



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 6, 2026 – 07:39 PM UTC

PDB ID : 4PW8 / pdb_00004pw8
Title : Human tryptophan 2,3-dioxygenase
Authors : Meng, B.; Wu, D.; Gu, J.H.; Ouyang, S.Y.; Ding, W.; Liu, Z.J.
Deposited on : 2014-03-19
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
Xtriage (Phenix)	:	2.0
EDS	:	3.0
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

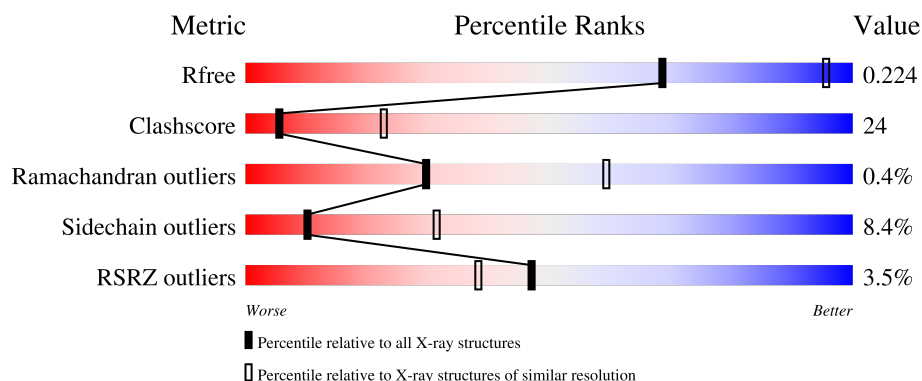
1 Overall quality at a glance

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	373	2% 54% 24% • 18%
1	B	373	3% 57% 16% 5% • 21%
1	C	373	2% 57% 20% 6% • 16%
1	D	373	3% 49% 22% 5% • 23%
1	E	373	3% 50% 22% 5% • 21%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	373	3% 53% 22% 5% 19%
1	G	373	4% 50% 24% 6% 18%
1	H	373	3% 61% 18% 19%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 20451 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 Tryptophan 2,3-dioxygenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	306	Total	C	N	O	S	0	0	0
			2592	1669	447	466	10			
1	B	296	Total	C	N	O	S	0	0	0
			2510	1627	432	441	10			
1	C	313	Total	C	N	O	S	0	0	0
			2660	1711	463	475	11			
1	D	288	Total	C	N	O	S	0	0	0
			2449	1587	424	428	10			
1	E	293	Total	C	N	O	S	0	0	0
			2481	1604	428	439	10			
1	F	303	Total	C	N	O	S	0	0	0
			2564	1658	444	452	10			
1	G	304	Total	C	N	O	S	0	0	0
			2582	1668	446	458	10			
1	H	301	Total	C	N	O	S	0	0	0
			2557	1651	441	455	10			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	16	SER	-	expression tag	UNP P48775
A	17	ASN	-	expression tag	UNP P48775
A	18	ALA	-	expression tag	UNP P48775
B	16	SER	-	expression tag	UNP P48775
B	17	ASN	-	expression tag	UNP P48775
B	18	ALA	-	expression tag	UNP P48775
C	16	SER	-	expression tag	UNP P48775
C	17	ASN	-	expression tag	UNP P48775
C	18	ALA	-	expression tag	UNP P48775
D	16	SER	-	expression tag	UNP P48775
D	17	ASN	-	expression tag	UNP P48775
D	18	ALA	-	expression tag	UNP P48775
E	16	SER	-	expression tag	UNP P48775

Continued on next page...

Continued from previous page...

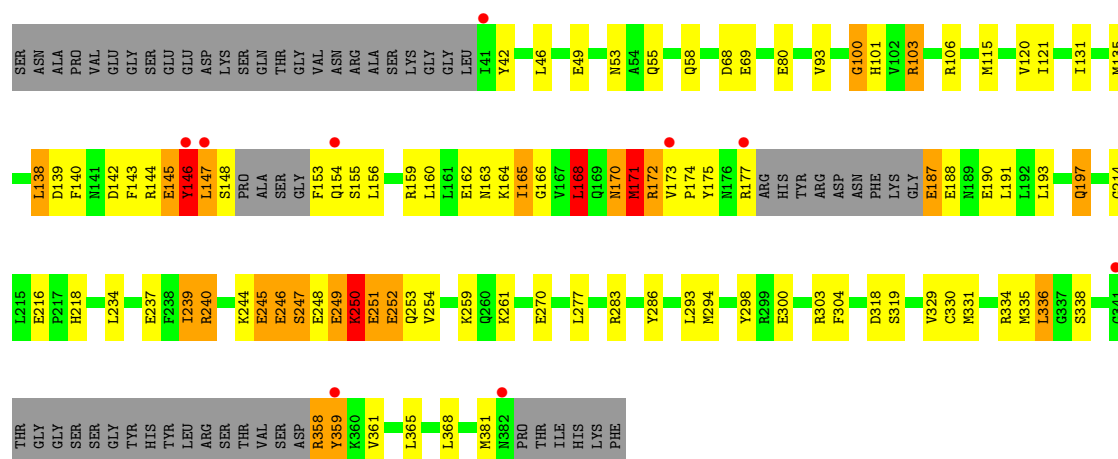
Chain	Residue	Modelled	Actual	Comment	Reference
E	17	ASN	-	expression tag	UNP P48775
E	18	ALA	-	expression tag	UNP P48775
F	16	SER	-	expression tag	UNP P48775
F	17	ASN	-	expression tag	UNP P48775
F	18	ALA	-	expression tag	UNP P48775
G	16	SER	-	expression tag	UNP P48775
G	17	ASN	-	expression tag	UNP P48775
G	18	ALA	-	expression tag	UNP P48775
H	16	SER	-	expression tag	UNP P48775
H	17	ASN	-	expression tag	UNP P48775
H	18	ALA	-	expression tag	UNP P48775

- Molecule 2 is COBALT (II) ION (CCD ID: CO) (formula: Co).

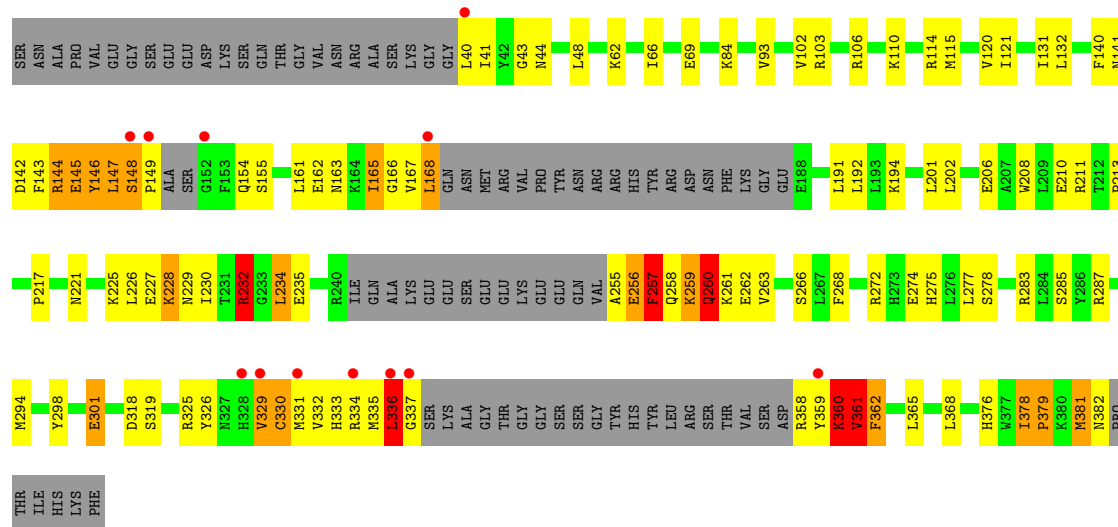
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	H	1	Total Co 1 1	0	0

- Molecule 3 is water.

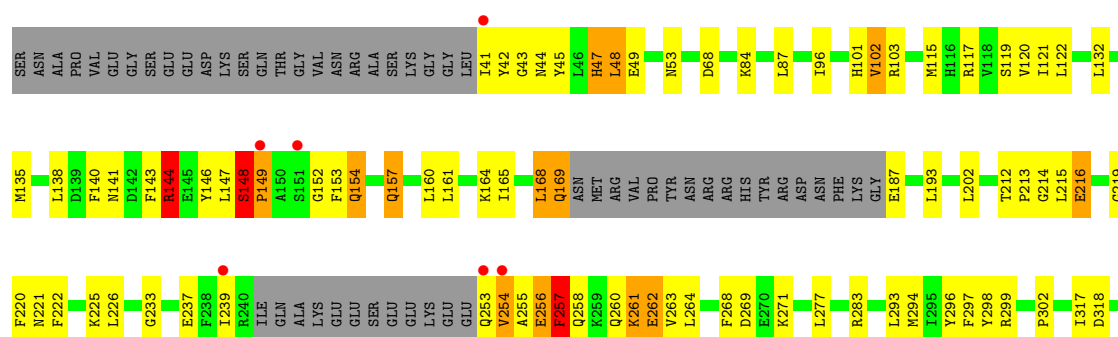
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	11	Total O 11 11	0	0
3	B	5	Total O 5 5	0	0
3	C	5	Total O 5 5	0	0
3	E	13	Total O 13 13	0	0
3	F	7	Total O 7 7	0	0
3	G	2	Total O 2 2	0	0
3	H	12	Total O 12 12	0	0

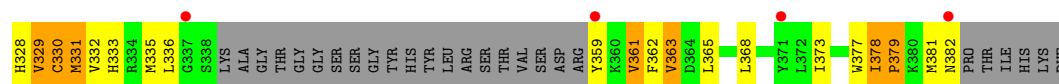


• Molecule 1: Tryptophan 2,3-dioxygenase

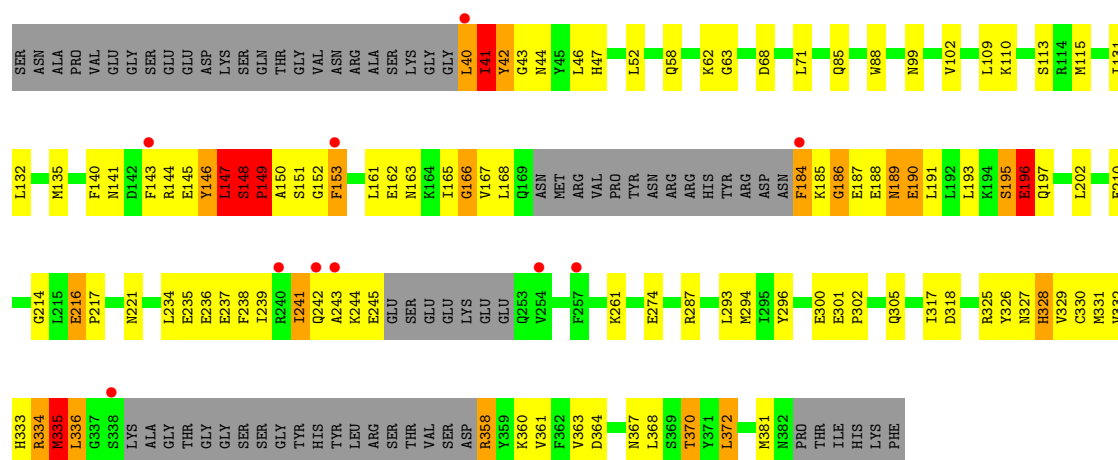


• Molecule 1: Tryptophan 2,3-dioxygenase

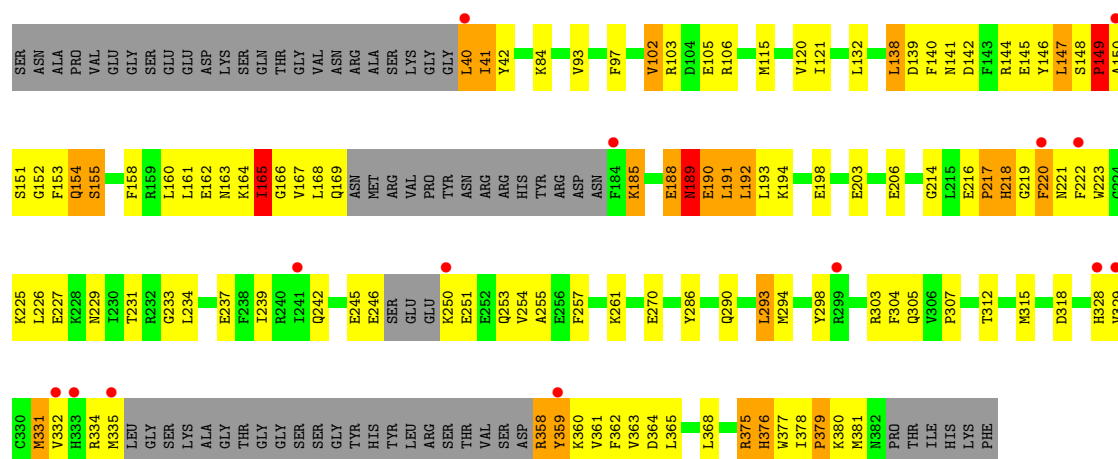




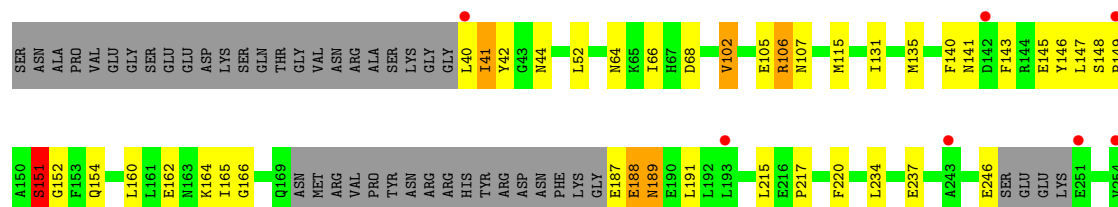
• Molecule 1: Tryptophan 2,3-dioxygenase

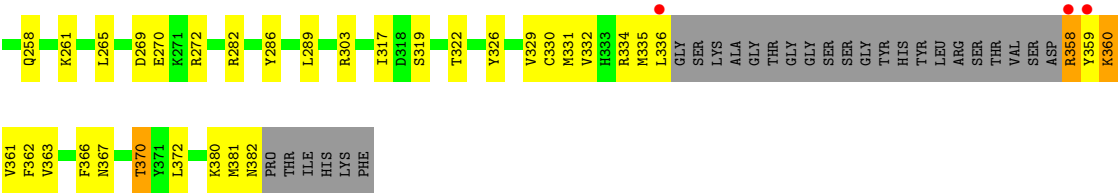


• Molecule 1: Tryptophan 2,3-dioxygenase



• Molecule 1: Tryptophan 2,3-dioxygenase





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	134.33Å 156.92Å 160.33Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.73 – 2.90 49.73 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.73-2.90) 96.0 (49.73-2.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.54 (at 2.91Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, R_{free}	0.204 , 0.222 0.208 , 0.224	Depositor DCC
R_{free} test set	2006 reflections (2.66%)	wwPDB-VP
Wilson B-factor (Å ²)	59.4	Xtriage
Anisotropy	0.493	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 52.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.014 for -h,l,k	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	20451	wwPDB-VP
Average B, all atoms (Å ²)	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.32% 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:
CO

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	1.09	2/2648 (0.1%)	1.00	13/3561 (0.4%)
1	B	0.84	4/2566 (0.2%)	1.06	15/3452 (0.4%)
1	C	0.88	0/2716	1.01	15/3649 (0.4%)
1	D	0.76	2/2503 (0.1%)	1.01	17/3365 (0.5%)
1	E	0.83	7/2536 (0.3%)	0.98	14/3412 (0.4%)
1	F	0.77	1/2620 (0.0%)	1.02	19/3522 (0.5%)
1	G	0.82	1/2638 (0.0%)	0.98	7/3545 (0.2%)
1	H	0.68	0/2612	0.91	5/3513 (0.1%)
All	All	0.84	17/20839 (0.1%)	1.00	105/28019 (0.4%)

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	C	0	1

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	379	PRO	N-CD	7.18	1.57	1.47
1	F	149	PRO	N-CD	5.71	1.55	1.47
1	E	148	SER	C-N	5.62	1.40	1.33
1	A	163	ASN	C-O	-5.56	1.17	1.24
1	E	379	PRO	N-CD	-5.54	1.40	1.47

The worst 5 of 105 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	186	GLY	N-CA-C	11.32	125.63	112.50
1	H	151	SER	N-CA-C	10.84	122.84	111.14
1	B	150	ALA	N-CA-C	-10.43	88.59	110.80
1	H	165	ILE	N-CA-C	10.36	120.27	110.53
1	B	218	HIS	N-CA-C	-10.25	97.31	110.53

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	146	TYR	Peptide

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	2592	0	2588	86	0
1	B	2510	0	2520	97	0
1	C	2660	0	2660	132	0
1	D	2449	0	2460	151	0
1	E	2481	0	2483	139	0
1	F	2564	0	2575	146	0
1	G	2582	0	2587	197	0
1	H	2557	0	2560	67	0
2	H	1	0	0	0	0
3	A	11	0	0	2	0
3	B	5	0	0	0	0
3	C	5	0	0	3	0
3	E	13	0	0	5	0
3	F	7	0	0	2	0
3	G	2	0	0	1	0
3	H	12	0	0	1	0
All	All	20451	0	20433	962	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 24.

The worst 5 of 962 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:145:GLU:O	1:C:146:TYR:CD1	1.83	1.29
1:G:146:TYR:C	1:G:147:LEU:HD23	1.63	1.22
1:H:187:GLU:HG3	1:H:188:GLU:H	1.02	1.16
1:B:375:ARG:HA	1:B:378:ILE:HD12	1.15	1.13
1:E:144:ARG:NH1	1:E:336:LEU:HD11	1.65	1.11

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	300/373 (80%)	294 (98%)	6 (2%)	0	100	100
1	B	288/373 (77%)	279 (97%)	8 (3%)	1 (0%)	36	65
1	C	305/373 (82%)	297 (97%)	6 (2%)	2 (1%)	18	48
1	D	278/373 (74%)	269 (97%)	9 (3%)	0	100	100
1	E	285/373 (76%)	278 (98%)	7 (2%)	0	100	100
1	F	295/373 (79%)	287 (97%)	5 (2%)	3 (1%)	12	39
1	G	296/373 (79%)	285 (96%)	8 (3%)	3 (1%)	12	39
1	H	293/373 (79%)	287 (98%)	6 (2%)	0	100	100
All	All	2340/2984 (78%)	2276 (97%)	55 (2%)	9 (0%)	30	59

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	150	ALA
1	C	247	SER
1	C	250	LYS
1	F	42	TYR
1	F	148	SER

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	285/340 (84%)	265 (93%)	20 (7%)	14	40
1	B	276/340 (81%)	248 (90%)	28 (10%)	7	24
1	C	292/340 (86%)	265 (91%)	27 (9%)	8	27
1	D	269/340 (79%)	244 (91%)	25 (9%)	8	27
1	E	273/340 (80%)	253 (93%)	20 (7%)	13	38
1	F	281/340 (83%)	252 (90%)	29 (10%)	7	23
1	G	283/340 (83%)	258 (91%)	25 (9%)	9	29
1	H	281/340 (83%)	266 (95%)	15 (5%)	20	52
All	All	2240/2720 (82%)	2051 (92%)	189 (8%)	10	31

5 of 189 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	329	VAL
1	F	334	ARG
1	E	363	VAL
1	F	184	PHE
1	G	41	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	127	GLN
1	G	128	GLN
1	H	127	GLN
1	B	260	GLN
1	B	218	HIS

5.3.3 RNA ⓘ

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 1 ligands modelled in this entry, 1 is 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 [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	306/373 (82%)	-0.22	7 (2%) 61 52	24, 44, 97, 131	0
1	B	296/373 (79%)	0.12	11 (3%) 45 37	29, 65, 129, 148	0
1	C	313/373 (83%)	-0.08	9 (2%) 53 45	36, 59, 101, 138	0
1	D	288/373 (77%)	0.29	12 (4%) 40 32	42, 76, 129, 147	0
1	E	293/373 (78%)	0.09	10 (3%) 48 39	33, 67, 125, 155	0
1	F	303/373 (81%)	0.10	10 (3%) 49 40	31, 66, 119, 143	0
1	G	304/373 (81%)	0.30	14 (4%) 37 29	32, 74, 126, 140	0
1	H	301/373 (80%)	-0.07	10 (3%) 49 40	27, 51, 103, 142	0
All	All	2404/2984 (80%)	0.06	83 (3%) 47 38	24, 62, 120, 155	0

The worst 5 of 83 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	337	GLY	6.8
1	B	254	VAL	4.4
1	G	335	MET	4.2
1	C	41	ILE	3.8
1	H	336	LEU	3.8

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
2	CO	H	401	1/1	0.97	0.13	57,57,57,57	0

6.5 Other polymers [i](#)

There are no such residues in this entry.