



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 05:45 am GMT

PDB ID : 1HMY  
Title : THE STRUCTURE OF UNLIGANDED REVERSE TRANSCRIPTASE  
FROM THE HUMAN IMMUNODEFICIENCY VIRUS TYPE 1  
Authors : Rodgers, D.W.; Gamblin, S.J.; Harris, B.A.; Ray, S.; Culp, J.S.; Hellmig, B.;  
Woolf, D.J.; Debouck, C.; Harrison, S.C.  
Deposited on : 1994-12-15  
Resolution : 3.20 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

<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  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

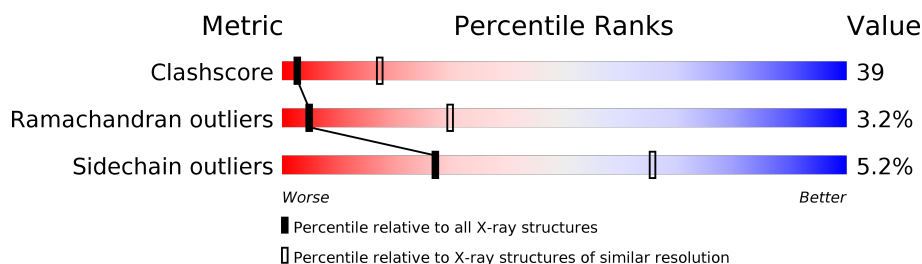
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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	1009 (3.20-3.20)
Ramachandran outliers	110173	1118 (3.22-3.18)
Sidechain outliers	110143	1117 (3.22-3.18)


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  $\geq 3$ , 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	560	
1	C	560	
1	E	560	
1	G	560	
2	B	440	
2	D	440	
2	F	440	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	H	440	 A horizontal bar chart showing the quality of chain H. The bar is divided into four segments: green (39%), yellow (45%), orange (10%), and red (6%). The segments are labeled with their respective percentages: 39%, 45%, and 10% (for the red segment). The red segment is the smallest, followed by orange, then yellow, and green is the largest.

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 29596 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 REVERSE TRANSCRIPTASE (SUBUNIT P66).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	536	Total	C	N	O	S	0	0	0
			4200	2711	698	784	7			
1	C	536	Total	C	N	O	S	0	0	0
			4200	2711	698	784	7			
1	E	536	Total	C	N	O	S	0	0	0
			4200	2711	698	784	7			
1	G	536	Total	C	N	O	S	0	0	0
			4200	2711	698	784	7			

- Molecule 2 is a protein called HIV-1 REVERSE TRANSCRIPTASE (SUBUNIT P51).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	395	Total	C	N	O	S	0	0	0
			3198	2079	531	582	6			
2	D	395	Total	C	N	O	S	0	0	0
			3198	2079	531	582	6			
2	F	395	Total	C	N	O	S	0	0	0
			3198	2079	531	582	6			
2	H	395	Total	C	N	O	S	0	0	0
			3198	2079	531	582	6			

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

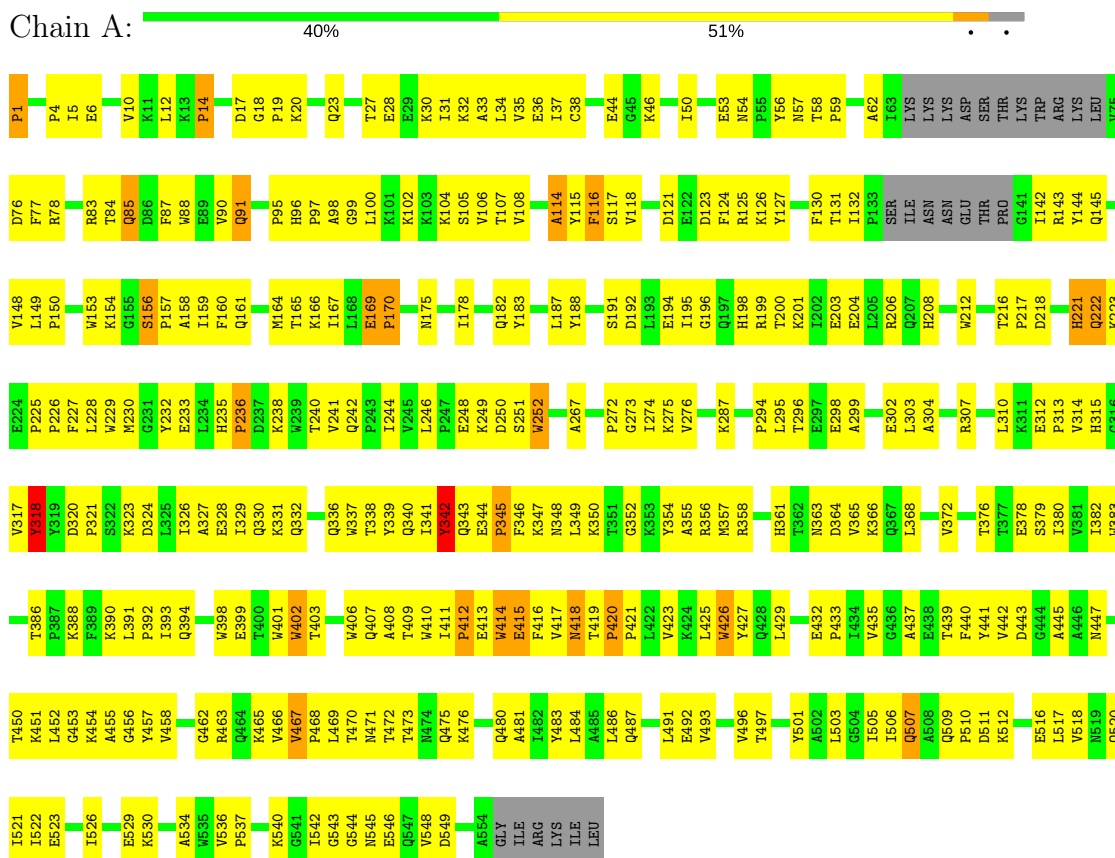
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	H	1	Total	Mg	0	0
			1	1		
3	B	1	Total	Mg	0	0
			1	1		
3	D	1	Total	Mg	0	0
			1	1		
3	F	1	Total	Mg	0	0
			1	1		

### 3 Residue-property plots

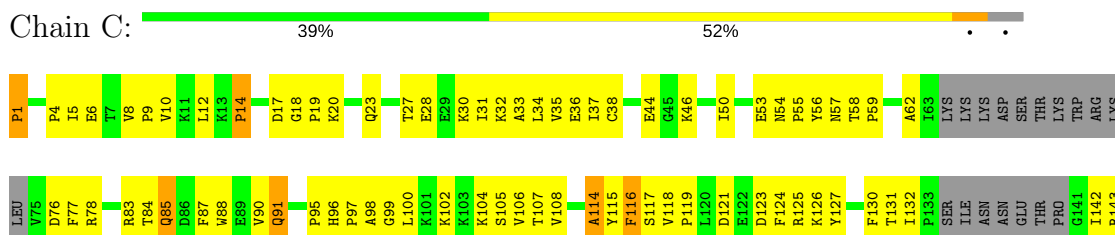
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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.

Note EDS was not executed.

#### • Molecule 1: HIV-1 REVERSE TRANSCRIPTASE (SUBUNIT P66)



#### • Molecule 1: HIV-1 REVERSE TRANSCRIPTASE (SUBUNIT P66)

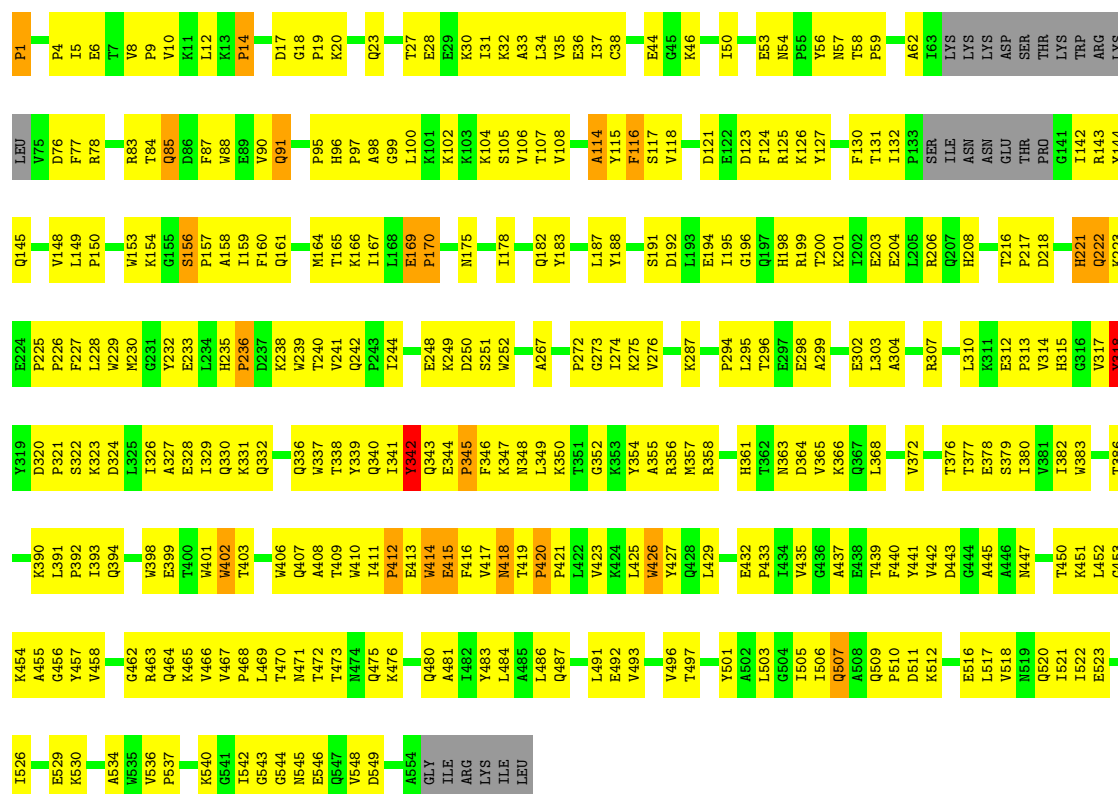




Response	Percentage
Yes	40%
No	51%
Don't know	•
Refuse to answer	•

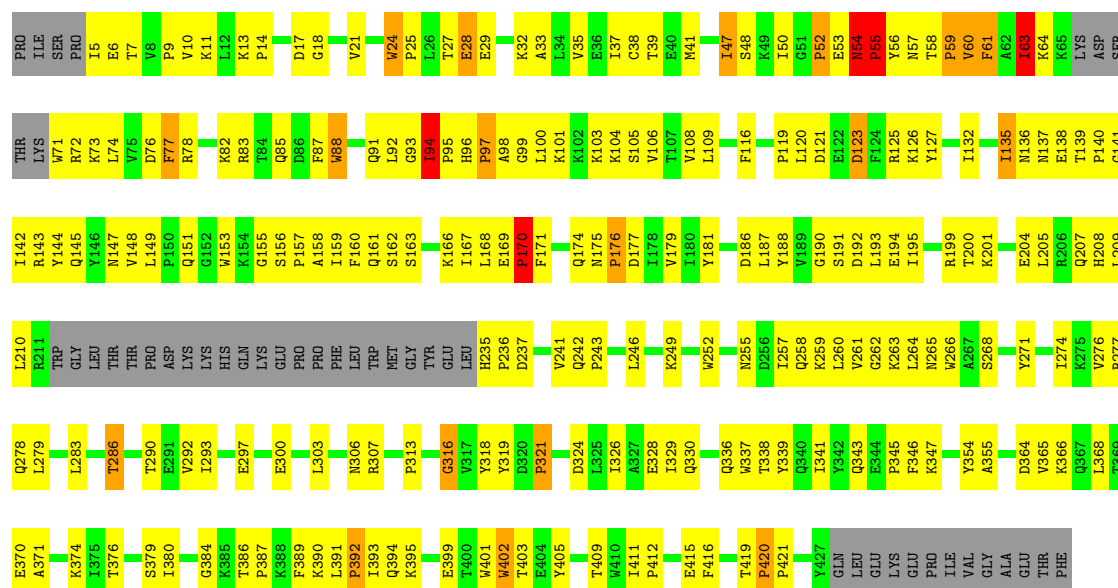


Response	Percentage
Yes	39%
No	52%
Don't know	7%
Refuse to answer	4%



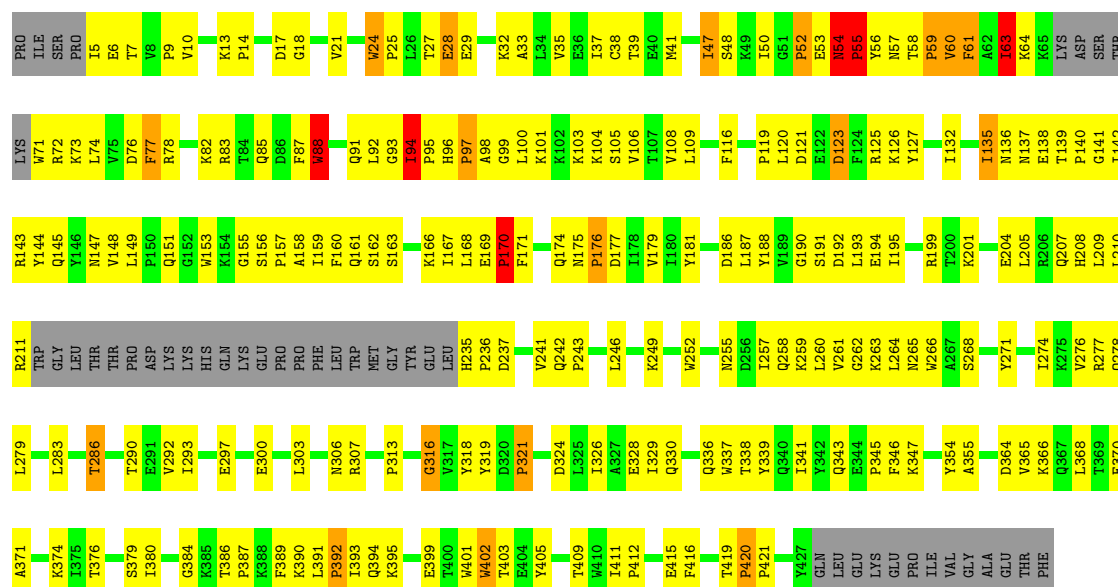
• Molecule 2: HIV-1 REVERSE TRANSCRIPTASE (SUBUNIT P51)

Chain B: 39% 45% 10%



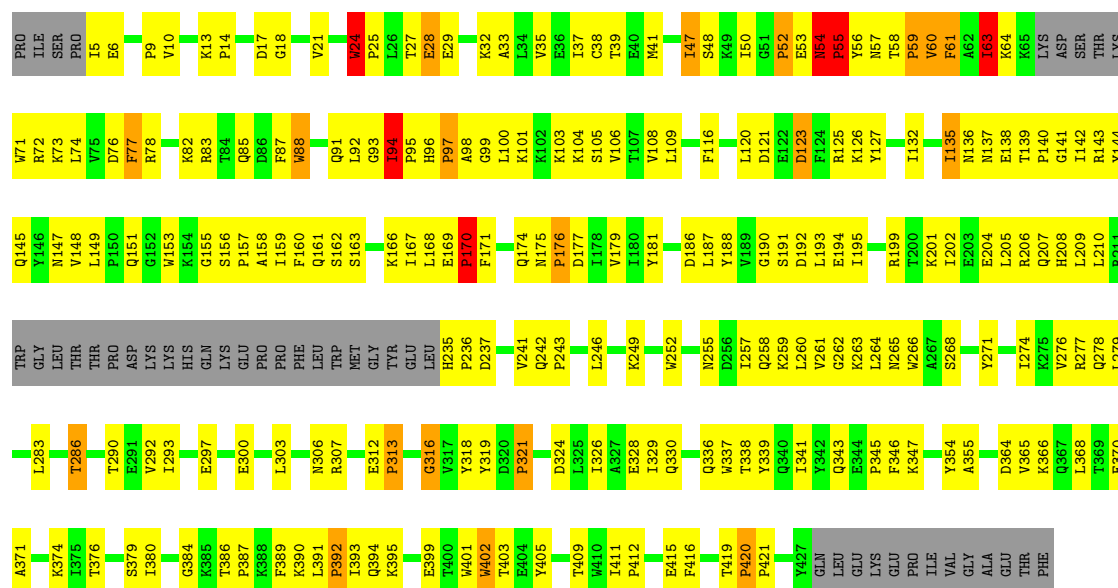
• Molecule 2: HIV-1 REVERSE TRANSCRIPTASE (SUBUNIT P51)

Chain D: 40% 45% 10%



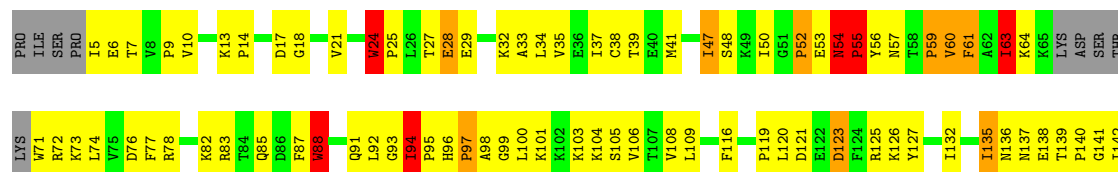
● Molecule 2: HIV-1 REVERSE TRANSCRIPTASE (SUBUNIT P51)

Chain F: 40% 45% 10%



● Molecule 2: HIV-1 REVERSE TRANSCRIPTASE (SUBUNIT P51)

Chain H: 39% 45% 10%





A371	L279	R143	R211
K374	L283	Y144	TRP
I375	L286	GLY	GLY
T376	T290	Y146	LEU
S379	E291	N147	THR
I380	V292	V148	THR
G384	I293	L149	PRO
K385	E297	F150	ASP
T386	E300	G151	LYS
P387	L303	G152	LYS
K388	N306	W153	HIS
F389	R307	GLN	GLN
K390	P313	G154	LYS
L391	G316	G155	LYS
P392	V317	S156	GLU
I393	Y318	P157	PRO
Q394	Y319	A158	PRO
K395	D320	I159	PHE
E399	P321	F160	LEU
T400	D324	Q161	TRP
W401	L325	S162	MET
W402	I326	S163	GLY
W403	A327	K166	TYR
T404	E328	GLU	GLU
Y405	I329	I167	LEU
T409	Q330	L168	H235
W410	Q336	E169	P236
I411	W337	F170	D237
P412	T338	F171	V241
E415	Y339	Q174	Q242
F416	I341	N175	P243
T419	Q340	P176	L246
P420	Y342	I177	K249
P421	Q343	I180	W252
Y427	E344	Y181	N255
GLN	P345	D186	D256
LEU	F346	L187	I257
GLU	K347	Y188	Q258
LYS	Y354	G190	K259
LYS	A355	S191	L260
GLU	D364	D192	V261
PRO	V365	L193	K262
ILE	K366	E194	L264
VAL	Q367	I195	N265
GLY	L368	R199	W266
ALA	T369	L200	A267
GLU	E370	K201	S268
THR		E204	Y271
PHE		L205	I274
		R206	K275
		Q207	V276
		H208	R277
		L209	Q278
		L210	

## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	168.70Å 162.80Å 331.80Å 90.00° 105.70° 90.00°	Depositor
Resolution (Å)	6.00 – 3.20	Depositor
% Data completeness (in resolution range)	(Not available) (6.00-3.20)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, $R_{free}$	0.254 , 0.297	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	29596	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	97.0	wwPDB-VP

## 5 Model quality

### 5.1 Standard geometry

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

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	0.71	1/4307 (0.0%)	0.98	5/5865 (0.1%)
1	C	0.71	1/4307 (0.0%)	0.98	4/5865 (0.1%)
1	E	0.71	1/4307 (0.0%)	0.98	5/5865 (0.1%)
1	G	0.71	1/4307 (0.0%)	0.98	4/5865 (0.1%)
2	B	0.74	0/3285	1.02	5/4466 (0.1%)
2	D	0.74	0/3285	1.02	5/4466 (0.1%)
2	F	0.74	1/3285 (0.0%)	1.02	5/4466 (0.1%)
2	H	0.74	1/3285 (0.0%)	1.02	5/4466 (0.1%)
All	All	0.72	6/30368 (0.0%)	0.99	38/41324 (0.1%)

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	0	2
1	C	0	2
1	E	0	2
1	G	0	2
2	B	0	1
2	D	0	1
2	F	0	1
2	H	0	1
All	All	0	12

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	222	GLN	C-O	5.35	1.33	1.23
1	A	222	GLN	C-O	5.33	1.33	1.23

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	222	GLN	C-O	5.32	1.33	1.23
1	E	222	GLN	C-O	5.32	1.33	1.23
2	F	24	TRP	CB-CG	-5.02	1.41	1.50

The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	94	ILE	C-N-CD	7.79	144.77	128.40
2	F	94	ILE	C-N-CD	7.77	144.72	128.40
2	B	94	ILE	C-N-CD	7.77	144.71	128.40
2	D	94	ILE	C-N-CD	7.76	144.69	128.40
2	D	54	ASN	C-N-CD	-7.01	105.17	120.60

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	318	TYR	Sidechain
1	A	342	TYR	Sidechain
2	B	61	PHE	Sidechain
1	C	318	TYR	Sidechain
1	C	342	TYR	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	4200	0	4064	338	2
1	C	4200	0	4064	405	24
1	E	4200	0	4064	342	2
1	G	4200	0	4064	401	8
2	B	3198	0	3184	242	4
2	D	3198	0	3184	241	8
2	F	3198	0	3184	245	0
2	H	3198	0	3184	242	20
3	B	1	0	0	0	0
3	D	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	F	1	0	0	0	0
3	H	1	0	0	0	0
All	All	29596	0	28992	2306	34

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 39.

The worst 5 of 2306 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:C:345:PRO:HG3	1:G:326:ILE:CD1	1.22	1.68
1:C:346:PHE:CD1	1:G:390:LYS:CE	1.87	1.58
1:C:345:PRO:CG	1:G:326:ILE:CD1	1.79	1.54
1:C:346:PHE:CD1	1:G:390:LYS:HE3	0.98	1.51
1:C:346:PHE:CE1	1:G:390:LYS:HG3	1.60	1.36

The worst 5 of 34 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:54:ASN:CA	2:H:88:TRP:CZ2[3_445]	0.68	1.52
1:C:53:GLU:O	2:H:88:TRP:NE1[3_445]	0.72	1.48
2:D:88:TRP:NE1	1:G:54:ASN:N[3_445]	0.91	1.29
1:C:53:GLU:O	2:H:88:TRP:CD1[3_445]	1.03	1.17
1:C:53:GLU:C	2:H:88:TRP:NE1[3_445]	1.08	1.12

## 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	530/560 (95%)	439 (83%)	74 (14%)	17 (3%)	<b>5</b> 30

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	530/560 (95%)	438 (83%)	75 (14%)	17 (3%)	5	30
1	E	530/560 (95%)	439 (83%)	74 (14%)	17 (3%)	5	30
1	G	530/560 (95%)	439 (83%)	74 (14%)	17 (3%)	5	30
2	B	389/440 (88%)	320 (82%)	57 (15%)	12 (3%)	5	31
2	D	389/440 (88%)	320 (82%)	57 (15%)	12 (3%)	5	31
2	F	389/440 (88%)	320 (82%)	57 (15%)	12 (3%)	5	31
2	H	389/440 (88%)	320 (82%)	57 (15%)	12 (3%)	5	31
All	All	3676/4000 (92%)	3035 (83%)	525 (14%)	116 (3%)	5	30

5 of 116 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	222	GLN
1	A	345	PRO
2	B	94	ILE
1	C	222	GLN
1	C	345	PRO

### 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	431/500 (86%)	411 (95%)	20 (5%)	31	70
1	C	431/500 (86%)	411 (95%)	20 (5%)	31	70
1	E	431/500 (86%)	410 (95%)	21 (5%)	29	68
1	G	431/500 (86%)	411 (95%)	20 (5%)	31	70
2	B	343/400 (86%)	323 (94%)	20 (6%)	23	62
2	D	343/400 (86%)	323 (94%)	20 (6%)	23	62
2	F	343/400 (86%)	323 (94%)	20 (6%)	23	62
2	H	343/400 (86%)	323 (94%)	20 (6%)	23	62
All	All	3096/3600 (86%)	2935 (95%)	161 (5%)	27	65

5 of 161 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	170	PRO
1	E	252	TRP
2	H	63	ILE
2	D	266	TRP
1	E	4	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 100 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	208	HIS
1	E	175	ASN
2	H	145	GLN
2	D	235	HIS
1	E	23	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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.