



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 20, 2024 – 05:02 PM EDT

PDB ID : 1FK9  
Title : CRYSTAL STRUCTURE OF HIV-1 REVERSE TRANSCRIPTASE IN  
COMPLEX WITH DMP-266(EFAVIRENZ)  
Authors : Ren, J.; Milton, J.; Weaver, K.L.; Short, S.A.; Stuart, D.I.; Stammers, D.K.  
Deposited on : 2000-08-09  
Resolution : 2.50 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](mailto: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 : 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

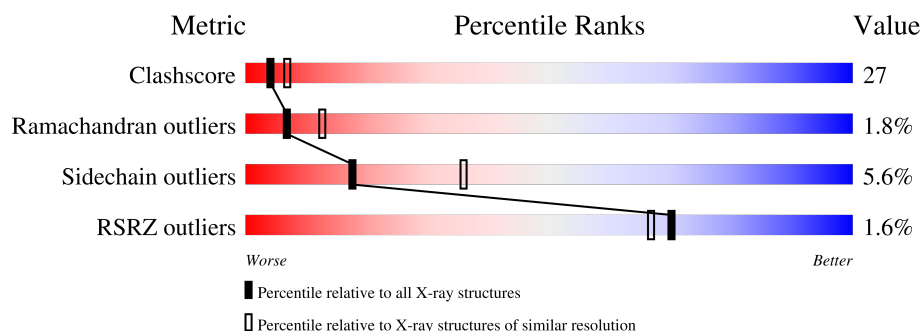
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	180529	6282 (2.50-2.50)
Ramachandran outliers	177936	6191 (2.50-2.50)
Sidechain outliers	177891	6193 (2.50-2.50)
RSRZ outliers	164620	5504 (2.50-2.50)

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  $\geq 3$ , 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	543	<div> <div> <div></div> <div>49%</div> <div>42%</div> <div>5%</div> <div>.</div> </div> </div>
2	B	440	<div> <div>2%</div> <div> <div></div> <div>48%</div> <div>41%</div> <div>.</div> <div>8%</div> </div> </div>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 7740 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.

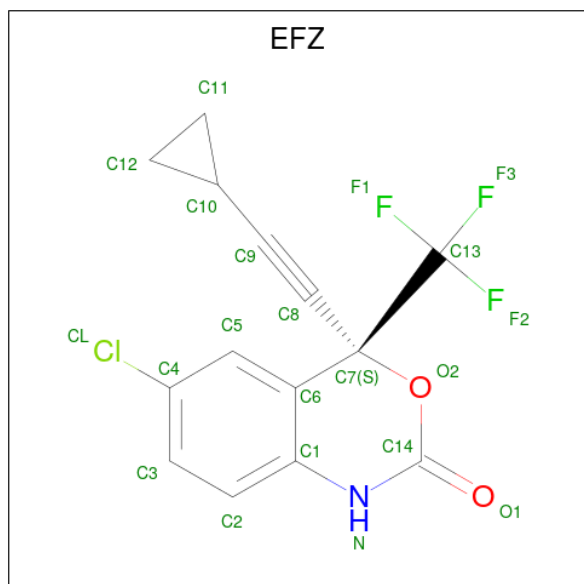
- Molecule 1 is a protein called HIV-1 RT, A-CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	520	Total	C	N	O	S	0	0	0
			4269	2767	706	788	8			

- Molecule 2 is a protein called HIV-1 RT, B-CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	403	Total	C	N	O	S	0	0	0
			3337	2172	552	607	6			

- Molecule 3 is (-)-6-CHLORO-4-CYCLOPROPYLETHYNYL-4-TRIFLUOROMETHYL-1,4-DIHYDRO-2H-3,1-BENZOXAZIN-2-ONE (three-letter code: EFZ) (formula: C<sub>14</sub>H<sub>9</sub>ClF<sub>3</sub>NO<sub>2</sub>).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	A	1	Total	C	Cl	F	N	O	0	0
			21	14	1	3	1	2		

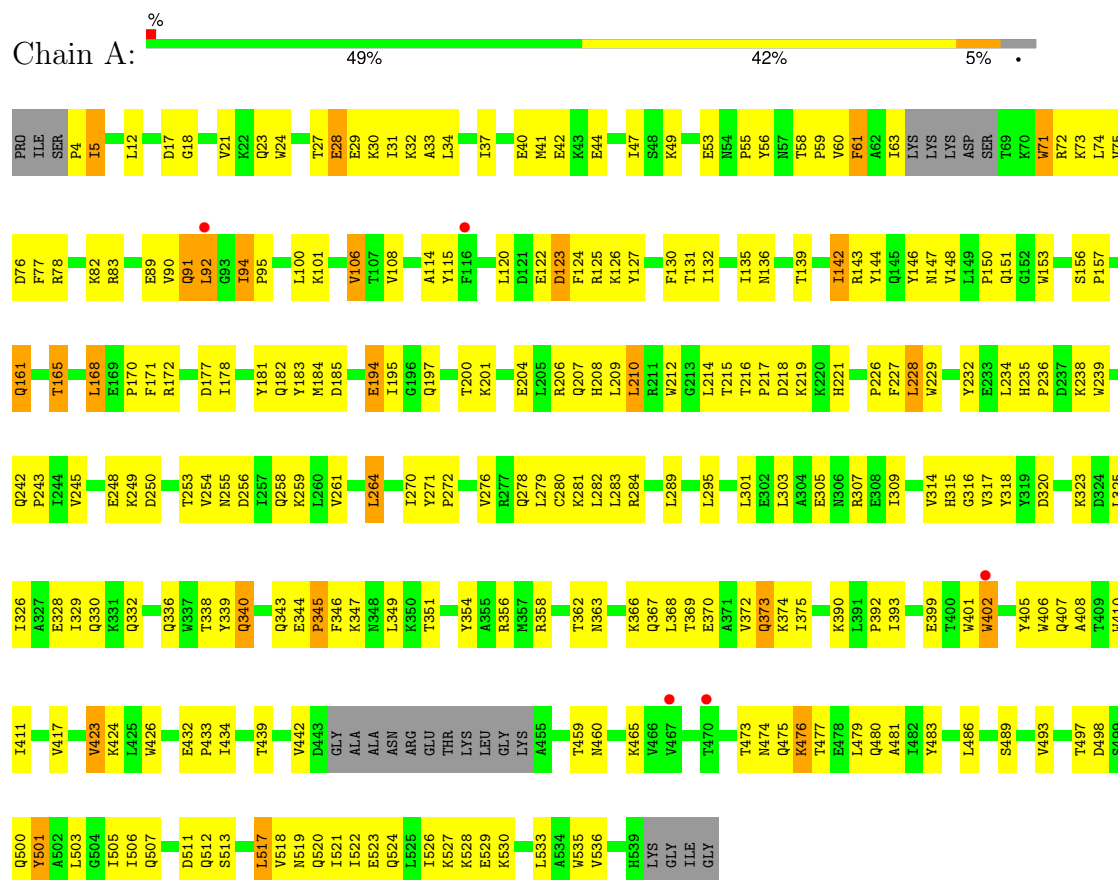
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	63	Total 63	O 63	0	0
4	B	50	Total 50	O 50	0	0

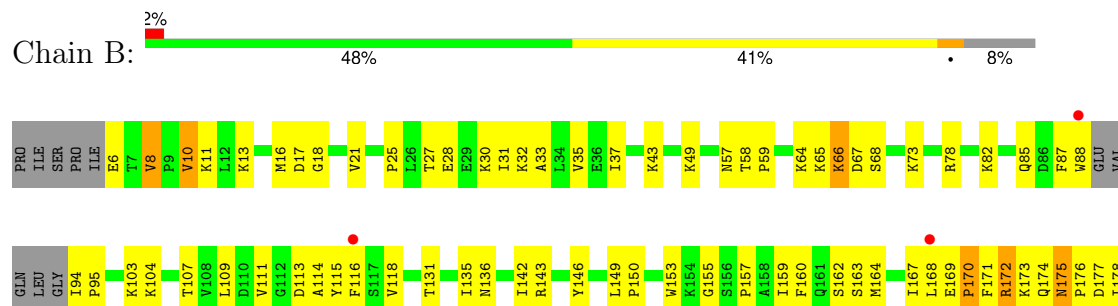
### 3 Residue-property plots

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: HIV-1 RT, A-CHAIN



#### • Molecule 2: HIV-1 RT, B-CHAIN



V179	E248	K337	W418
Y183	K249	T338	T419
M184	W252	Y339	P420
L187	T253	I341	P421
	V254	Y342	L422
S191	I257	Q343	W423
D192	Q258	E344	K424
L193	Q259	F345	L425
E194	L260	F346	W426
I195		K347	Y427
G196	W266	N348	Q428
Q197		I349	LEU
H198	Q269	K350	GLU
R199	I270	T351	LYS
T200	I274	G352	GLU
K201	K275		PRO
I202	V276	R356	ILE
E203	R277	MET	VAL
E204		ARG	GLY
L205		GLY	ALA
R206	C280	ALA	GLU
Q207	K281	HIS	THR
H208	L282	T362	PHE
L209	L283	N363	
L210	R284	D364	
R211		V365	
W212	K287	K366	
G213	A288	Q367	
L214	L289	L368	
T215		T369	
T216	I293	E370	
P217	P294	A371	
D218	L295	V372	
K219	T296	Q373	
K220	E297		
HIS		I380	
GLN	E300	G384	
LYS	L301	K385	
GLU	E302	T386	
PRO	L303	P387	
PRO	A304	K388	
PHE	E305	F389	
LEU	N306		
TRP	R307	Q394	
MET	E308	T397	
G231	I309		
	L310	W401	
H235		W402	
P236	D320	T403	
D237		E404	
K238	K323	Y405	
W239	D324	W406	
T240	L325		
V241	T326	W410	
Q242	A327		
P243		W414	
	Q332		
P247			

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	139.80Å 115.00Å 65.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.50 30.00 – 2.50	Depositor EDS
% Data completeness (in resolution range)	85.2 (30.00-2.50) 85.1 (30.00-2.50)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.68 (at 2.51Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.218 , 0.301 0.213 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	57.8	Xtriage
Anisotropy	0.179	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 69.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7740	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.68% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: EFZ, CSD

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	0.42	0/4375	0.64	0/5950
2	B	0.39	0/3430	0.62	0/4658
All	All	0.40	0/7805	0.63	0/10608

There are no bond length outliers.

There are no bond angle outliers.

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	4269	0	4293	245	0
2	B	3337	0	3368	180	0
3	A	21	0	9	1	0
4	A	63	0	0	7	0
4	B	50	0	0	4	0
All	All	7740	0	7670	415	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 27.

The worst 5 of 415 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)
2:B:111:VAL:HA	2:B:214:LEU:HD11	1.37	1.01
2:B:94:ILE:HD12	2:B:95:PRO:HD2	1.43	0.99
1:A:181:TYR:HE1	1:A:183:TYR:HB2	1.34	0.92
2:B:173:LYS:HA	2:B:176:PRO:HG3	1.50	0.92
1:A:5:ILE:HD12	1:A:5:ILE:H	1.35	0.89

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	513/543 (94%)	461 (90%)	43 (8%)	9 (2%)	7	12
2	B	395/440 (90%)	351 (89%)	37 (9%)	7 (2%)	7	12
All	All	908/983 (92%)	812 (89%)	80 (9%)	16 (2%)	7	12

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	91	GLN
1	A	92	LEU
1	A	219	LYS
1	A	402	TRP
2	B	162	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	468/485 (96%)	439 (94%)	29 (6%)	15	31
2	B	368/400 (92%)	350 (95%)	18 (5%)	21	42
All	All	836/885 (94%)	789 (94%)	47 (6%)	17	36

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	533	LEU
2	B	207	GLN
2	B	6	GLU
2	B	66	LYS
2	B	219	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	520	GLN
2	B	151	GLN
2	B	407	GLN
2	B	147	ASN
2	B	175	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue 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
1	CSD	A	280	1	4,7,8	2.74	1 (25%)	1,8,10	5.13	1 (100%)

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
1	CSD	A	280	1	-	2/2/6/8	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	280	CSD	OD1-SG	5.30	1.52	1.47

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	280	CSD	OD1-SG-CB	5.13	115.05	105.60

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	280	CSD	N-CA-CB-SG
1	A	280	CSD	CA-CB-SG-OD1

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides 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	$\# Z  > 2$
3	EFZ	A	999	-	23,23,23	2.67	6 (26%)	36,36,36	1.56	2 (5%)

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
3	EFZ	A	999	-	-	1/10/32/32	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	999	EFZ	C10-C9	7.98	1.70	1.45
3	A	999	EFZ	C12-C11	-5.68	1.28	1.48
3	A	999	EFZ	C11-C10	-4.30	1.23	1.48
3	A	999	EFZ	C12-C10	-4.25	1.24	1.48
3	A	999	EFZ	C7-C6	2.90	1.55	1.51

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	999	EFZ	C11-C10-C9	-5.75	105.52	119.09
3	A	999	EFZ	C12-C10-C9	-5.73	105.57	119.09

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	999	EFZ	C7-C8-C9-C10

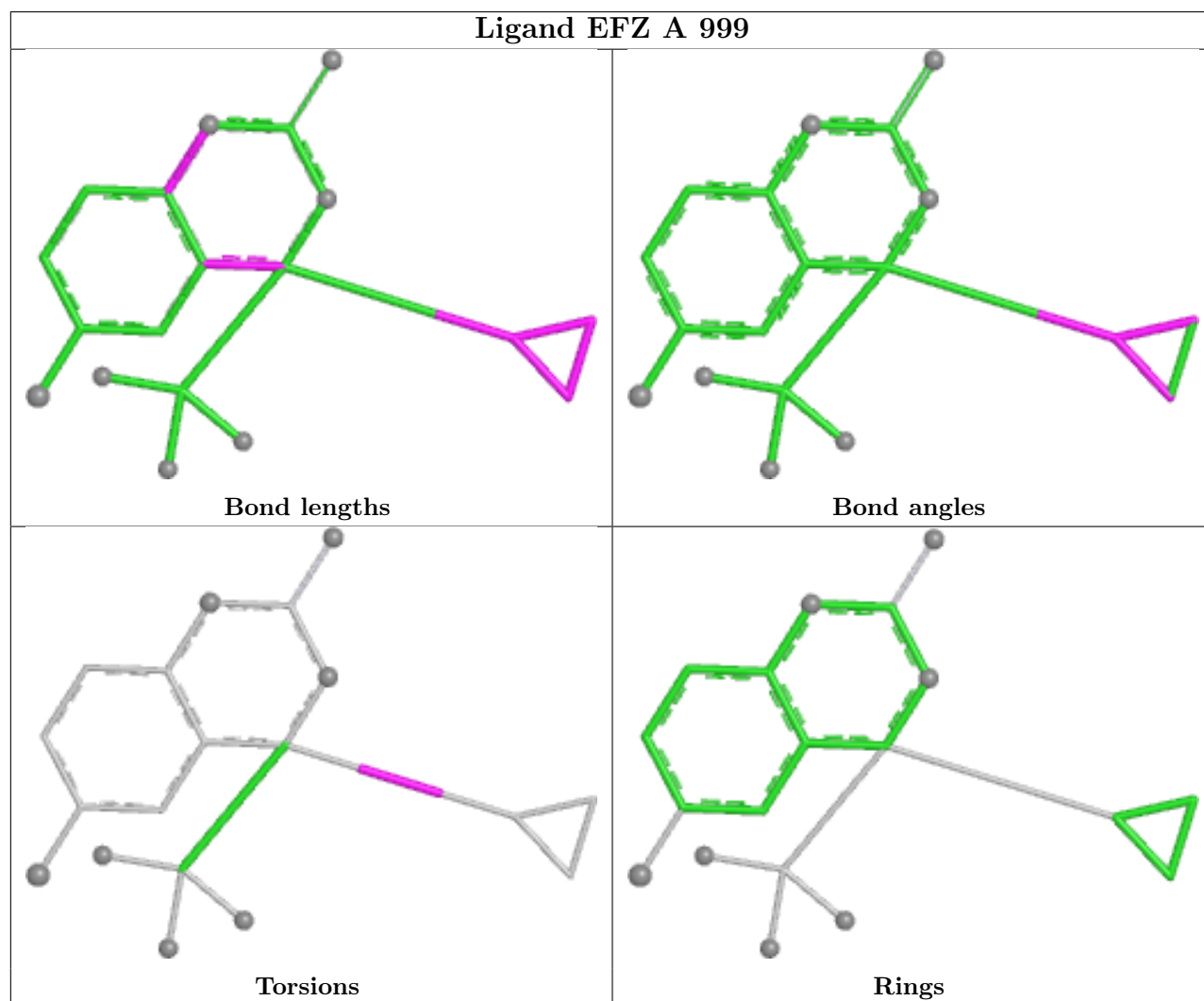
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	999	EFZ	1	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 ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	519/543 (95%)	-0.11	5 (0%) 79 76	28, 63, 103, 139	0
2	B	403/440 (91%)	-0.08	10 (2%) 58 55	27, 62, 111, 148	0
All	All	922/983 (93%)	-0.10	15 (1%) 70 67	27, 63, 109, 148	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	214	LEU	6.5
2	B	116	PHE	3.6
2	B	428	GLN	2.8
2	B	215	THR	2.8
1	A	467	VAL	2.5

### 6.2 Non-standard residues in protein, DNA, RNA chains [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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	CSD	A	280	8/9	0.95	0.07	48,50,55,57	0

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

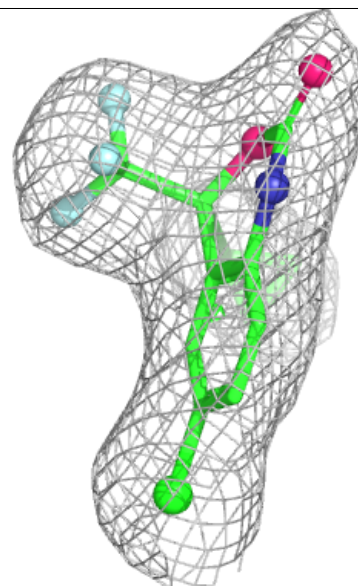
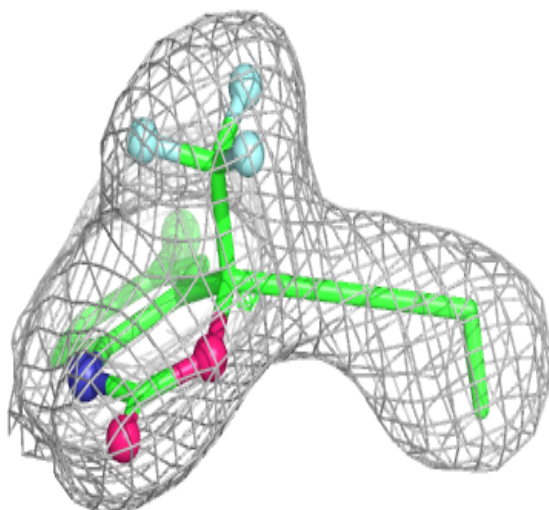
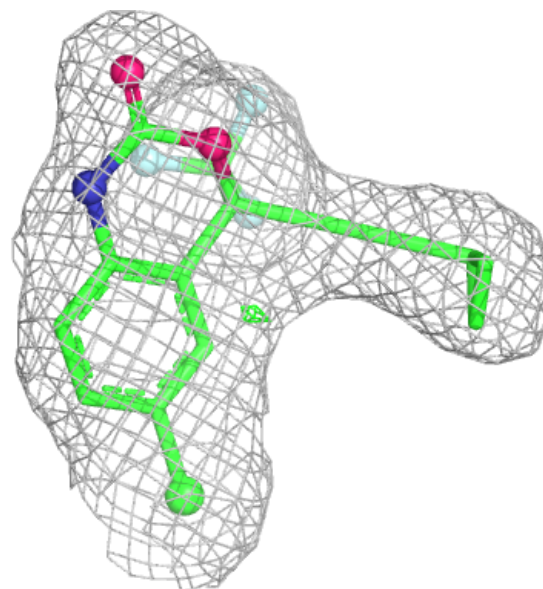
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	EFZ	A	999	21/21	0.96	0.06	35,42,49,54	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.

### Electron density around EFZ A 999:

2mF<sub>o</sub>-DF<sub>c</sub> (at 0.7 rmsd) in gray  
mF<sub>o</sub>-DF<sub>c</sub> (at 3 rmsd) in purple (negative)  
and green (positive)





## 6.5 Other polymers [i](#)

There are no such residues in this entry.