM UTC

PDB ID : 1S6P / pdb_00001s6p
Title : CRYSTAL STRUCTURE OF HUMAN IMMUNODEFICIENCY VIRUS
TYPE 1 REVERSE TRANSCRIPTASE (RT) IN COMPLEX WITH
JANSSEN-R100943
Authors : Das, K.; Arnold, E.
Deposited on : 2004-01-26
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

<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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

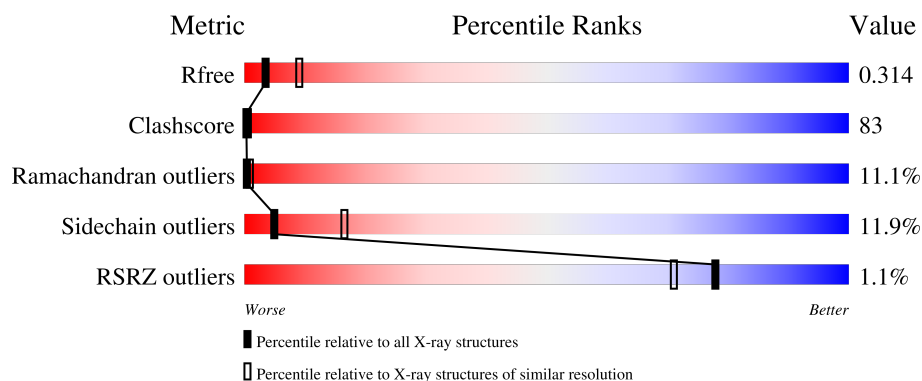
1 Overall quality at a glance

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))
R_{free}	180053	2481 (2.90-2.90)
Clashscore	190562	2690 (2.90-2.90)
Ramachandran outliers	187476	2623 (2.90-2.90)
Sidechain outliers	187428	2625 (2.90-2.90)
RSRZ outliers	180081	2481 (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 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	560	19% 58% 20% ..
2	B	430	14% 54% 28% ..

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 8202 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 POL polyprotein [Contains: Reverse transcriptase].

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	552	Total	C	N	O	S	18	0	0
			4498	2913	748	830	7			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	280	SER	CYS	engineered mutation	UNP P03366

- Molecule 2 is a protein called POL polyprotein [Contains: Reverse transcriptase].

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	427	Total	C	N	O	S	17	0	0
			3529	2300	584	638	7			

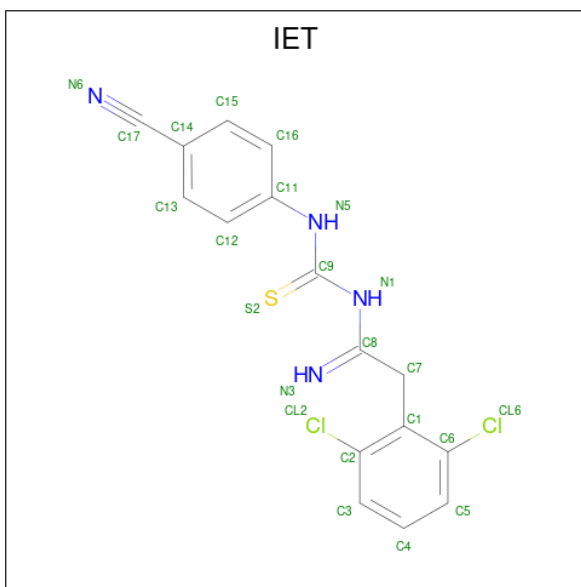
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	280	SER	CYS	engineered mutation	UNP P03366

- Molecule 3 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

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

- Molecule 4 is 1-(4-CYANO-PHENYL)-3-[2-(2,6-DICHLORO-PHENYL)-1-IMINO-ETHYL]-THIOUREA (CCD ID: IET) (formula: C₁₆H₁₂Cl₂N₄S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	Cl	N	S	0	0
			23	16	2	4	1		

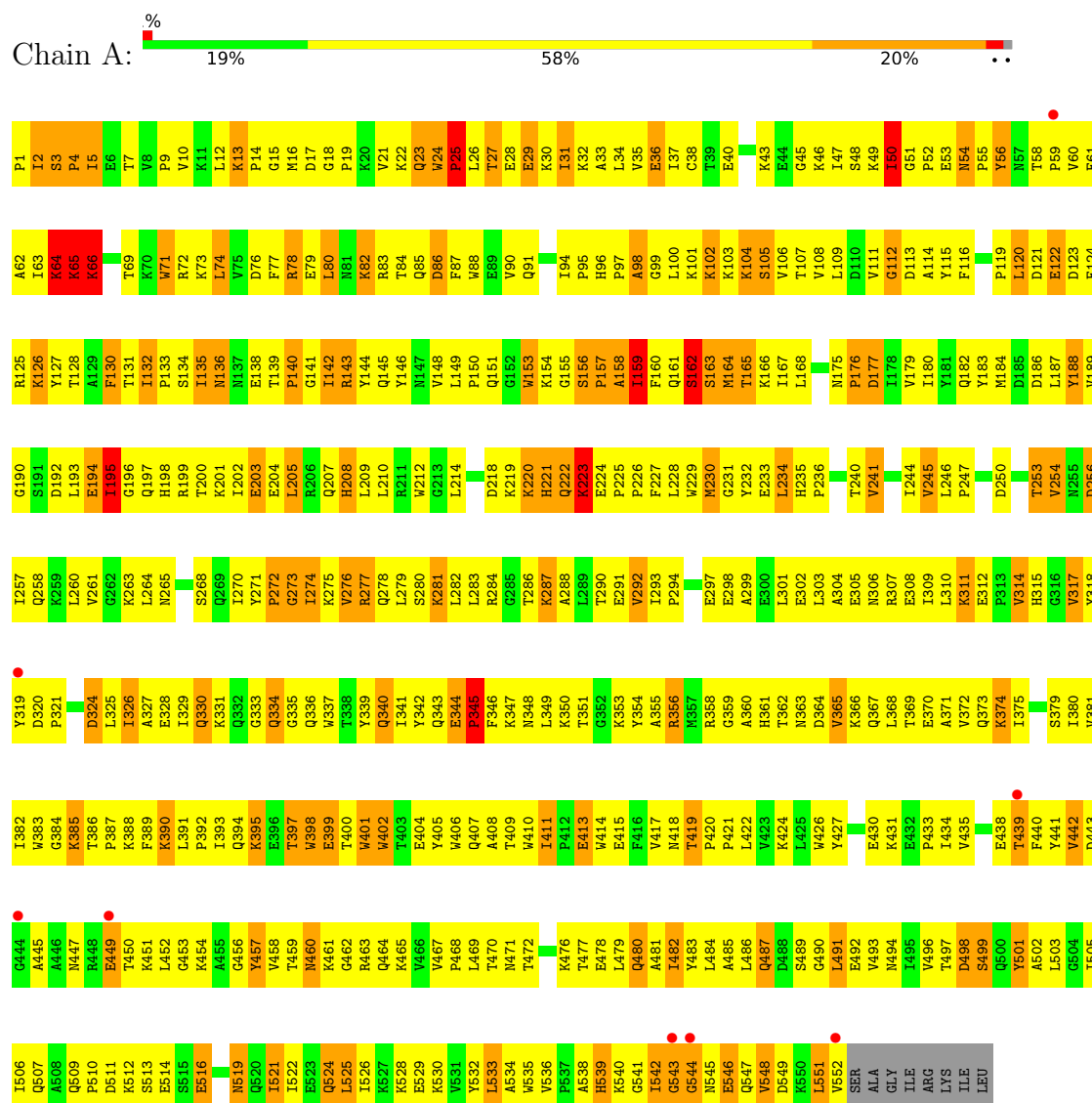
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	81	Total	O	0	0
			81	81		
5	B	70	Total	O	0	0
			70	70		

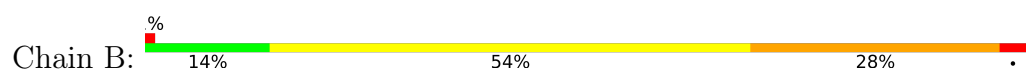
3 Residue-property plots

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: POL polyprotein [Contains: Reverse transcriptase]



• Molecule 2: POL polyprotein [Contains: Reverse transcriptase]



P1	K65	G190	D250	P313	I375
I2	K66	S191	S251	V314	I376
I5	D67	D192	W252	H315	I377
E6	S68	L193	T253	G316	E378
T7	T69	E194	V254	V317	S379
V8	K70	I195	W255	V318	I380
P9	W71	G196	D256	V319	V381
V10	R72	Q197	I257	D320	I382
K11	K73	H198	Q258	P321	W383
L12	L74	R199	K259	S322	G384
K13	V75	T200	L260	K323	K385
P14	D76	K201	V261	D324	T386
G15	F77	I202	G262	L325	P387
M16	R78	E203	K263	I326	K388
D17	E79	E204	L264	A327	F389
G18	L80	L205	N265	E328	K390
	M81	R206	W266	I329	L391
	N82	Q207	A267	Q330	P392
K22	R83	H208	S268	K331	I393
Q23	T84	L209	Q269	Q332	Q394
W24	Q85	L210	I270	G333	K395
	D86	W212	Y271	Q334	
T27	F87	G213	G273	Q336	W398
E28	W88	L214	I274	Q337	E399
E29	E89	T215	K275	T338	T400
K30	V90	T216	V276	Y339	W401
I31	Q91	P217	K277	Q340	W402
K32	L92	D218	Q278	I341	T403
A33	P157	K219	L279	Y342	E404
L34	I159	K220	S280	Q343	Y405
V35	F160	H221	K281	E344	W406
E36	L100	Q222	L282	P345	
I37	C38	K223	L283	F346	T409
T39	K102	E224	R284		W410
E40	K103	P225	Q285	L349	I411
M41	K104	P226		K350	P412
E42	T107	F227	L289	T351	
K43	V108	L228	T290		F416
E44	G45	W229	E291	Y354	V417
K45	L109	M230	V292	A355	M418
K46	D110	G231	I293	K356	T419
I47	V111	E232	P294	M357	P420
S48	G112	E233	L295	K358	P421
K49	D113	L234	T296	G359	L422
I50	A114	H235	E297	K360	V423
G51	Y115	P236	E298	H361	K424
P52	F116	D237	A299	T362	L425
E53	S117	K238	E300	N363	W426
N54	W118	W239	L301	D364	Y427
P55	P119	T240	E302	V365	GLN
Y56		Y241		K366	LEU
N57	E122	Q242	E305	Q367	GLU
T58	D123	P243	N306	L368	
P59	F124	I244	R307	T369	
V60	R125	E245	E308	E370	
F61	K126	L246	I309	A371	
A62	Y127	P247	L310	V372	
I63	T128	E248	K311	Q373	
K64	A129	K249	E312	K374	

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	226.70Å 67.40Å 104.30Å 90.00° 107.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.90 20.00 – 2.90	Depositor EDS
% Data completeness (in resolution range)	89.8 (20.00-2.90) 91.8 (20.00-2.90)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.59 (at 2.83Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.245 , 0.312 0.252 , 0.314	Depositor DCC
R_{free} test set	1616 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	67.3	Xtriage
Anisotropy	0.215	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 100.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	8202	wwPDB-VP
Average B, all atoms (Å ²)	76.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.75% 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 [i](#)

5.1 Standard geometry [i](#)

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

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	0.95	7/4616 (0.2%)	1.27	52/6271 (0.8%)
2	B	0.81	7/3634 (0.2%)	1.41	60/4940 (1.2%)
All	All	0.89	14/8250 (0.2%)	1.33	112/11211 (1.0%)

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	1	0
2	B	0	1
All	All	1	1

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	31	ILE	CG1-CD1	40.87	3.11	1.51
1	A	221	HIS	CA-C	9.78	1.57	1.52
2	B	224	GLU	C-N	-7.84	1.24	1.33
1	A	498	ASP	C-N	-6.97	1.24	1.33
2	B	225	PRO	CA-C	6.73	1.58	1.52

The worst 5 of 112 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	31	ILE	CB-CG1-CD1	-27.80	55.42	113.80
2	B	419	THR	CA-C-N	-13.21	106.77	120.38
2	B	419	THR	C-N-CA	-13.21	106.77	120.38
2	B	225	PRO	CB-CA-C	-11.61	96.76	110.92

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	156	SER	N-CA-C	-10.50	98.28	112.35

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	31	ILE	CB

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	130	PHE	Sidechain

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	4498	0	4560	724	0
2	B	3529	0	3568	649	0
3	A	1	0	0	0	0
4	A	23	0	12	5	0
5	A	81	0	0	8	0
5	B	70	0	0	5	0
All	All	8202	0	8140	1338	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 83.

The worst 5 of 1338 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:64:LYS:O	1:A:66:LYS:HG2	1.34	1.25
2:B:139:THR:CG2	2:B:140:PRO:HD2	1.66	1.24
1:A:222:GLN:O	1:A:224:GLU:HG3	1.41	1.16
1:A:497:THR:HG22	1:A:499:SER:H	1.09	1.16
2:B:358:ARG:HG2	2:B:359:GLY:H	1.09	1.16

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	550/560 (98%)	404 (74%)	92 (17%)	54 (10%)	0	1
2	B	425/430 (99%)	277 (65%)	94 (22%)	54 (13%)	0	0
All	All	975/990 (98%)	681 (70%)	186 (19%)	108 (11%)	0	1

5 of 108 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	5	ILE
1	A	25	PRO
1	A	52	PRO
1	A	65	LYS
1	A	112	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	494/500 (99%)	443 (90%)	51 (10%)	7	23
2	B	389/392 (99%)	335 (86%)	54 (14%)	3	11
All	All	883/892 (99%)	778 (88%)	105 (12%)	5	16

5 of 105 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	67	ASP
2	B	183	TYR
2	B	383	TRP
2	B	74	LEU
2	B	142	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	255	ASN
2	B	348	ASN
2	B	258	GLN
2	B	330	GLN
2	B	407	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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	IET	A	701	-	24,24,24	5.14	14 (58%)	29,32,32	2.41	7 (24%)

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	IET	A	701	-	-	6/12/14/14	0/2/2/2

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	701	IET	C7-C8	-17.97	1.39	1.51
4	A	701	IET	C6-CL6	-9.50	1.51	1.73
4	A	701	IET	C2-CL2	-9.46	1.51	1.73
4	A	701	IET	C2-C1	4.32	1.46	1.39
4	A	701	IET	C16-C11	4.00	1.46	1.39

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	701	IET	C8-N1-C9	-7.23	121.42	128.72
4	A	701	IET	C1-C7-C8	5.97	124.82	113.64
4	A	701	IET	N5-C9-N1	5.24	123.15	115.31
4	A	701	IET	C7-C1-C6	-3.97	116.63	121.97
4	A	701	IET	C7-C1-C2	3.28	126.39	121.97

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

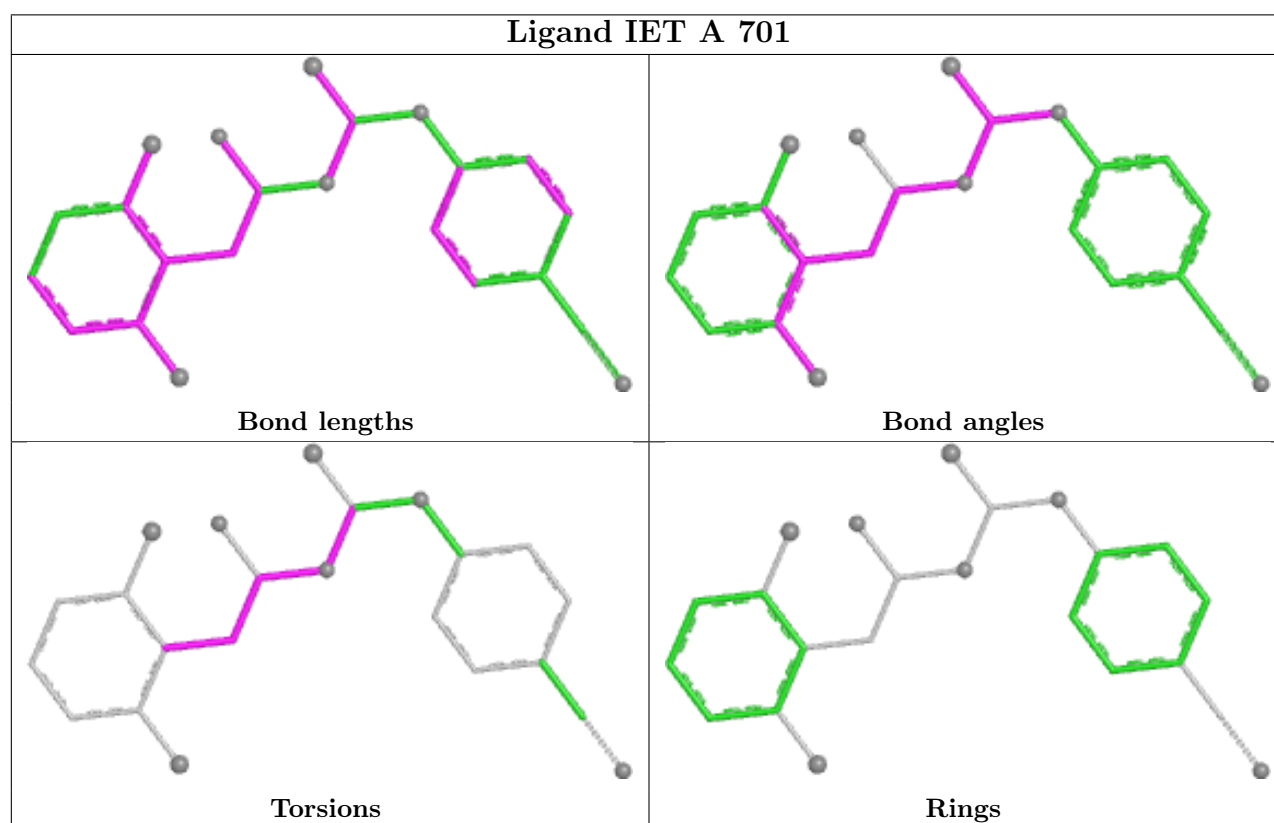
Mol	Chain	Res	Type	Atoms
4	A	701	IET	C7-C8-N1-C9
4	A	701	IET	N5-C9-N1-C8
4	A	701	IET	S2-C9-N1-C8
4	A	701	IET	C2-C1-C7-C8
4	A	701	IET	C1-C7-C8-N1

There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	701	IET	5	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	552/560 (98%)	-0.07	8 (1%) 73 65	37, 82, 109, 111	5 (0%)
2	B	427/430 (99%)	-0.16	3 (0%) 84 79	19, 69, 110, 111	4 (0%)
All	All	979/990 (98%)	-0.11	11 (1%) 78 71	19, 77, 110, 111	9 (0%)

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	444	GLY	4.8
1	A	449	GLU	4.3
2	B	231	GLY	4.1
1	A	59	PRO	3.6
1	A	319	TYR	2.6

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](#)

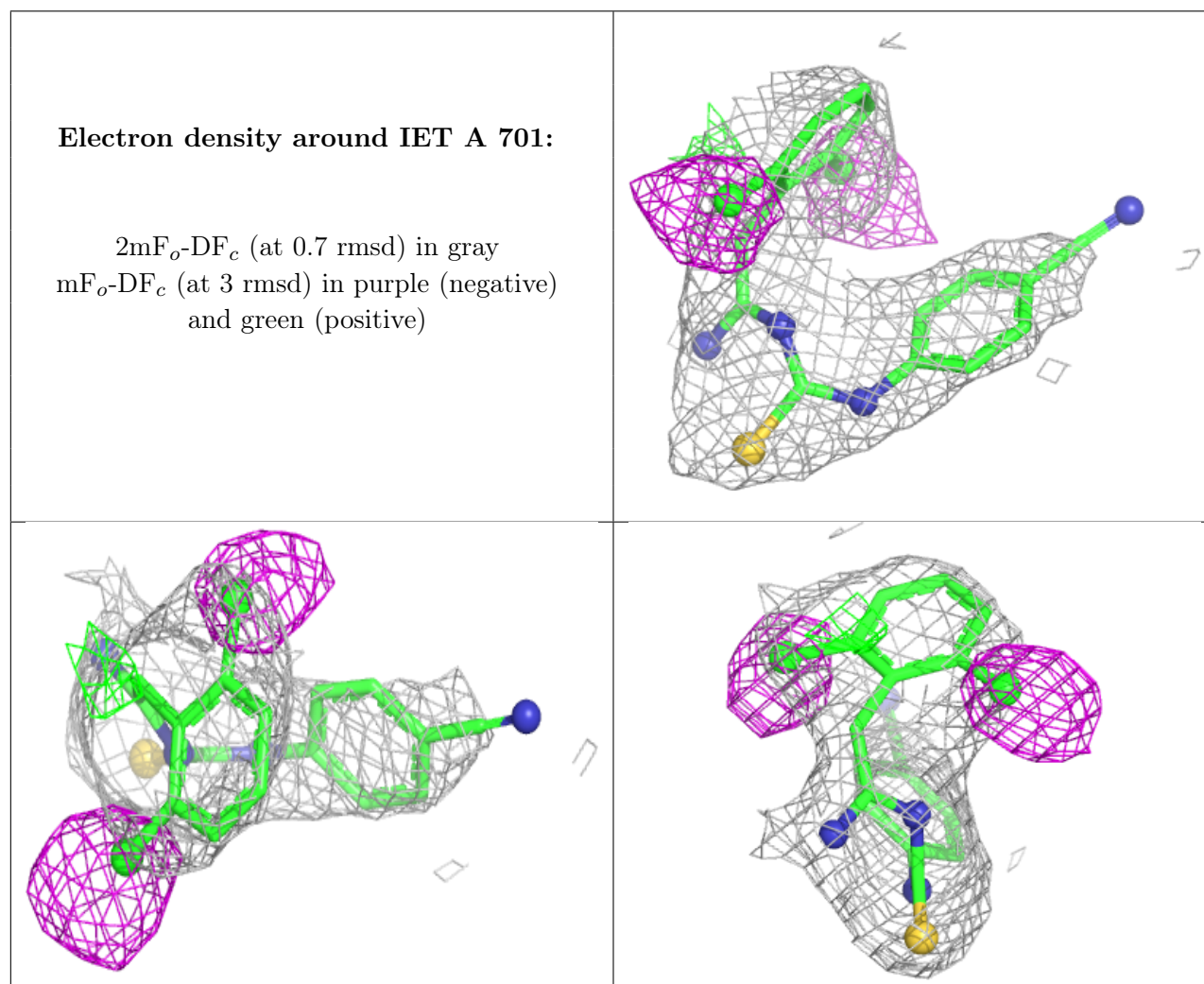
There are no oligosaccharides in this entry.

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($\text{\AA}^2$)	Q<0.9
4	IET	A	701	23/23	0.80	0.13	63,78,84,92	0
3	MG	A	601	1/1	0.85	0.08	79,79,79,79	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.