



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 07:31 am BST

PDB ID : 1QGN  
Title : CYSTATHIONINE GAMMA-SYNTHASE FROM NICOTIANA TABACUM  
Authors : Steegborn, C.; Messerschmidt, A.; Laber, B.; Streber, W.; Huber, R.; Clausen, T.  
Deposited on : 1999-05-02  
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](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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

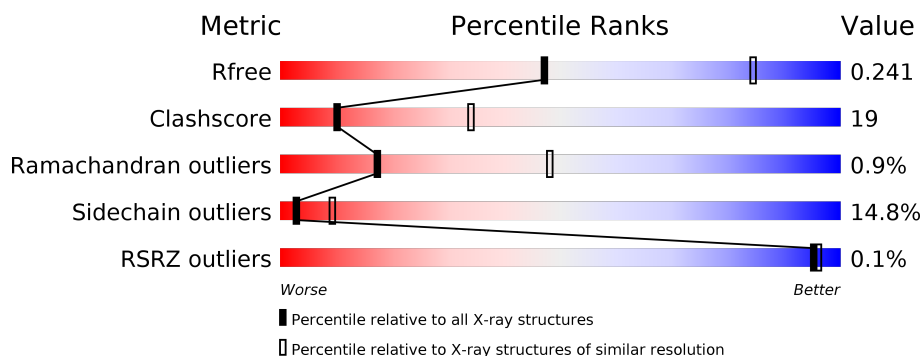
# 1 Overall quality at a glance

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))
$R_{free}$	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (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 on 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	445	<div> <div>51%</div> <div>33%</div> <div>5%</div> <div>11%</div> </div>
1	B	445	<div> <div>55%</div> <div>29%</div> <div>5%</div> <div>11%</div> </div>
1	C	445	<div> <div>51%</div> <div>33%</div> <div>5%</div> <div>11%</div> </div>
1	D	445	<div> <div>53%</div> <div>30%</div> <div>6%</div> <div>11%</div> </div>
1	E	445	<div> <div>56%</div> <div>29%</div> <div>•</div> <div>11%</div> </div>
1	F	445	<div> <div>52%</div> <div>32%</div> <div>6%</div> <div>11%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	445	<div><div></div><div>54%</div><div>30%</div><div>5%</div><div>11%</div></div>
1	H	445	<div><div></div><div>55%</div><div>30%</div><div>5%</div><div>11%</div></div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 24702 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 PROTEIN (CYSTATHIONINE GAMMA-SYNTHASE).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	398	Total	C	N	O	S	0	0	0
			3032	1933	514	569	16			
1	B	398	Total	C	N	O	S	0	0	0
			3032	1933	514	569	16			
1	C	398	Total	C	N	O	S	0	0	0
			3032	1933	514	569	16			
1	D	398	Total	C	N	O	S	0	0	0
			3032	1933	514	569	16			
1	E	398	Total	C	N	O	S	0	0	0
			3032	1933	514	569	16			
1	F	398	Total	C	N	O	S	0	0	0
			3032	1933	514	569	16			
1	G	398	Total	C	N	O	S	0	0	0
			3032	1933	514	569	16			
1	H	398	Total	C	N	O	S	0	0	0
			3032	1933	514	569	16			

- Molecule 2 is PYRIDOXAL-5'-PHOSPHATE (three-letter code: PLP) (formula: C<sub>8</sub>H<sub>10</sub>NO<sub>6</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	B	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	C	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	D	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	E	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	F	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	G	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	H	1	Total	C	N	O	P	0	0
			15	8	1	5	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	44	Total	O	0	0
			44	44		
3	B	41	Total	O	0	0
			41	41		
3	C	41	Total	O	0	0
			41	41		
3	D	37	Total	O	0	0
			37	37		

*Continued on next page...*

*Continued from previous page...*

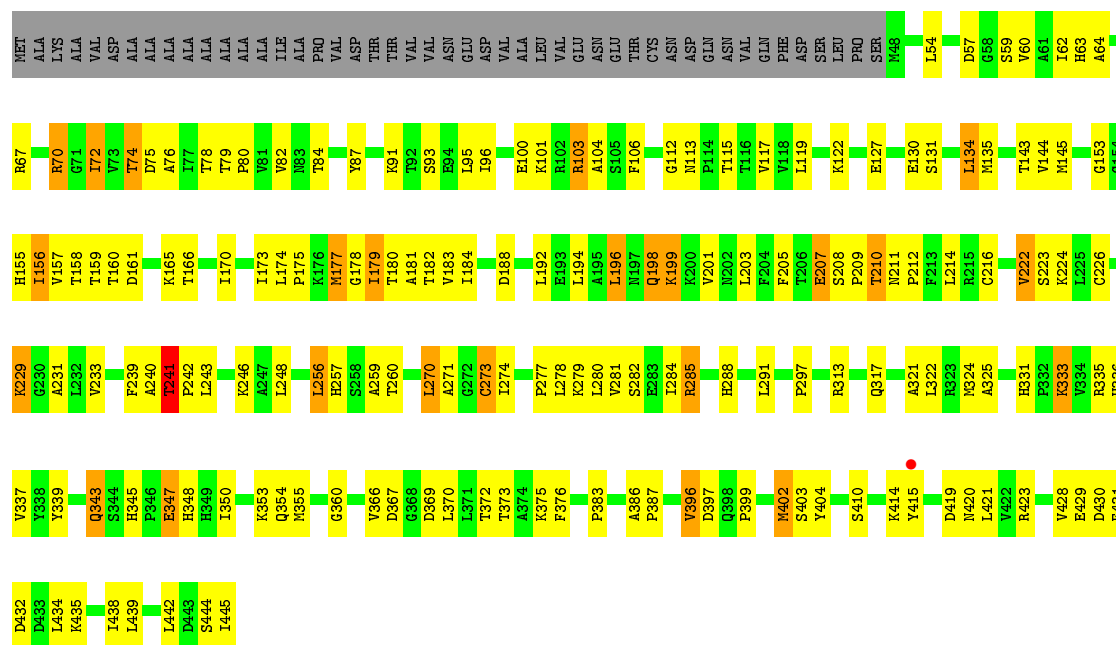
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	39	Total 39	O 39	0	0
3	F	42	Total 42	O 42	0	0
3	G	41	Total 41	O 41	0	0
3	H	41	Total 41	O 41	0	0





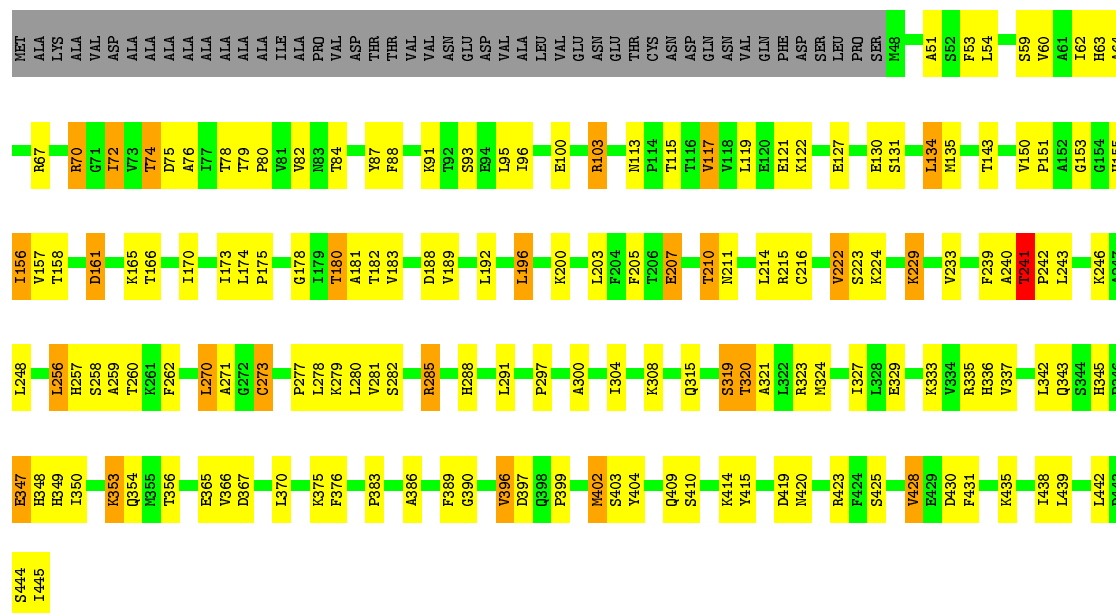
• Molecule 1: PROTEIN (CYSTATHIONINE GAMMA-SYNTHASE)

Chain C: 51% 33% 5% 11%



• Molecule 1: PROTEIN (CYSTATHIONINE GAMMA-SYNTHASE)

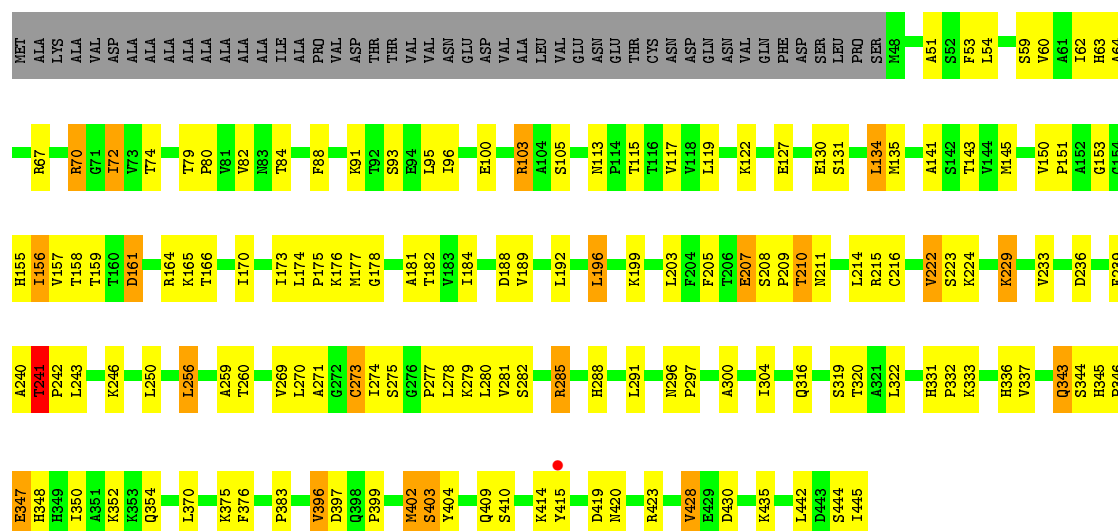
Chain D: 53% 30% 6% 11%



• Molecule 1: PROTEIN (CYSTATHIONINE GAMMA-SYNTHASE)

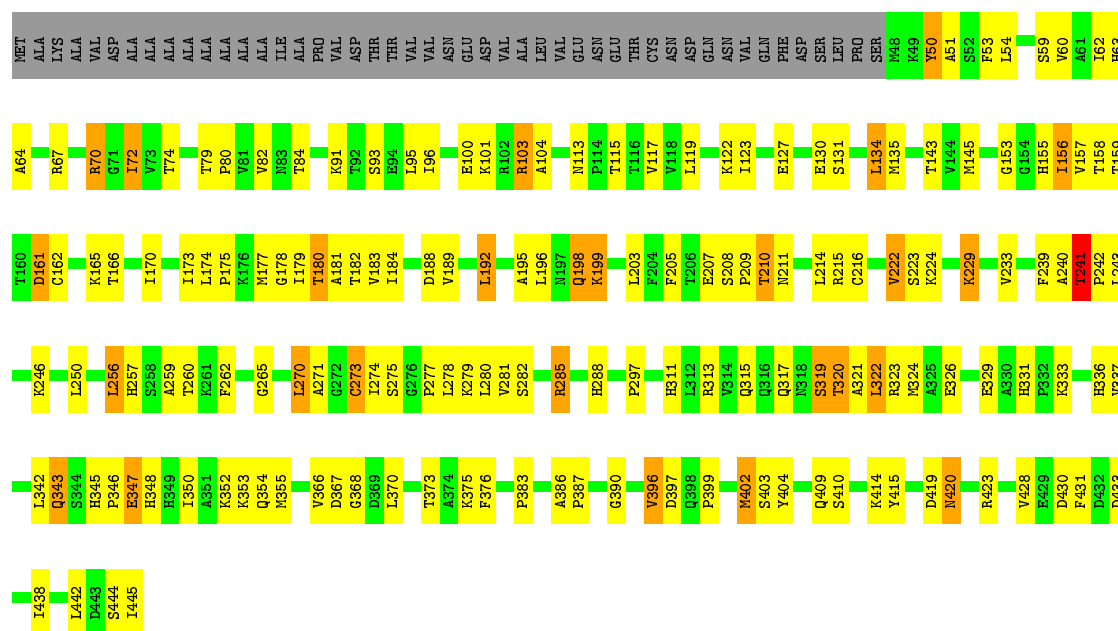


Chain E: 



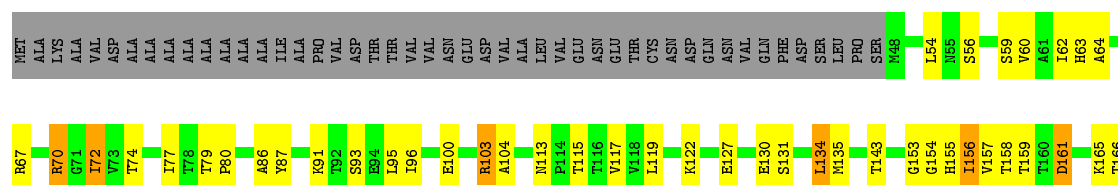
• Molecule 1: PROTEIN (CYSTATHIONINE GAMMA-SYNTHASE)

Chain F: 



• Molecule 1: PROTEIN (CYSTATHIONINE GAMMA-SYNTHASE)

Chain G: 





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	120.00Å 129.50Å 309.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 2.90 24.95 – 2.79	Depositor EDS
% Data completeness (in resolution range)	88.2 (8.00-2.90) 86.2 (24.95-2.79)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.37 (at 2.80Å)	Xtriage
Refinement program	X-PLOR	Depositor
R, $R_{free}$	0.201 , 0.250 0.231 , 0.241	Depositor DCC
$R_{free}$ test set	5276 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	50.9	Xtriage
Anisotropy	0.383	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 29.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	24702	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.12% 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: PLP

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	0.36	0/3091	0.59	1/4190 (0.0%)
1	B	0.35	0/3091	0.58	1/4190 (0.0%)
1	C	0.33	0/3091	0.58	1/4190 (0.0%)
1	D	0.35	0/3091	0.59	1/4190 (0.0%)
1	E	0.35	0/3091	0.58	1/4190 (0.0%)
1	F	0.34	0/3091	0.57	1/4190 (0.0%)
1	G	0.34	0/3091	0.58	1/4190 (0.0%)
1	H	0.34	0/3091	0.58	1/4190 (0.0%)
All	All	0.34	0/24728	0.58	8/33520 (0.0%)

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	414	LYS	CD-CE-NZ	5.57	124.50	111.70
1	F	414	LYS	CD-CE-NZ	5.56	124.48	111.70
1	B	414	LYS	CD-CE-NZ	5.52	124.40	111.70
1	E	414	LYS	CD-CE-NZ	5.50	124.35	111.70
1	D	414	LYS	CD-CE-NZ	5.47	124.29	111.70

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	3032	0	3054	125	0
1	B	3032	0	3055	117	0
1	C	3032	0	3055	129	0
1	D	3032	0	3055	120	0
1	E	3032	0	3055	116	0
1	F	3032	0	3055	131	0
1	G	3032	0	3055	121	0
1	H	3032	0	3055	114	0
2	A	15	0	6	0	0
2	B	15	0	6	0	0
2	C	15	0	6	0	0
2	D	15	0	6	0	0
2	E	15	0	6	0	0
2	F	15	0	6	0	0
2	G	15	0	6	0	0
2	H	15	0	6	0	0
3	A	44	0	0	2	0
3	B	41	0	0	3	0
3	C	41	0	0	3	0
3	D	37	0	0	2	0
3	E	39	0	0	1	0
3	F	42	0	0	3	0
3	G	41	0	0	1	0
3	H	41	0	0	3	0
All	All	24702	0	24487	925	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 19.

The worst 5 of 925 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:241:THR:HG22	1:C:243:LEU:H	1.22	1.02
1:H:241:THR:HG22	1:H:243:LEU:H	1.25	1.02
1:B:241:THR:HG22	1:B:243:LEU:H	1.23	1.01
1:E:241:THR:HG22	1:E:243:LEU:H	1.20	1.00
1:G:241:THR:HG22	1:G:243:LEU:H	1.26	0.99

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	396/445 (89%)	360 (91%)	31 (8%)	5 (1%)	12	37
1	B	396/445 (89%)	363 (92%)	29 (7%)	4 (1%)	15	45
1	C	396/445 (89%)	361 (91%)	32 (8%)	3 (1%)	19	51
1	D	396/445 (89%)	364 (92%)	28 (7%)	4 (1%)	15	45
1	E	396/445 (89%)	363 (92%)	30 (8%)	3 (1%)	19	51
1	F	396/445 (89%)	356 (90%)	35 (9%)	5 (1%)	12	37
1	G	396/445 (89%)	363 (92%)	31 (8%)	2 (0%)	29	61
1	H	396/445 (89%)	364 (92%)	29 (7%)	3 (1%)	19	51
All	All	3168/3560 (89%)	2894 (91%)	245 (8%)	29 (1%)	17	48

5 of 29 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	188	ASP
1	B	188	ASP
1	C	188	ASP
1	D	320	THR
1	E	188	ASP

### 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	327/364 (90%)	275 (84%)	52 (16%)	2	7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	327/364 (90%)	280 (86%)	47 (14%)	3	9
1	C	327/364 (90%)	278 (85%)	49 (15%)	3	9
1	D	327/364 (90%)	276 (84%)	51 (16%)	2	8
1	E	327/364 (90%)	282 (86%)	45 (14%)	3	10
1	F	327/364 (90%)	277 (85%)	50 (15%)	2	8
1	G	327/364 (90%)	284 (87%)	43 (13%)	4	12
1	H	327/364 (90%)	277 (85%)	50 (15%)	2	8
All	All	2616/2912 (90%)	2229 (85%)	387 (15%)	3	9

5 of 387 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	279	LYS
1	E	222	VAL
1	H	222	VAL
1	D	323	ARG
1	D	428	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	349	HIS
1	E	63	HIS
1	H	343	GLN
1	D	420	ASN
1	E	420	ASN

### 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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

8 ligands are 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$
2	PLP	B	500	1	15,15,16	3.54	4 (26%)	20,22,23	2.33	8 (40%)
2	PLP	D	500	1	15,15,16	3.47	5 (33%)	20,22,23	2.26	8 (40%)
2	PLP	E	500	1	15,15,16	3.43	5 (33%)	20,22,23	2.27	9 (45%)
2	PLP	G	500	1	15,15,16	3.72	4 (26%)	20,22,23	2.28	7 (35%)
2	PLP	A	500	1	15,15,16	3.62	6 (40%)	20,22,23	2.29	8 (40%)
2	PLP	C	500	1	15,15,16	3.83	5 (33%)	20,22,23	2.32	8 (40%)
2	PLP	H	500	1	15,15,16	3.60	4 (26%)	20,22,23	2.31	8 (40%)
2	PLP	F	500	1	15,15,16	3.43	5 (33%)	20,22,23	2.33	8 (40%)

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
2	PLP	B	500	1	-	0/6/6/8	0/1/1/1
2	PLP	D	500	1	-	0/6/6/8	0/1/1/1
2	PLP	E	500	1	-	0/6/6/8	0/1/1/1
2	PLP	G	500	1	-	0/6/6/8	0/1/1/1
2	PLP	A	500	1	-	0/6/6/8	0/1/1/1
2	PLP	C	500	1	-	0/6/6/8	0/1/1/1
2	PLP	H	500	1	-	0/6/6/8	0/1/1/1
2	PLP	F	500	1	-	0/6/6/8	0/1/1/1

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



Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	G	500	PLP	C5-C4	11.27	1.53	1.40
2	B	500	PLP	C5-C4	10.99	1.52	1.40
2	C	500	PLP	C5-C4	10.97	1.52	1.40
2	F	500	PLP	C5-C4	10.88	1.52	1.40
2	H	500	PLP	C5-C4	10.78	1.52	1.40

The worst 5 of 64 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	500	PLP	C2A-C2-C3	6.16	128.50	120.89
2	G	500	PLP	C2A-C2-C3	6.07	128.39	120.89
2	A	500	PLP	C2A-C2-C3	6.05	128.36	120.89
2	F	500	PLP	C2A-C2-C3	5.86	128.13	120.89
2	E	500	PLP	C2A-C2-C3	5.84	128.10	120.89

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, 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	398/445 (89%)	-0.19	0 100 100	12, 29, 57, 95	0
1	B	398/445 (89%)	-0.36	0 100 100	9, 28, 57, 92	0
1	C	398/445 (89%)	-0.27	1 (0%) 94 94	11, 31, 61, 95	0
1	D	398/445 (89%)	-0.32	0 100 100	9, 28, 58, 91	0
1	E	398/445 (89%)	-0.11	1 (0%) 94 94	9, 28, 57, 87	0
1	F	398/445 (89%)	-0.42	0 100 100	8, 31, 60, 89	0
1	G	398/445 (89%)	-0.10	1 (0%) 94 94	7, 31, 60, 89	0
1	H	398/445 (89%)	-0.38	0 100 100	9, 29, 56, 97	0
All	All	3184/3560 (89%)	-0.27	3 (0%) 95 96	7, 30, 58, 97	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	415	TYR	3.1
1	E	415	TYR	2.8
1	G	400	ALA	2.2

### 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 carbohydrates 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( $\text{\AA}^2$ )	Q<0.9
2	PLP	A	500	15/16	0.93	0.23	14,29,35,36	0
2	PLP	G	500	15/16	0.94	0.23	20,30,39,42	0
2	PLP	C	500	15/16	0.94	0.20	19,31,37,38	0
2	PLP	D	500	15/16	0.96	0.17	13,23,28,30	0
2	PLP	H	500	15/16	0.96	0.19	17,30,34,36	0
2	PLP	B	500	15/16	0.97	0.15	8,26,32,34	0
2	PLP	E	500	15/16	0.97	0.19	15,28,36,37	0
2	PLP	F	500	15/16	0.98	0.15	17,33,37,37	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.