



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:58 AM GMT

PDB ID : 3CTM  
Title : Crystal Structure of a Carbonyl Reductase from Candida Parapsilosis with anti-Prelog Stereo-specificity  
Authors : Zhang, R.; Zhu, G.; Li, X.; Xu, Y.; Zhang, X.C.; Rao, Z.  
Deposited on : 2008-04-14  
Resolution : 2.69 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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

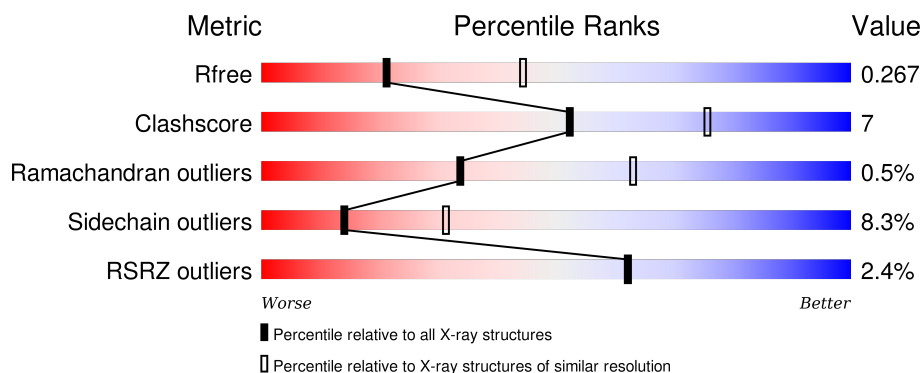
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.69 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))
$R_{free}$	91344	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (2.70-2.70)

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.

Mol	Chain	Length	Quality of chain
1	A	279	<div> <div>3%</div> <div>79%</div> <div>16%</div> <div>• •</div> </div>
1	B	279	<div> <div>2%</div> <div>77%</div> <div>18%</div> <div>• •</div> </div>
1	C	279	<div> <div>3%</div> <div>77%</div> <div>13%</div> <div>• 5%</div> </div>
1	D	279	<div> <div>%</div> <div>77%</div> <div>18%</div> <div>• •</div> </div>
1	E	279	<div> <div>2%</div> <div>80%</div> <div>13%</div> <div>• 5%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	279	<div> <div>4%</div> <div>77%</div> <div>16%</div> <div>• •</div> </div>
1	G	279	<div> <div>3%</div> <div>80%</div> <div>14%</div> <div>• 5%</div> </div>
1	H	279	<div> <div>79%</div> <div>15%</div> <div>• •</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 16392 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 Carbonyl Reductase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	272	Total	C	N	O	S	0	0	0
			2072	1318	339	409	6			
1	B	271	Total	C	N	O	S	0	0	0
			2051	1306	336	402	7			
1	C	264	Total	C	N	O	S	0	0	0
			2014	1284	330	395	5			
1	D	269	Total	C	N	O	S	0	0	0
			2039	1299	334	400	6			
1	E	265	Total	C	N	O	S	0	0	0
			2020	1286	331	397	6			
1	F	268	Total	C	N	O	S	0	0	0
			2031	1293	334	398	6			
1	G	265	Total	C	N	O	S	0	0	0
			2021	1288	331	397	5			
1	H	268	Total	C	N	O	S	0	0	0
			2031	1295	333	397	6			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	10	Total	O	0	0
			10	10		
2	B	15	Total	O	0	0
			15	15		
2	C	10	Total	O	0	0
			10	10		
2	D	16	Total	O	0	0
			16	16		
2	E	11	Total	O	0	0
			11	11		
2	F	23	Total	O	0	0
			23	23		

*Continued on next page...*

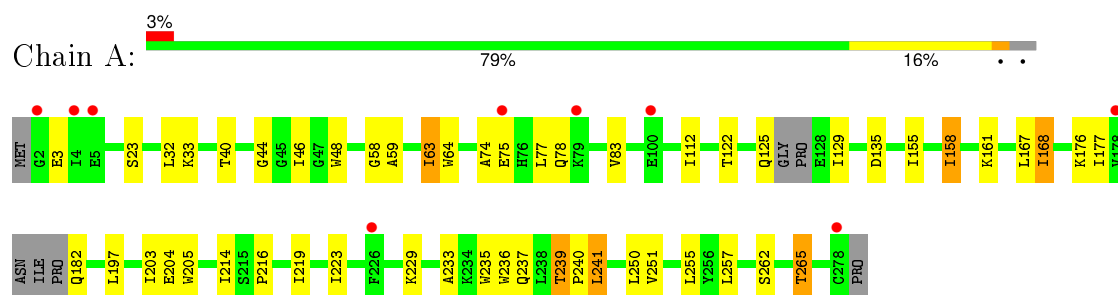
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	15	Total	O	0	0
			15	15		
2	H	13	Total	O	0	0
			13	13		

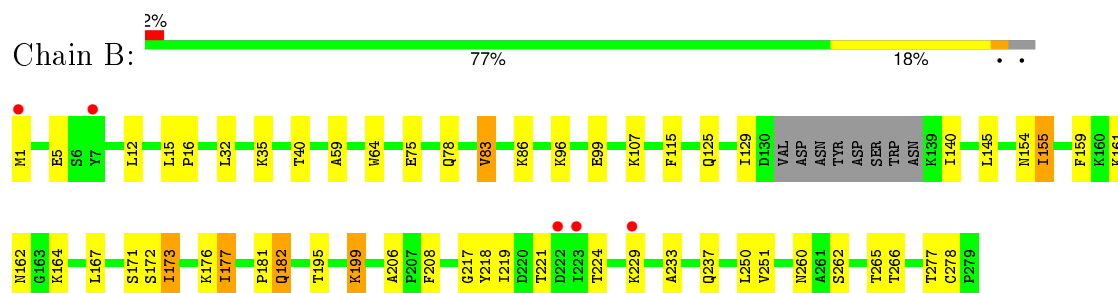
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

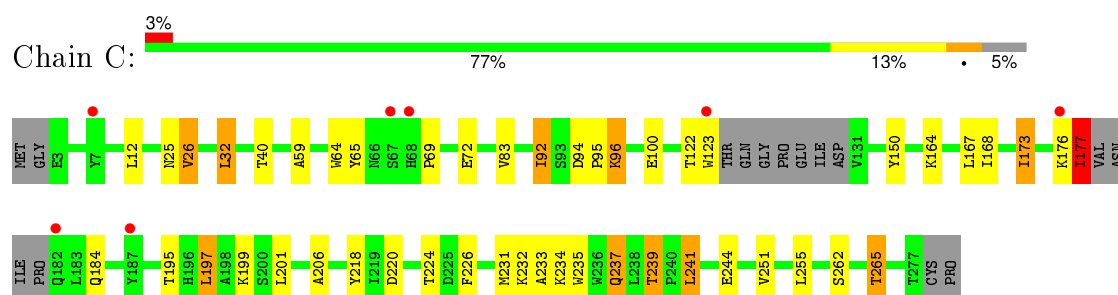
#### • Molecule 1: Carbonyl Reductase



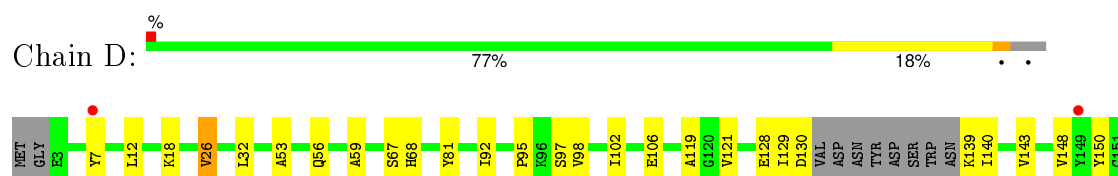
#### • Molecule 1: Carbonyl Reductase



#### • Molecule 1: Carbonyl Reductase

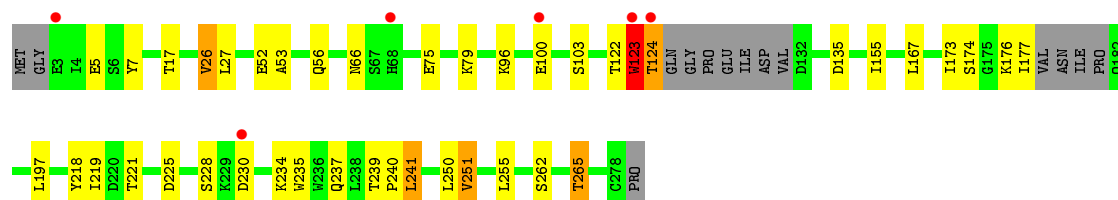
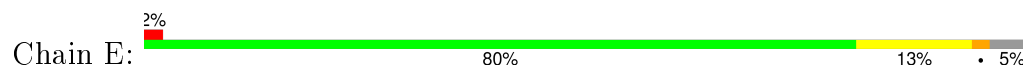


#### • Molecule 1: Carbonyl Reductase

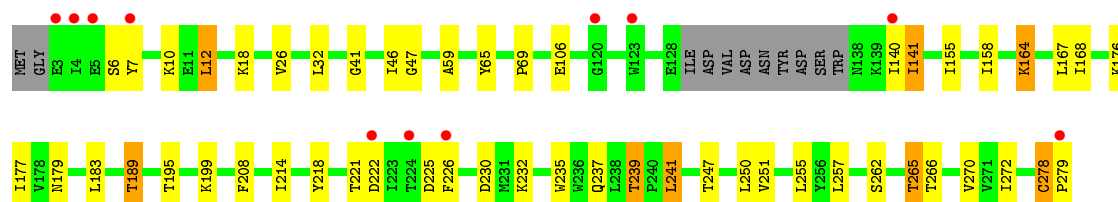
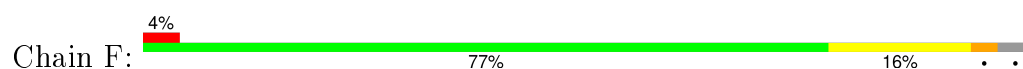




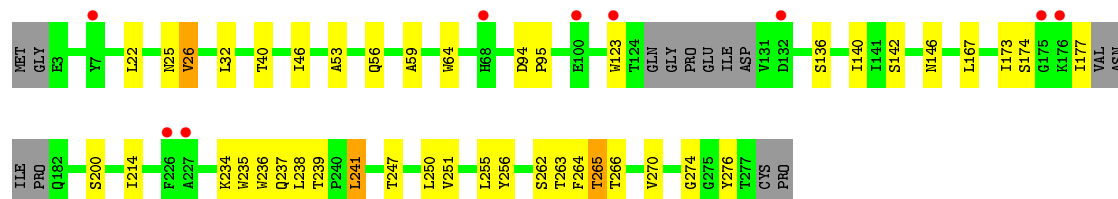
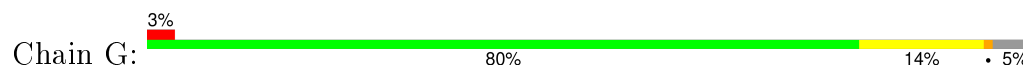
• Molecule 1: Carbonyl Reductase



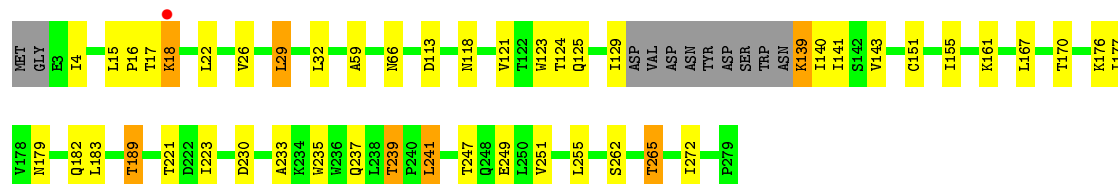
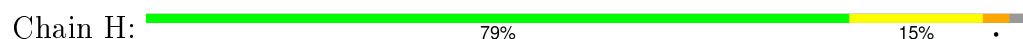
• Molecule 1: Carbonyl Reductase



• Molecule 1: Carbonyl Reductase



• Molecule 1: Carbonyl Reductase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	104.73Å 142.76Å 151.84Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.67 – 2.69 41.66 – 2.69	Depositor EDS
% Data completeness (in resolution range)	99.5 (41.67-2.69) 99.5 (41.66-2.69)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.86 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.205 , 0.264 0.209 , 0.267	Depositor DCC
$R_{free}$ test set	3215 reflections (5.35%)	DCC
Wilson B-factor (Å <sup>2</sup> )	43.6	Xtriage
Anisotropy	0.025	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 41.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 63276 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	16392	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.03% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.



## 5 Model quality

### 5.1 Standard geometry

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	0/2117	0.55	0/2876
1	B	0.46	0/2097	0.56	0/2849
1	C	0.49	1/2059 (0.0%)	0.57	1/2797 (0.0%)
1	D	0.48	0/2085	0.58	1/2834 (0.0%)
1	E	0.46	1/2065 (0.0%)	0.58	1/2805 (0.0%)
1	F	0.46	0/2077	0.58	0/2823
1	G	0.44	0/2066	0.56	0/2807
1	H	0.47	0/2077	0.57	0/2823
All	All	0.46	2/16643 (0.0%)	0.57	3/22614 (0.0%)

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	E	1	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	177	ILE	CG1-CD1	11.53	2.29	1.50
1	E	123	TRP	CB-CG	5.75	1.60	1.50

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	177	ILE	CB-CG1-CD1	-8.80	89.25	113.90
1	E	123	TRP	N-CA-C	5.21	125.07	111.00
1	D	12	LEU	CA-CB-CG	5.20	127.26	115.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	E	123	TRP	CA

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	122	THR	Peptide

## 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	2072	0	2025	27	0
1	B	2051	0	2026	29	0
1	C	2014	0	1972	38	0
1	D	2039	0	2011	31	0
1	E	2020	0	1975	22	0
1	F	2031	0	2002	36	0
1	G	2021	0	1979	34	0
1	H	2031	0	2007	37	0
2	A	10	0	0	0	0
2	B	15	0	0	1	0
2	C	10	0	0	0	0
2	D	16	0	0	1	0
2	E	11	0	0	0	0
2	F	23	0	0	0	0
2	G	15	0	0	1	0
2	H	13	0	0	0	0
All	All	16392	0	15997	215	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 7.

The worst 5 of 215 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:177:ILE:HG21	1:C:177:ILE:CD1	1.66	1.23
1:C:177:ILE:CG2	1:C:177:ILE:CD1	2.30	1.09

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:177:ILE:CG1	1:C:177:ILE:CD1	2.30	1.09
1:G:265:THR:HG22	1:H:272:ILE:HG23	1.44	0.96
1:F:278:CYS:H	1:F:279:PRO:HD2	1.30	0.94

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	266/279 (95%)	254 (96%)	10 (4%)	2 (1%)	24	51
1	B	267/279 (96%)	254 (95%)	13 (5%)	0	100	100
1	C	258/279 (92%)	250 (97%)	8 (3%)	0	100	100
1	D	265/279 (95%)	252 (95%)	11 (4%)	2 (1%)	24	51
1	E	259/279 (93%)	249 (96%)	8 (3%)	2 (1%)	24	51
1	F	264/279 (95%)	246 (93%)	15 (6%)	3 (1%)	17	42
1	G	259/279 (93%)	252 (97%)	6 (2%)	1 (0%)	39	69
1	H	264/279 (95%)	254 (96%)	10 (4%)	0	100	100
All	All	2102/2232 (94%)	2011 (96%)	81 (4%)	10 (0%)	34	63

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	123	TRP
1	F	278	CYS
1	A	129	ILE
1	G	123	TRP
1	E	228	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	222/228 (97%)	201 (90%)	21 (10%)	11	24
1	B	220/228 (96%)	202 (92%)	18 (8%)	14	32
1	C	215/228 (94%)	194 (90%)	21 (10%)	10	23
1	D	219/228 (96%)	200 (91%)	19 (9%)	13	29
1	E	216/228 (95%)	196 (91%)	20 (9%)	11	25
1	F	218/228 (96%)	203 (93%)	15 (7%)	19	43
1	G	216/228 (95%)	205 (95%)	11 (5%)	29	59
1	H	218/228 (96%)	198 (91%)	20 (9%)	11	25
All	All	1744/1824 (96%)	1599 (92%)	145 (8%)	14	31

5 of 145 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	128	GLU
1	E	26	VAL
1	H	139	LYS
1	D	173	ILE
1	D	239	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	104	GLN
1	E	91	ASN
1	H	118	ASN
1	D	105	GLN
1	E	66	ASN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 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 ⓘ

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	272/279 (97%)	0.12	9 (3%)	50	50	22, 47, 78, 91	0
1	B	271/279 (97%)	-0.22	5 (1%)	71	72	15, 34, 66, 89	0
1	C	264/279 (94%)	0.02	7 (2%)	58	58	20, 47, 75, 86	0
1	D	269/279 (96%)	-0.13	3 (1%)	82	83	19, 39, 70, 100	0
1	E	265/279 (94%)	-0.07	6 (2%)	64	64	18, 39, 75, 91	0
1	F	268/279 (96%)	-0.16	11 (4%)	41	41	15, 32, 70, 80	0
1	G	265/279 (94%)	0.00	9 (3%)	49	49	19, 44, 79, 92	0
1	H	268/279 (96%)	-0.26	1 (0%)	93	94	17, 34, 69, 89	0
All	All	2142/2232 (95%)	-0.09	51 (2%)	62	62	15, 39, 74, 100	0

The worst 5 of 51 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	2	GLY	5.9
1	E	123	TRP	5.5
1	E	124	THR	5.1
1	G	175	GLY	4.7
1	F	3	GLU	4.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.