



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 10:22 PM GMT

PDB ID : 1T6P  
Title : Crystal Structure of Phenylalanine Ammonia Lyase from Rhodosporidium toruloides  
Authors : Calabrese, J.C.; Jordan, D.B.  
Deposited on : 2004-05-06  
Resolution : 2.70 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

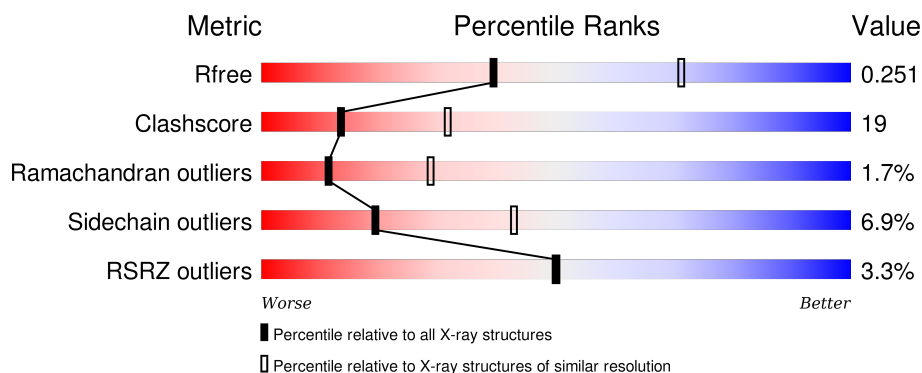
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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	714	<div> <div>3%</div> <div>61% 26% 8%</div> </div>
1	B	714	<div> <div>2%</div> <div>59% 30% 8%</div> </div>
1	C	714	<div> <div>3%</div> <div>53% 35% 9%</div> </div>
1	D	714	<div> <div>3%</div> <div>63% 25% 9%</div> </div>
1	E	714	<div> <div>4%</div> <div>58% 30% 8%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	714	
1	G	714	
1	H	714	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
1	175	A	211	X	-	-	-
1	175	B	211	X	-	-	-
1	175	C	211	X	-	-	-
1	175	D	211	X	-	-	-
1	175	E	211	X	-	-	-
1	175	F	211	X	-	-	-
1	175	G	211	X	-	-	-
1	175	H	211	X	-	-	-

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 40380 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 phenylalanine ammonia-lyase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	654	Total	C	N	O	S	Se	0	0	0
			4978	3131	884	946	4	13			
1	B	660	Total	C	N	O	S	Se	0	0	0
			5019	3153	893	956	4	13			
1	C	650	Total	C	N	O	S	Se	0	0	0
			4946	3110	882	937	4	13			
1	D	650	Total	C	N	O	S	Se	0	0	0
			4948	3111	880	940	4	13			
1	E	655	Total	C	N	O	S	Se	0	0	0
			4982	3133	886	946	4	13			
1	F	650	Total	C	N	O	S	Se	0	0	0
			4951	3113	882	939	4	13			
1	G	656	Total	C	N	O	S	Se	0	0	0
			4995	3140	890	948	4	13			
1	H	644	Total	C	N	O	S	Se	0	0	0
			4910	3088	872	933	4	13			

There are 144 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	51	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	107	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	169	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	211	175	ALA	SEE REMARK 999	UNP P11544
A	211	175	SER	SEE REMARK 999	UNP P11544
A	211	175	GLY	SEE REMARK 999	UNP P11544
A	250	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	278	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	288	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	299	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	304	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	422	MSE	MET	MODIFIED RESIDUE	UNP P11544

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	443	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	448	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	497	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	564	MSE	MET	MODIFIED RESIDUE	UNP P11544
A	714	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	1	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	51	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	107	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	169	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	211	175	ALA	SEE REMARK 999	UNP P11544
B	211	175	SER	SEE REMARK 999	UNP P11544
B	211	175	GLY	SEE REMARK 999	UNP P11544
B	250	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	278	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	288	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	299	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	304	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	422	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	443	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	448	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	497	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	564	MSE	MET	MODIFIED RESIDUE	UNP P11544
B	714	MSE	MET	MODIFIED RESIDUE	UNP P11544
C	1	MSE	MET	MODIFIED RESIDUE	UNP P11544
C	51	MSE	MET	MODIFIED RESIDUE	UNP P11544
C	107	MSE	MET	MODIFIED RESIDUE	UNP P11544
C	169	MSE	MET	MODIFIED RESIDUE	UNP P11544
C	211	175	ALA	SEE REMARK 999	UNP P11544
C	211	175	SER	SEE REMARK 999	UNP P11544
C	211	175	GLY	SEE REMARK 999	UNP P11544
C	250	MSE	MET	MODIFIED RESIDUE	UNP P11544
C	278	MSE	MET	MODIFIED RESIDUE	UNP P11544
C	288	MSE	MET	MODIFIED RESIDUE	UNP P11544
C	299	MSE	MET	MODIFIED RESIDUE	UNP P11544
C	304	MSE	MET	MODIFIED RESIDUE	UNP P11544
C	422	MSE	MET	MODIFIED RESIDUE	UNP P11544
C	443	MSE	MET	MODIFIED RESIDUE	UNP P11544
C	448	MSE	MET	MODIFIED RESIDUE	UNP P11544
C	497	MSE	MET	MODIFIED RESIDUE	UNP P11544
C	564	MSE	MET	MODIFIED RESIDUE	UNP P11544
C	714	MSE	MET	MODIFIED RESIDUE	UNP P11544
D	1	MSE	MET	MODIFIED RESIDUE	UNP P11544

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	51	MSE	MET	MODIFIED RESIDUE	UNP P11544
D	107	MSE	MET	MODIFIED RESIDUE	UNP P11544
D	169	MSE	MET	MODIFIED RESIDUE	UNP P11544
D	211	175	ALA	SEE REMARK 999	UNP P11544
D	211	175	SER	SEE REMARK 999	UNP P11544
D	211	175	GLY	SEE REMARK 999	UNP P11544
D	250	MSE	MET	MODIFIED RESIDUE	UNP P11544
D	278	MSE	MET	MODIFIED RESIDUE	UNP P11544
D	288	MSE	MET	MODIFIED RESIDUE	UNP P11544
D	299	MSE	MET	MODIFIED RESIDUE	UNP P11544
D	304	MSE	MET	MODIFIED RESIDUE	UNP P11544
D	422	MSE	MET	MODIFIED RESIDUE	UNP P11544
D	443	MSE	MET	MODIFIED RESIDUE	UNP P11544
D	448	MSE	MET	MODIFIED RESIDUE	UNP P11544
D	497	MSE	MET	MODIFIED RESIDUE	UNP P11544
D	564	MSE	MET	MODIFIED RESIDUE	UNP P11544
D	714	MSE	MET	MODIFIED RESIDUE	UNP P11544
E	1	MSE	MET	MODIFIED RESIDUE	UNP P11544
E	51	MSE	MET	MODIFIED RESIDUE	UNP P11544
E	107	MSE	MET	MODIFIED RESIDUE	UNP P11544
E	169	MSE	MET	MODIFIED RESIDUE	UNP P11544
E	211	175	ALA	SEE REMARK 999	UNP P11544
E	211	175	SER	SEE REMARK 999	UNP P11544
E	211	175	GLY	SEE REMARK 999	UNP P11544
E	250	MSE	MET	MODIFIED RESIDUE	UNP P11544
E	278	MSE	MET	MODIFIED RESIDUE	UNP P11544
E	288	MSE	MET	MODIFIED RESIDUE	UNP P11544
E	299	MSE	MET	MODIFIED RESIDUE	UNP P11544
E	304	MSE	MET	MODIFIED RESIDUE	UNP P11544
E	422	MSE	MET	MODIFIED RESIDUE	UNP P11544
E	443	MSE	MET	MODIFIED RESIDUE	UNP P11544
E	448	MSE	MET	MODIFIED RESIDUE	UNP P11544
E	497	MSE	MET	MODIFIED RESIDUE	UNP P11544
E	564	MSE	MET	MODIFIED RESIDUE	UNP P11544
E	714	MSE	MET	MODIFIED RESIDUE	UNP P11544
F	1	MSE	MET	MODIFIED RESIDUE	UNP P11544
F	51	MSE	MET	MODIFIED RESIDUE	UNP P11544
F	107	MSE	MET	MODIFIED RESIDUE	UNP P11544
F	169	MSE	MET	MODIFIED RESIDUE	UNP P11544
F	211	175	ALA	SEE REMARK 999	UNP P11544
F	211	175	SER	SEE REMARK 999	UNP P11544
F	211	175	GLY	SEE REMARK 999	UNP P11544

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	250	MSE	MET	MODIFIED RESIDUE	UNP P11544
F	278	MSE	MET	MODIFIED RESIDUE	UNP P11544
F	288	MSE	MET	MODIFIED RESIDUE	UNP P11544
F	299	MSE	MET	MODIFIED RESIDUE	UNP P11544
F	304	MSE	MET	MODIFIED RESIDUE	UNP P11544
F	422	MSE	MET	MODIFIED RESIDUE	UNP P11544
F	443	MSE	MET	MODIFIED RESIDUE	UNP P11544
F	448	MSE	MET	MODIFIED RESIDUE	UNP P11544
F	497	MSE	MET	MODIFIED RESIDUE	UNP P11544
F	564	MSE	MET	MODIFIED RESIDUE	UNP P11544
F	714	MSE	MET	MODIFIED RESIDUE	UNP P11544
G	1	MSE	MET	MODIFIED RESIDUE	UNP P11544
G	51	MSE	MET	MODIFIED RESIDUE	UNP P11544
G	107	MSE	MET	MODIFIED RESIDUE	UNP P11544
G	169	MSE	MET	MODIFIED RESIDUE	UNP P11544
G	211	175	ALA	SEE REMARK 999	UNP P11544
G	211	175	SER	SEE REMARK 999	UNP P11544
G	211	175	GLY	SEE REMARK 999	UNP P11544
G	250	MSE	MET	MODIFIED RESIDUE	UNP P11544
G	278	MSE	MET	MODIFIED RESIDUE	UNP P11544
G	288	MSE	MET	MODIFIED RESIDUE	UNP P11544
G	299	MSE	MET	MODIFIED RESIDUE	UNP P11544
G	304	MSE	MET	MODIFIED RESIDUE	UNP P11544
G	422	MSE	MET	MODIFIED RESIDUE	UNP P11544
G	443	MSE	MET	MODIFIED RESIDUE	UNP P11544
G	448	MSE	MET	MODIFIED RESIDUE	UNP P11544
G	497	MSE	MET	MODIFIED RESIDUE	UNP P11544
G	564	MSE	MET	MODIFIED RESIDUE	UNP P11544
G	714	MSE	MET	MODIFIED RESIDUE	UNP P11544
H	1	MSE	MET	MODIFIED RESIDUE	UNP P11544
H	51	MSE	MET	MODIFIED RESIDUE	UNP P11544
H	107	MSE	MET	MODIFIED RESIDUE	UNP P11544
H	169	MSE	MET	MODIFIED RESIDUE	UNP P11544
H	211	175	ALA	SEE REMARK 999	UNP P11544
H	211	175	SER	SEE REMARK 999	UNP P11544
H	211	175	GLY	SEE REMARK 999	UNP P11544
H	250	MSE	MET	MODIFIED RESIDUE	UNP P11544
H	278	MSE	MET	MODIFIED RESIDUE	UNP P11544
H	288	MSE	MET	MODIFIED RESIDUE	UNP P11544
H	299	MSE	MET	MODIFIED RESIDUE	UNP P11544
H	304	MSE	MET	MODIFIED RESIDUE	UNP P11544
H	422	MSE	MET	MODIFIED RESIDUE	UNP P11544

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	443	MSE	MET	MODIFIED RESIDUE	UNP P11544
H	448	MSE	MET	MODIFIED RESIDUE	UNP P11544
H	497	MSE	MET	MODIFIED RESIDUE	UNP P11544
H	564	MSE	MET	MODIFIED RESIDUE	UNP P11544
H	714	MSE	MET	MODIFIED RESIDUE	UNP P11544

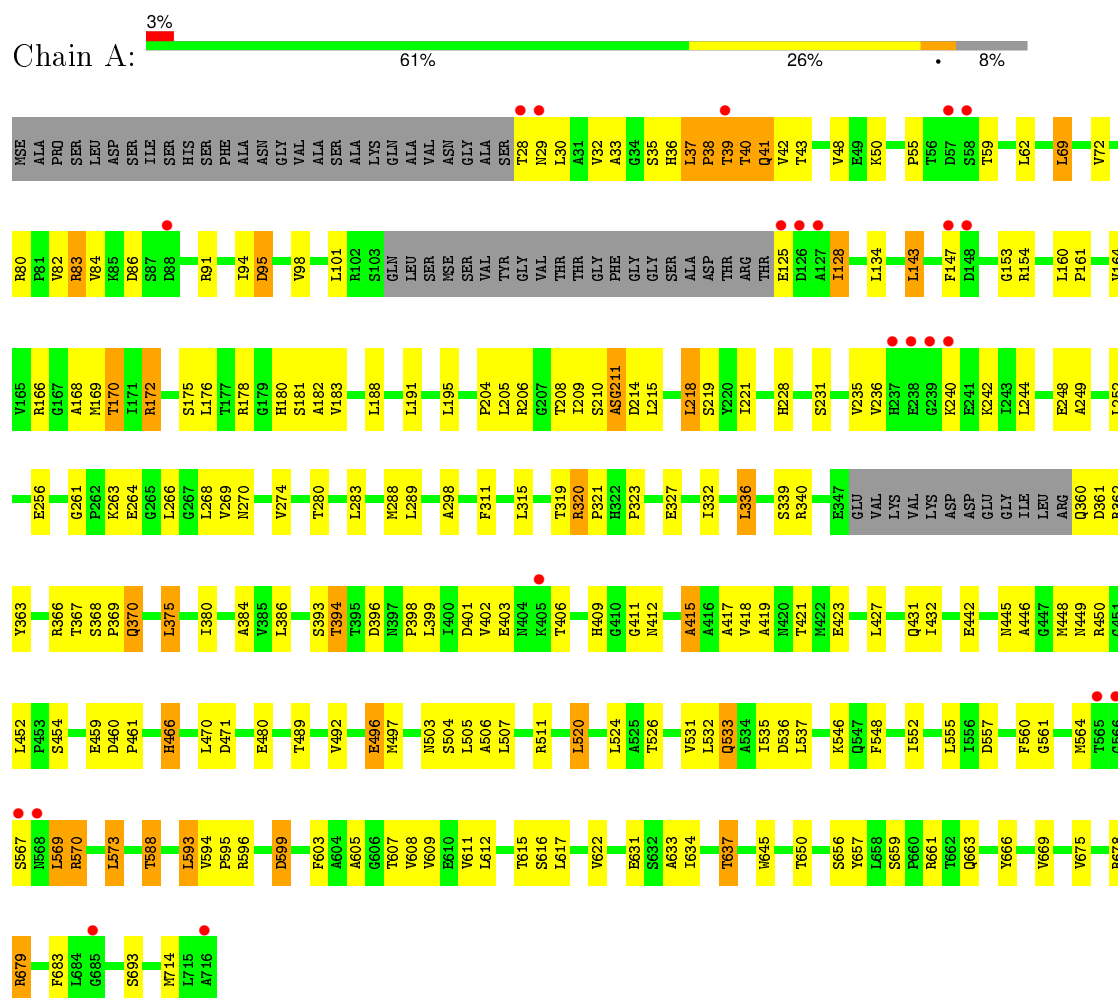
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	84	Total O 84 84	0	0
2	B	79	Total O 79 79	0	0
2	C	72	Total O 72 72	0	0
2	D	101	Total O 101 101	0	0
2	E	64	Total O 64 64	0	0
2	F	83	Total O 83 83	0	0
2	G	96	Total O 96 96	0	0
2	H	72	Total O 72 72	0	0

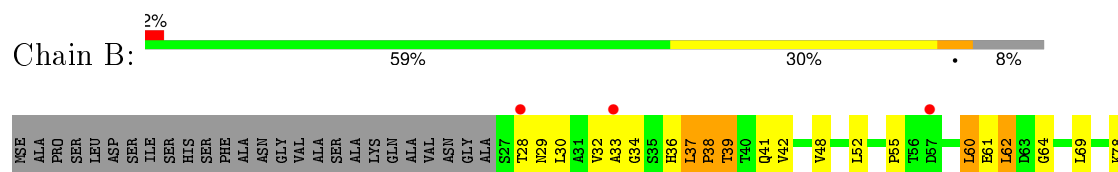
### 3 Residue-property plots

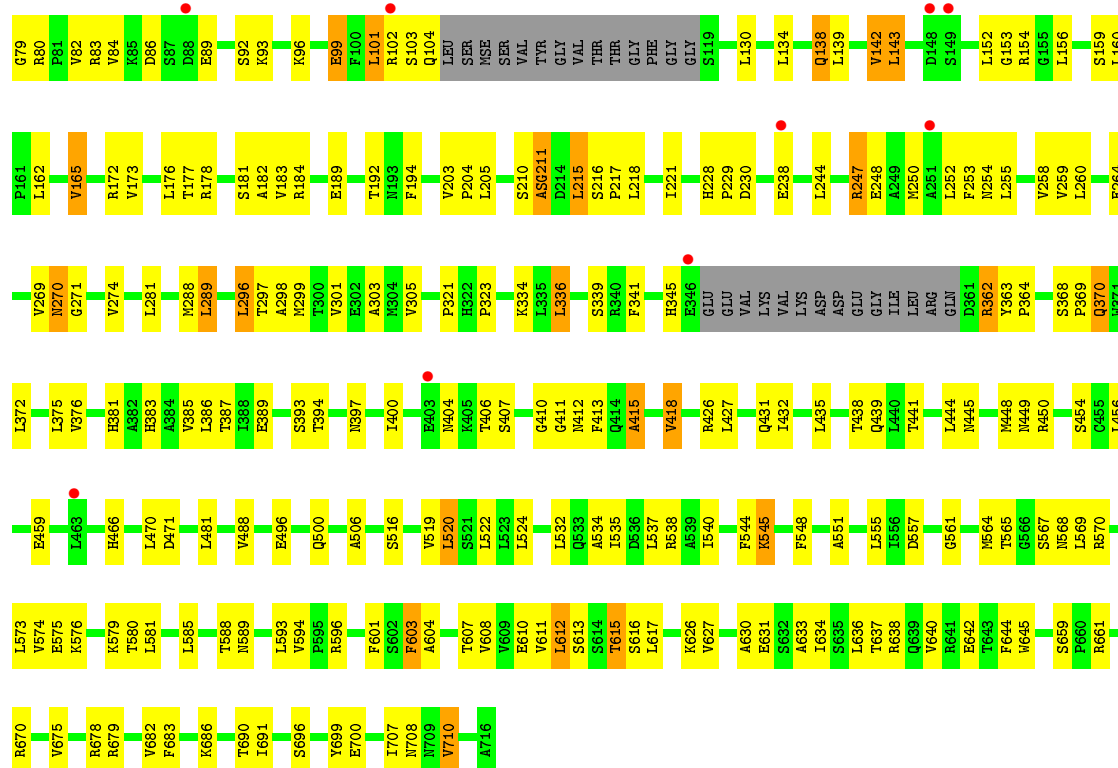
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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 ( $\text{RSRZ} > 2$ ). 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: phenylalanine ammonia-lyase

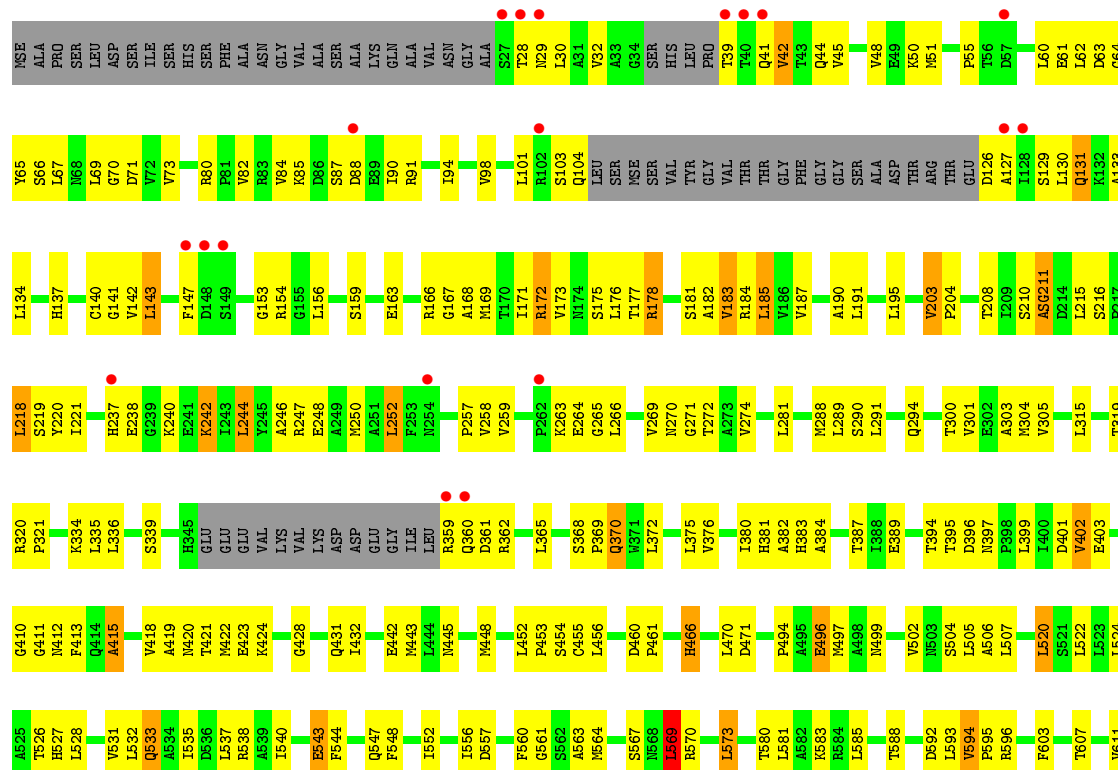


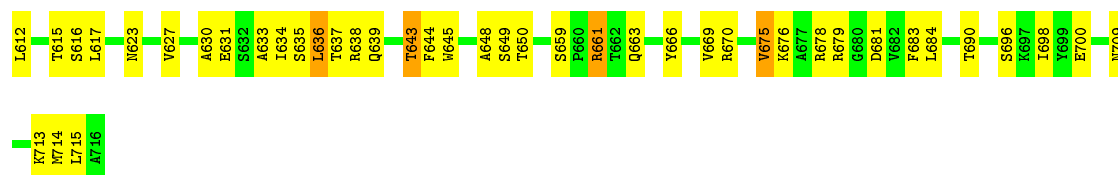
- Molecule 1: phenylalanine ammonia-lyase



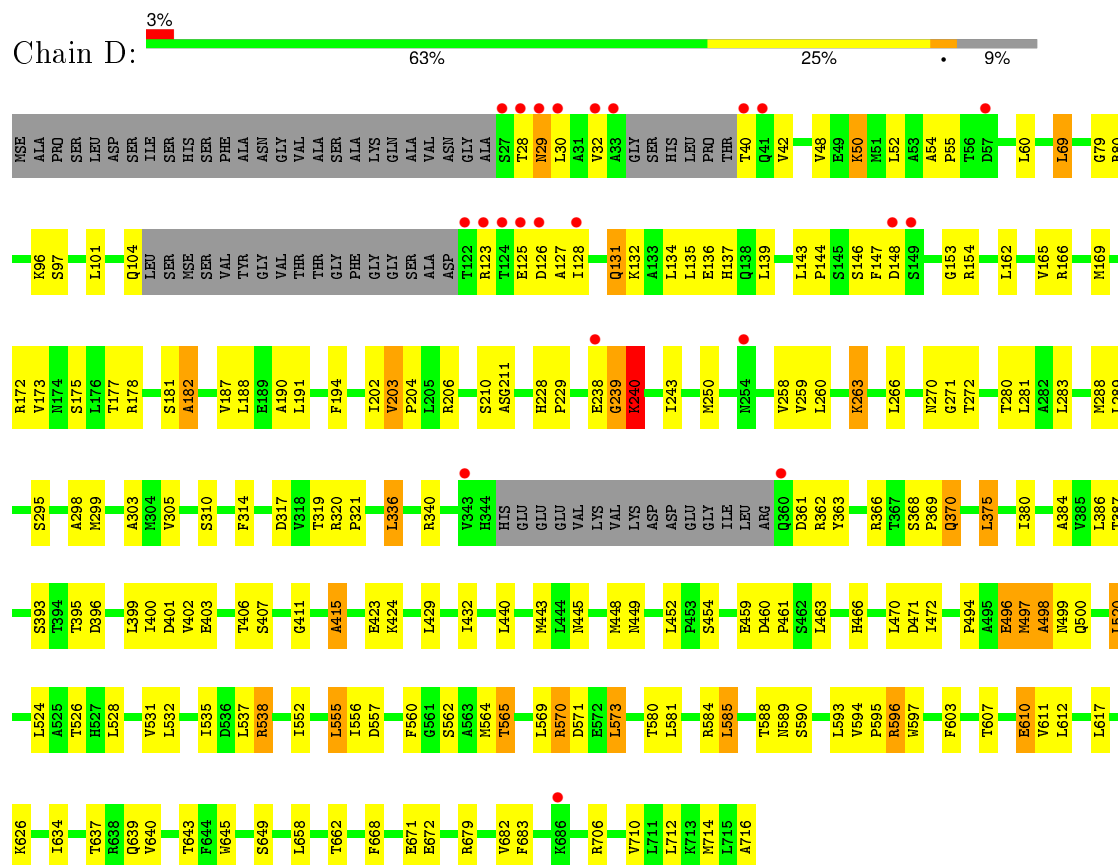


• Molecule 1: phenylalanine ammonia-lyase

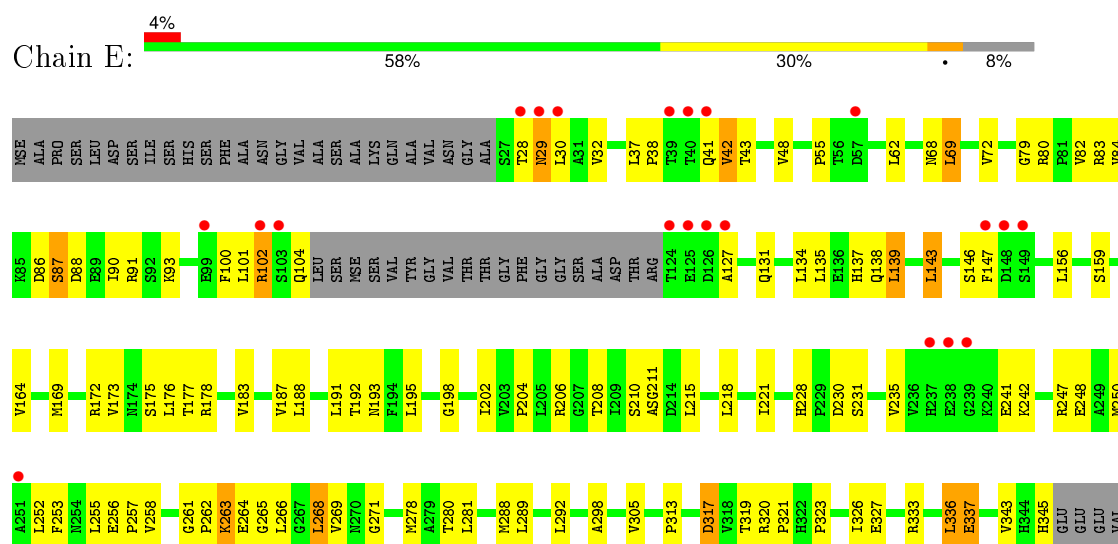


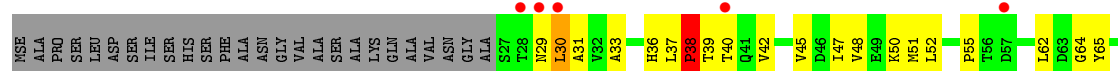


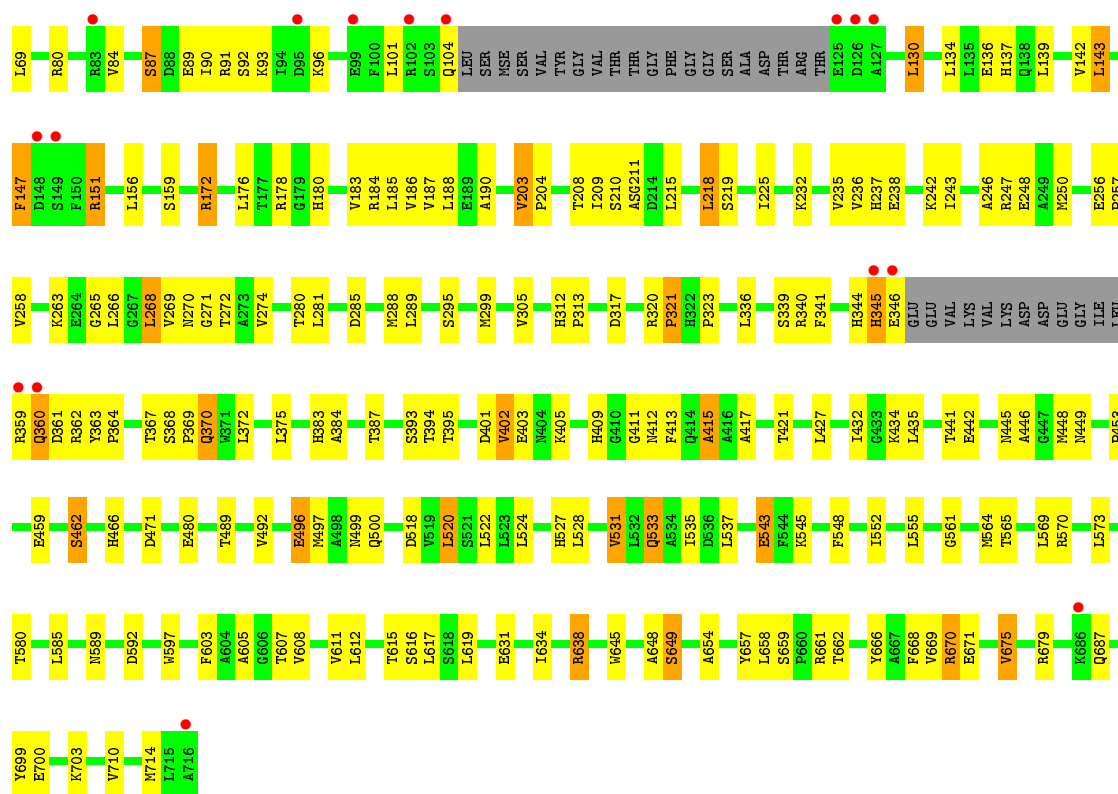
- Molecule 1: phenylalanine ammonia-lyase



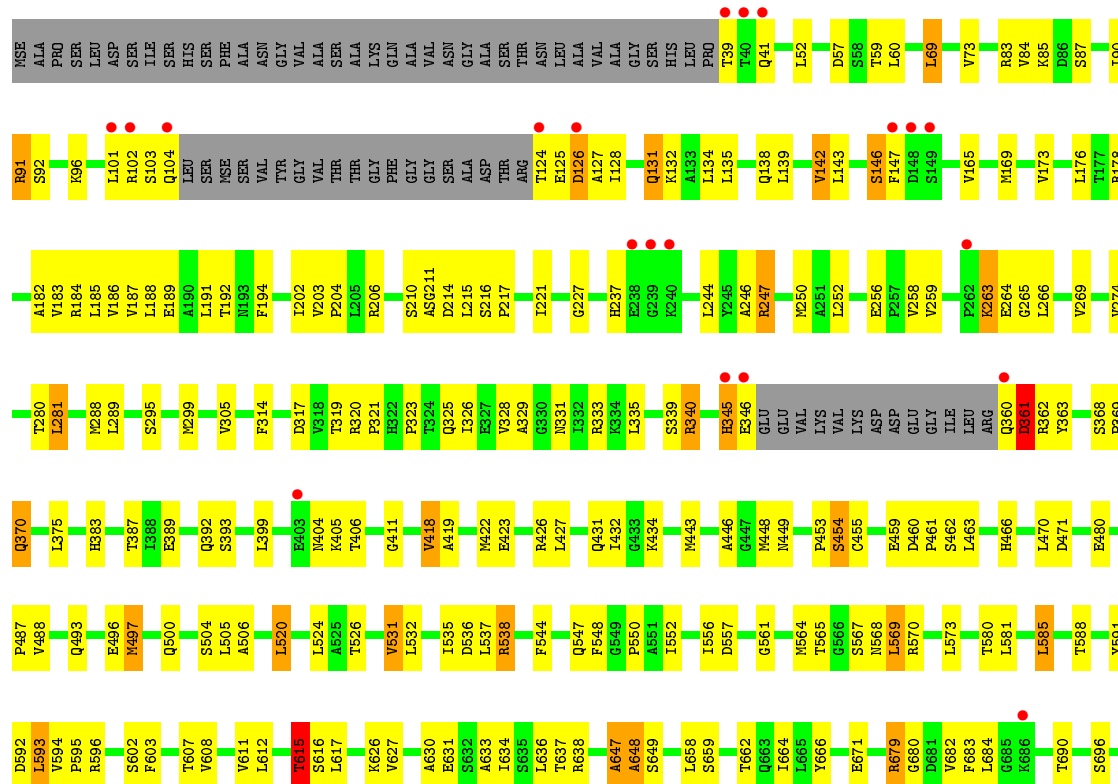
- Molecule 1: phenylalanine ammonia-lyase







• Molecule 1: phenylalanine ammonia-lyase



E700	
I707	
N708	
N709	
V710	
L711	
L712	
K713	
A716	

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	105.19Å 180.97Å 149.22Å 90.00° 96.95° 90.00°	Depositor
Resolution (Å)	25.10 – 2.70 25.12 – 2.70	Depositor EDS
% Data completeness (in resolution range)	(Not available) (25.10-2.70) 96.7 (25.12-2.70)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.64 (at 2.72Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.194 , 0.253 0.195 , 0.251	Depositor DCC
$R_{free}$ test set	3693 reflections (2.51%)	DCC
Wilson B-factor (Å <sup>2</sup> )	36.3	Xtriage
Anisotropy	0.796	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 54.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	3 of 146898 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	40380	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.05% 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

Bond lengths and bond angles in the following residue types are not validated in this section:  
175

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	0.37	0/5036	0.66	0/6815
1	B	0.37	0/5077	0.67	0/6871
1	C	0.37	0/5001	0.66	1/6764 (0.0%)
1	D	0.38	0/5002	0.68	0/6766
1	E	0.37	0/5040	0.66	0/6821
1	F	0.38	0/5006	0.67	1/6771 (0.0%)
1	G	0.37	0/5053	0.68	0/6837
1	H	0.37	0/4966	0.67	0/6718
All	All	0.37	0/40181	0.67	2/54363 (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	A	2	1
1	B	2	1
1	C	2	1
1	D	2	1
1	E	2	1
1	F	2	1
1	G	2	1
1	H	2	1
All	All	16	8

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	215	LEU	CA-CB-CG	5.35	127.60	115.30
1	C	569	LEU	CA-CB-CG	5.27	127.41	115.30

5 of 16 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	211	175	C3,C2
1	B	211	175	C3,C2
1	C	211	175	C3,C2
1	D	211	175	C3,C2
1	E	211	175	C3,C2

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	210	SER	Peptide
1	B	210	SER	Peptide
1	C	210	SER	Peptide
1	D	210	SER	Peptide
1	E	210	SER	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	4978	0	5040	201	0
1	B	5019	0	5081	203	0
1	C	4946	0	5018	261	0
1	D	4948	0	5020	158	0
1	E	4982	0	5048	192	0
1	F	4951	0	5020	199	0
1	G	4995	0	5060	205	0
1	H	4910	0	4974	188	0
2	A	84	0	0	4	0
2	B	79	0	0	3	0
2	C	72	0	0	5	0
2	D	101	0	0	4	0
2	E	64	0	0	5	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	F	83	0	0	2	0
2	G	96	0	0	8	0
2	H	72	0	0	4	0
All	All	40380	0	40261	1496	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 1496 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:F:594:VAL:HG23	1:F:595:PRO:HD3	1.26	1.09
1:B:247:ARG:HB3	1:B:247:ARG:HH11	1.13	1.08
1:A:244:LEU:HD11	1:A:249:ALA:HB2	1.31	1.07
1:F:334:LYS:HD3	1:F:661:ARG:HH21	1.21	1.04
1:H:615:THR:HG23	1:H:617:LEU:H	1.24	1.02

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	647/714 (91%)	593 (92%)	46 (7%)	8 (1%)	16	39
1	B	653/714 (92%)	598 (92%)	41 (6%)	14 (2%)	9	23
1	C	641/714 (90%)	587 (92%)	45 (7%)	9 (1%)	14	35
1	D	641/714 (90%)	591 (92%)	38 (6%)	12 (2%)	10	25
1	E	648/714 (91%)	594 (92%)	43 (7%)	11 (2%)	11	29
1	F	641/714 (90%)	588 (92%)	41 (6%)	12 (2%)	10	25
1	G	649/714 (91%)	604 (93%)	34 (5%)	11 (2%)	11	29

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	H	637/714 (89%)	590 (93%)	34 (5%)	13 (2%)	9	24
All	All	5157/5712 (90%)	4745 (92%)	322 (6%)	90 (2%)	11	29

5 of 90 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	39	THR
1	A	454	SER
1	A	466	HIS
1	B	39	THR
1	B	260	LEU

### 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	542/573 (95%)	504 (93%)	38 (7%)	19	42
1	B	547/573 (96%)	511 (93%)	36 (7%)	21	45
1	C	538/573 (94%)	505 (94%)	33 (6%)	23	49
1	D	539/573 (94%)	496 (92%)	43 (8%)	15	33
1	E	543/573 (95%)	502 (92%)	41 (8%)	16	37
1	F	539/573 (94%)	505 (94%)	34 (6%)	22	48
1	G	544/573 (95%)	505 (93%)	39 (7%)	18	41
1	H	535/573 (93%)	500 (94%)	35 (6%)	21	46
All	All	4327/4584 (94%)	4028 (93%)	299 (7%)	19	43

5 of 299 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	537	LEU
1	E	336	LEU
1	H	256	GLU
1	D	565	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	62	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 65 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	29	ASN
1	E	527	HIS
1	H	370	GLN
1	E	137	HIS
1	E	412	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

8 non-standard protein/DNA/RNA residues 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
1	175	A	211	1	11,14,15	1.39	2 (18%)	6,19,21	3.85	3 (50%)
1	175	B	211	1	11,14,15	1.34	2 (18%)	6,19,21	4.40	3 (50%)
1	175	C	211	1	11,14,15	1.28	1 (9%)	6,19,21	3.83	3 (50%)
1	175	D	211	1	11,14,15	1.58	2 (18%)	6,19,21	3.88	1 (16%)
1	175	E	211	1	11,14,15	1.21	1 (9%)	6,19,21	3.83	3 (50%)
1	175	F	211	1	11,14,15	1.40	1 (9%)	6,19,21	3.84	3 (50%)
1	175	G	211	1	11,14,15	1.34	1 (9%)	6,19,21	3.91	3 (50%)
1	175	H	211	1	11,14,15	1.22	1 (9%)	6,19,21	3.71	3 (50%)

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	175	A	211	1	2/2/5/7	0/3/25/26	0/1/1/1
1	175	B	211	1	2/2/5/7	0/3/25/26	0/1/1/1
1	175	C	211	1	2/2/5/7	0/3/25/26	0/1/1/1
1	175	D	211	1	2/2/5/7	0/3/25/26	0/1/1/1
1	175	E	211	1	2/2/5/7	0/3/25/26	0/1/1/1
1	175	F	211	1	2/2/5/7	0/3/25/26	0/1/1/1
1	175	G	211	1	2/2/5/7	0/3/25/26	0/1/1/1
1	175	H	211	1	2/2/5/7	0/3/25/26	0/1/1/1

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	211	175	C3-C5	-3.79	1.46	1.52
1	F	211	175	C3-C5	-3.18	1.47	1.52
1	G	211	175	C3-C5	-2.72	1.47	1.52
1	B	211	175	C3-C5	-2.67	1.48	1.52
1	A	211	175	C3-C5	-2.60	1.48	1.52

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	211	175	O6-C5-N0	-2.59	122.85	125.75
1	F	211	175	O6-C5-N0	-2.36	123.10	125.75
1	A	211	175	O6-C5-N0	-2.35	123.11	125.75
1	E	211	175	O6-C5-N0	-2.22	123.26	125.75
1	H	211	175	O6-C5-N0	-2.21	123.28	125.75

5 of 16 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	211	175	C3
1	A	211	175	C2
1	H	211	175	C3
1	H	211	175	C2
1	G	211	175	C3

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	211	175	1	0
1	B	211	175	2	0
1	C	211	175	1	0

## 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	640/714 (89%)	-0.25	22 (3%)	49	49	12, 36, 66, 87	0
1	B	646/714 (90%)	-0.21	12 (1%)	70	70	12, 37, 64, 78	0
1	C	636/714 (89%)	-0.18	19 (2%)	54	54	15, 36, 68, 92	0
1	D	636/714 (89%)	-0.27	22 (3%)	48	48	8, 33, 63, 113	0
1	E	641/714 (89%)	-0.16	27 (4%)	40	39	13, 39, 66, 94	0
1	F	636/714 (89%)	-0.17	25 (3%)	43	43	13, 36, 64, 98	0
1	G	642/714 (89%)	-0.25	21 (3%)	50	50	14, 35, 62, 94	0
1	H	630/714 (88%)	-0.25	20 (3%)	51	51	13, 36, 62, 91	0
All	All	5107/5712 (89%)	-0.22	168 (3%)	50	50	8, 36, 65, 113	0

The worst 5 of 168 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	124	THR	6.5
1	C	148	ASP	5.8
1	D	122	THR	5.7
1	D	123	ARG	5.7
1	D	148	ASP	5.5

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

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
1	175	A	211	14/15	0.97	0.10	-	27,32,34,35	0
1	175	H	211	14/15	0.95	0.13	-	24,32,35,35	0
1	175	G	211	14/15	0.95	0.13	-	25,31,35,36	0
1	175	C	211	14/15	0.95	0.14	-	30,35,36,37	0
1	175	E	211	14/15	0.94	0.11	-	30,37,38,38	0
1	175	D	211	14/15	0.93	0.15	-	31,35,37,37	0
1	175	F	211	14/15	0.96	0.11	-	25,31,34,36	0
1	175	B	211	14/15	0.96	0.12	-	23,29,32,32	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.