



# Full wwPDB X-ray Structure Validation Report ⓘ

May 22, 2020 – 07:23 pm BST

PDB ID : 4C5S  
Title : Structural Investigations into the Stereochemistry and Activity of a Phenylalanine-2,3-Aminomutase from *Taxus chinensis*  
Authors : Wybenga, G.G.; Szymanski, W.; Wu, B.; Feringa, B.L.; Janssen, D.B.; Dijkstra, B.W.  
Deposited on : 2013-09-16  
Resolution : 1.85 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

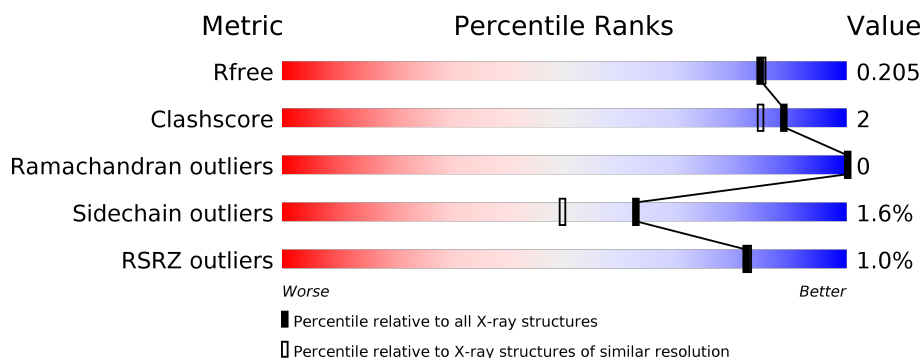
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.85 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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	705	<div> <div>%</div> <div> <div></div> <div>84%</div> <div>6%</div> <div>9%</div> </div> </div>
1	B	705	<div> <div>%</div> <div> <div></div> <div>85%</div> <div>5%</div> <div>9%</div> </div> </div>
1	C	705	<div> <div>%</div> <div> <div></div> <div>84%</div> <div>7%</div> <div>9%</div> </div> </div>
1	D	705	<div> <div>%</div> <div> <div></div> <div>86%</div> <div>5%</div> <div>9%</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 20797 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	639	Total	C	N	O	S	0	0	0
			4955	3143	854	935	23			
1	B	641	Total	C	N	O	S	0	0	0
			4971	3153	857	938	23			
1	C	642	Total	C	N	O	S	0	0	0
			4979	3157	857	942	23			
1	D	641	Total	C	N	O	S	0	0	0
			4971	3153	857	938	23			

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	expression tag	UNP Q68G84
A	-18	GLY	-	expression tag	UNP Q68G84
A	-17	SER	-	expression tag	UNP Q68G84
A	-16	SER	-	expression tag	UNP Q68G84
A	-15	HIS	-	expression tag	UNP Q68G84
A	-14	HIS	-	expression tag	UNP Q68G84
A	-13	HIS	-	expression tag	UNP Q68G84
A	-12	HIS	-	expression tag	UNP Q68G84
A	-11	HIS	-	expression tag	UNP Q68G84
A	-10	HIS	-	expression tag	UNP Q68G84
A	-9	SER	-	expression tag	UNP Q68G84
A	-8	SER	-	expression tag	UNP Q68G84
A	-7	GLY	-	expression tag	UNP Q68G84
A	-6	LEU	-	expression tag	UNP Q68G84
A	-5	VAL	-	expression tag	UNP Q68G84
A	-4	PRO	-	expression tag	UNP Q68G84
A	-3	ARG	-	expression tag	UNP Q68G84
A	-2	GLY	-	expression tag	UNP Q68G84
A	-1	SER	-	expression tag	UNP Q68G84
A	0	HIS	-	expression tag	UNP Q68G84
A	175	MDO	ALA	chromophore	UNP Q68G84

*Continued on next page...*

*Continued from previous page...*

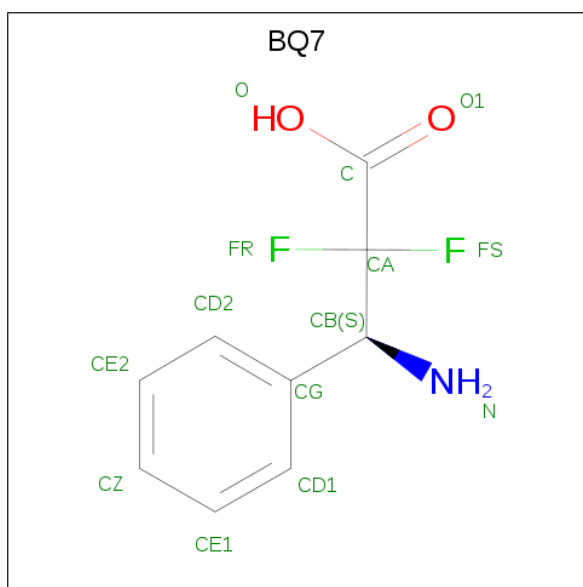
Chain	Residue	Modelled	Actual	Comment	Reference
A	175	MDO	SER	chromophore	UNP Q68G84
A	175	MDO	GLY	chromophore	UNP Q68G84
A	80	ALA	TYR	engineered mutation	UNP P42212
B	-19	MET	-	expression tag	UNP Q68G84
B	-18	GLY	-	expression tag	UNP Q68G84
B	-17	SER	-	expression tag	UNP Q68G84
B	-16	SER	-	expression tag	UNP Q68G84
B	-15	HIS	-	expression tag	UNP Q68G84
B	-14	HIS	-	expression tag	UNP Q68G84
B	-13	HIS	-	expression tag	UNP Q68G84
B	-12	HIS	-	expression tag	UNP Q68G84
B	-11	HIS	-	expression tag	UNP Q68G84
B	-10	HIS	-	expression tag	UNP Q68G84
B	-9	SER	-	expression tag	UNP Q68G84
B	-8	SER	-	expression tag	UNP Q68G84
B	-7	GLY	-	expression tag	UNP Q68G84
B	-6	LEU	-	expression tag	UNP Q68G84
B	-5	VAL	-	expression tag	UNP Q68G84
B	-4	PRO	-	expression tag	UNP Q68G84
B	-3	ARG	-	expression tag	UNP Q68G84
B	-2	GLY	-	expression tag	UNP Q68G84
B	-1	SER	-	expression tag	UNP Q68G84
B	0	HIS	-	expression tag	UNP Q68G84
B	175	MDO	ALA	chromophore	UNP Q68G84
B	175	MDO	SER	chromophore	UNP Q68G84
B	175	MDO	GLY	chromophore	UNP Q68G84
B	80	ALA	TYR	engineered mutation	UNP P42212
C	-19	MET	-	expression tag	UNP Q68G84
C	-18	GLY	-	expression tag	UNP Q68G84
C	-17	SER	-	expression tag	UNP Q68G84
C	-16	SER	-	expression tag	UNP Q68G84
C	-15	HIS	-	expression tag	UNP Q68G84
C	-14	HIS	-	expression tag	UNP Q68G84
C	-13	HIS	-	expression tag	UNP Q68G84
C	-12	HIS	-	expression tag	UNP Q68G84
C	-11	HIS	-	expression tag	UNP Q68G84
C	-10	HIS	-	expression tag	UNP Q68G84
C	-9	SER	-	expression tag	UNP Q68G84
C	-8	SER	-	expression tag	UNP Q68G84
C	-7	GLY	-	expression tag	UNP Q68G84
C	-6	LEU	-	expression tag	UNP Q68G84
C	-5	VAL	-	expression tag	UNP Q68G84

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	-4	PRO	-	expression tag	UNP Q68G84
C	-3	ARG	-	expression tag	UNP Q68G84
C	-2	GLY	-	expression tag	UNP Q68G84
C	-1	SER	-	expression tag	UNP Q68G84
C	0	HIS	-	expression tag	UNP Q68G84
C	175	MDO	ALA	chromophore	UNP Q68G84
C	175	MDO	SER	chromophore	UNP Q68G84
C	175	MDO	GLY	chromophore	UNP Q68G84
C	80	ALA	TYR	engineered mutation	UNP P42212
D	-19	MET	-	expression tag	UNP Q68G84
D	-18	GLY	-	expression tag	UNP Q68G84
D	-17	SER	-	expression tag	UNP Q68G84
D	-16	SER	-	expression tag	UNP Q68G84
D	-15	HIS	-	expression tag	UNP Q68G84
D	-14	HIS	-	expression tag	UNP Q68G84
D	-13	HIS	-	expression tag	UNP Q68G84
D	-12	HIS	-	expression tag	UNP Q68G84
D	-11	HIS	-	expression tag	UNP Q68G84
D	-10	HIS	-	expression tag	UNP Q68G84
D	-9	SER	-	expression tag	UNP Q68G84
D	-8	SER	-	expression tag	UNP Q68G84
D	-7	GLY	-	expression tag	UNP Q68G84
D	-6	LEU	-	expression tag	UNP Q68G84
D	-5	VAL	-	expression tag	UNP Q68G84
D	-4	PRO	-	expression tag	UNP Q68G84
D	-3	ARG	-	expression tag	UNP Q68G84
D	-2	GLY	-	expression tag	UNP Q68G84
D	-1	SER	-	expression tag	UNP Q68G84
D	0	HIS	-	expression tag	UNP Q68G84
D	175	MDO	ALA	chromophore	UNP Q68G84
D	175	MDO	SER	chromophore	UNP Q68G84
D	175	MDO	GLY	chromophore	UNP Q68G84
D	80	ALA	TYR	engineered mutation	UNP P42212

- Molecule 2 is (3S)-3-amino-2,2-difluoro-3-phenylpropanoic acid (three-letter code: BQ7) (formula: C<sub>9</sub>H<sub>9</sub>F<sub>2</sub>NO<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	F	N	O	0	0
			14	9	2	1	2		
2	B	1	Total	C	F	N	O	0	0
			14	9	2	1	2		
2	C	1	Total	C	F	N	O	0	0
			14	9	2	1	2		
2	D	1	Total	C	F	N	O	0	0
			14	9	2	1	2		

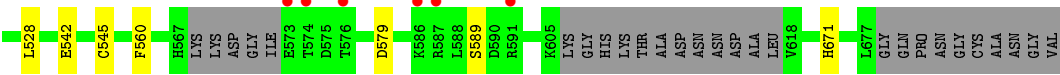
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	246	Total	O	0	0
			246	246		
3	B	226	Total	O	0	0
			226	226		
3	C	202	Total	O	0	0
			202	202		
3	D	191	Total	O	0	0
			191	191		

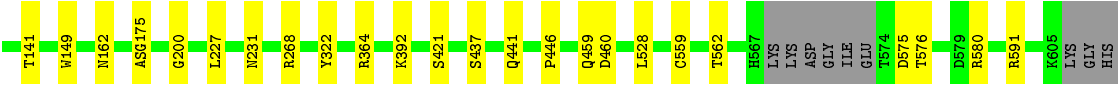
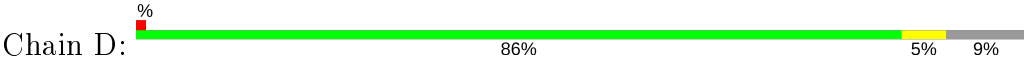


- Molecule 1: PHENYLALANINE AMMONIA-LYASE





● Molecule 1: PHENYLALANINE AMMONIA-LYASE





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	99.75Å 146.19Å 100.39Å 90.00° 99.31° 90.00°	Depositor
Resolution (Å)	99.07 – 1.85 48.48 – 1.85	Depositor EDS
% Data completeness (in resolution range)	99.9 (99.07-1.85) 99.9 (48.48-1.85)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.27 (at 1.86Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.178 , 0.205 0.178 , 0.205	Depositor DCC
$R_{free}$ test set	12073 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.3	Xtriage
Anisotropy	0.006	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 41.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.016 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	20797	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.02% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: BQ7, MDO

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	0.55	3/5026 (0.1%)	0.62	0/6820
1	B	0.54	1/5042 (0.0%)	0.63	2/6841 (0.0%)
1	C	0.54	4/5049 (0.1%)	0.62	0/6849
1	D	0.55	1/5042 (0.0%)	0.63	0/6841
All	All	0.54	9/20159 (0.0%)	0.62	2/27351 (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	0	2
1	B	0	2
1	C	0	2
1	D	0	2
All	All	0	8

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	330	TRP	CD2-CE2	5.62	1.48	1.41
1	C	330	TRP	CD2-CE2	5.54	1.48	1.41
1	A	68	TRP	CD2-CE2	5.37	1.47	1.41
1	C	294	TRP	CD2-CE2	5.35	1.47	1.41
1	C	149	TRP	CD2-CE2	5.27	1.47	1.41
1	C	68	TRP	CD2-CE2	5.25	1.47	1.41
1	A	294	TRP	CD2-CE2	5.22	1.47	1.41
1	A	149	TRP	CD2-CE2	5.01	1.47	1.41
1	D	149	TRP	CD2-CE2	5.01	1.47	1.41

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	664	ARG	NE-CZ-NH1	6.07	123.34	120.30
1	B	664	ARG	NE-CZ-NH2	-5.26	117.67	120.30

There are no chirality outliers.

All (8) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	175	MDO	Mainchain,Peptide
1	B	175	MDO	Mainchain,Peptide
1	C	175	MDO	Mainchain,Peptide
1	D	175	MDO	Mainchain,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	4955	0	5019	26	0
1	B	4971	0	5038	18	0
1	C	4979	0	5039	28	0
1	D	4971	0	5038	19	0
2	A	14	0	7	0	0
2	B	14	0	7	0	0
2	C	14	0	7	0	0
2	D	14	0	7	0	0
3	A	246	0	0	0	0
3	B	226	0	0	0	0
3	C	202	0	0	0	0
3	D	191	0	0	0	0
All	All	20797	0	20162	74	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 2.

All (74) 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:A:174:SER:C	1:A:175:MDO:H2	1.39	1.23
1:D:92:ARG:HB3	1:D:92:ARG:CZ	2.19	0.72
1:A:47:HIS:CE1	1:A:162:ASN:HD21	2.07	0.71
1:A:174:SER:C	1:A:175:MDO:CA	2.60	0.68
1:A:98:SER:O	1:A:102:GLU:HG2	1.97	0.65
1:A:446:PRO:HD3	1:B:446:PRO:HD3	1.83	0.60
1:A:27:ILE:CD1	1:A:43:LEU:HB2	2.34	0.58
1:A:422:VAL:HG11	1:C:99:GLU:HG3	1.87	0.56
1:B:422:VAL:HG11	1:D:99:GLU:HG3	1.86	0.55
1:B:662:ASP:OD1	1:B:664:ARG:HD3	2.06	0.55
1:B:12:LYS:HE3	1:B:677:LEU:HB2	1.87	0.54
1:A:92:ARG:HH22	1:C:542:GLU:CD	2.11	0.54
1:A:29:VAL:HG12	1:A:141:THR:HG21	1.89	0.53
1:C:446:PRO:HD3	1:D:446:PRO:HD3	1.91	0.53
1:B:98:SER:O	1:B:102:GLU:HG2	2.08	0.53
1:C:114:THR:OG1	1:C:122:ASP:HB2	2.09	0.53
1:D:437:SER:O	1:D:441:GLN:HG2	2.10	0.52
1:A:99:GLU:HG3	1:C:422:VAL:HG11	1.92	0.51
1:B:12:LYS:HE3	1:B:677:LEU:CB	2.41	0.50
1:A:437:SER:O	1:A:441:GLN:HG2	2.11	0.50
1:D:12:LYS:HE3	1:D:677:LEU:HB2	1.93	0.49
1:C:88:ALA:HB2	1:C:458:ASN:HB2	1.93	0.49
1:A:174:SER:CA	1:A:175:MDO:N	2.71	0.49
1:B:94:THR:HB	1:D:421:SER:HB2	1.95	0.49
1:A:94:THR:HB	1:C:421:SER:HB2	1.94	0.48
1:B:162:ASN:ND2	1:B:200:GLY:HA2	2.29	0.47
1:B:567:HIS:NE2	1:C:579:ASP:OD1	2.39	0.47
1:B:163:VAL:HG22	1:B:199:ILE:HG12	1.97	0.47
1:A:174:SER:CA	1:A:175:MDO:H2	2.23	0.47
1:D:227:LEU:O	1:D:231:ASN:HB2	2.15	0.47
1:C:255:GLU:HG2	1:C:331:LEU:HD13	1.96	0.47
1:C:51:VAL:HG21	1:C:159:LEU:HD13	1.97	0.47
1:C:437:SER:O	1:C:441:GLN:HG2	2.15	0.46
1:A:227:LEU:O	1:A:231:ASN:HB2	2.14	0.46
1:D:559:CYS:O	1:D:562:THR:HG22	2.16	0.46
1:B:27:ILE:CD1	1:B:43:LEU:HB2	2.46	0.46
1:D:576:THR:O	1:D:580:ARG:HG3	2.16	0.46
1:A:92:ARG:NH2	1:C:542:GLU:OE1	2.49	0.46
1:C:55:ALA:O	1:C:56:GLU:C	2.54	0.46
1:C:459:GLN:HG2	1:D:322:TYR:CE2	2.51	0.45
1:A:545:CYS:SG	1:A:589:SER:HA	2.57	0.45
1:C:545:CYS:SG	1:C:589:SER:HA	2.57	0.45

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:322:TYR:CZ	1:D:459:GLN:HG2	2.52	0.45
1:D:58:CYS:HB2	1:D:61:ARG:NH1	2.31	0.45
1:B:131:SER:HB2	1:B:238:ALA:HB1	1.98	0.45
1:C:388:ALA:HA	1:C:440:LEU:HG	1.99	0.44
1:C:459:GLN:HG2	1:D:322:TYR:CZ	2.52	0.44
1:D:92:ARG:CB	1:D:92:ARG:CZ	2.94	0.44
1:A:59:ARG:HD2	1:A:63:GLU:CD	2.38	0.44
1:B:51:VAL:HG21	1:B:159:LEU:HD13	2.00	0.44
1:C:29:VAL:HG12	1:C:141:THR:HG21	2.00	0.43
1:C:174:SER:HB2	1:C:465:ALA:H	1.83	0.43
1:C:227:LEU:O	1:C:231:ASN:HB2	2.18	0.43
1:C:322:TYR:CE2	1:D:459:GLN:HG2	2.54	0.43
1:A:576:THR:O	1:A:580:ARG:HG3	2.19	0.42
1:A:25:LYS:HA	1:A:25:LYS:HE2	2.01	0.42
1:A:174:SER:HB2	1:A:465:ALA:H	1.83	0.42
1:B:61:ARG:NH1	1:B:141:THR:O	2.53	0.42
1:A:563:ILE:O	1:A:567:HIS:HB2	2.20	0.42
1:B:437:SER:O	1:B:441:GLN:HG2	2.19	0.42
1:C:27:ILE:CD1	1:C:43:LEU:HB2	2.49	0.42
1:D:162:ASN:ND2	1:D:200:GLY:HA2	2.34	0.42
1:A:419:ASP:OD2	1:C:96:ARG:NH1	2.53	0.41
1:D:29:VAL:HG12	1:D:141:THR:HG21	2.02	0.41
1:C:509:ASN:O	1:C:513:THR:HG23	2.20	0.41
1:A:178:ASP:HB3	1:A:181:PRO:HG2	2.01	0.41
1:B:299:SER:HA	1:B:300:PRO:HD3	1.95	0.41
1:D:98:SER:O	1:D:102:GLU:HG2	2.21	0.41
1:C:400:GLU:OE2	1:D:460:ASP:OD2	2.38	0.41
1:B:387:VAL:HG12	1:B:440:LEU:HD11	2.02	0.40
1:A:255:GLU:HG2	1:A:331:LEU:HD13	2.04	0.40
1:A:207:ALA:N	1:A:208:PRO:HD2	2.37	0.40
1:B:560:PHE:HB2	1:C:560:PHE:HB2	2.04	0.40
1:C:387:VAL:HG12	1:C:440:LEU:HD11	2.03	0.40

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	630/705 (89%)	618 (98%)	12 (2%)	0	100	100
1	B	632/705 (90%)	621 (98%)	11 (2%)	0	100	100
1	C	631/705 (90%)	620 (98%)	11 (2%)	0	100	100
1	D	632/705 (90%)	618 (98%)	14 (2%)	0	100	100
All	All	2525/2820 (90%)	2477 (98%)	48 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	540/591 (91%)	532 (98%)	8 (2%)	65	53
1	B	542/591 (92%)	534 (98%)	8 (2%)	65	53
1	C	543/591 (92%)	534 (98%)	9 (2%)	60	47
1	D	542/591 (92%)	533 (98%)	9 (2%)	60	47
All	All	2167/2364 (92%)	2133 (98%)	34 (2%)	62	49

All (34) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	59	ARG
1	A	102	GLU
1	A	107	CYS
1	A	268	ARG
1	A	353	ASN
1	A	392	LYS
1	A	528	LEU
1	A	562	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	107	CYS
1	B	209	GLU
1	B	268	ARG
1	B	353	ASN
1	B	364	ARG
1	B	392	LYS
1	B	671	HIS
1	B	673	LEU
1	C	43	LEU
1	C	58	CYS
1	C	92	ARG
1	C	209	GLU
1	C	268	ARG
1	C	353	ASN
1	C	392	LYS
1	C	528	LEU
1	C	671	HIS
1	D	59	ARG
1	D	92	ARG
1	D	107	CYS
1	D	268	ARG
1	D	364	ARG
1	D	392	LYS
1	D	528	LEU
1	D	575	ASP
1	D	591	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	162	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

4 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	MDO	D	175	1,2	12,13,14	2.81	5 (41%)	15,18,20	2.27	3 (20%)
1	MDO	A	175	1,2	12,13,14	2.77	4 (33%)	15,18,20	2.28	5 (33%)
1	MDO	C	175	1,2	12,13,14	2.83	6 (50%)	15,18,20	2.82	3 (20%)
1	MDO	B	175	1,2	12,13,14	2.67	4 (33%)	15,18,20	2.78	3 (20%)

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	MDO	D	175	1,2	-	2/4/23/24	0/1/1/1
1	MDO	A	175	1,2	-	3/4/23/24	0/1/1/1
1	MDO	C	175	1,2	-	2/4/23/24	0/1/1/1
1	MDO	B	175	1,2	-	2/4/23/24	0/1/1/1

All (19) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	175	MDO	O2-C2	6.91	1.37	1.23
1	A	175	MDO	O2-C2	6.73	1.37	1.23
1	B	175	MDO	O2-C2	6.37	1.36	1.23
1	C	175	MDO	O2-C2	6.21	1.36	1.23
1	A	175	MDO	O-C	4.38	1.44	1.19
1	B	175	MDO	O-C	4.29	1.44	1.19
1	D	175	MDO	O-C	4.12	1.43	1.19
1	C	175	MDO	O-C	4.12	1.43	1.19
1	C	175	MDO	CA-C1	3.06	1.55	1.51
1	C	175	MDO	C2-N3	-3.06	1.32	1.39
1	C	175	MDO	CA2-N2	-2.96	1.33	1.39
1	A	175	MDO	C2-N3	-2.86	1.33	1.39
1	B	175	MDO	C2-N3	-2.72	1.33	1.39
1	A	175	MDO	CA2-N2	-2.69	1.34	1.39
1	B	175	MDO	CA2-N2	-2.67	1.34	1.39

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	175	MDO	C2-N3	-2.64	1.33	1.39
1	D	175	MDO	CA2-N2	-2.58	1.34	1.39
1	D	175	MDO	CA-C1	2.33	1.54	1.51
1	C	175	MDO	C1-N2	2.06	1.35	1.32

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	175	MDO	CA2-C2-N3	7.83	107.08	103.37
1	B	175	MDO	CA2-C2-N3	6.95	106.66	103.37
1	B	175	MDO	O2-C2-CA2	-6.48	127.32	130.96
1	C	175	MDO	O2-C2-CA2	-5.68	127.77	130.96
1	D	175	MDO	CA2-C2-N3	5.56	106.00	103.37
1	A	175	MDO	CA2-C2-N3	5.06	105.77	103.37
1	A	175	MDO	O2-C2-CA2	-4.79	128.27	130.96
1	D	175	MDO	O2-C2-CA2	-4.75	128.29	130.96
1	B	175	MDO	O-C-CA3	-4.11	113.99	126.39
1	C	175	MDO	O-C-CA3	-4.06	114.12	126.39
1	D	175	MDO	O-C-CA3	-3.83	114.84	126.39
1	A	175	MDO	O-C-CA3	-3.21	116.69	126.39
1	A	175	MDO	CA-C1-N3	-2.08	122.25	124.75
1	A	175	MDO	CA-C1-N2	2.08	126.71	124.05

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	175	MDO	C-CA3-N3-C2
1	D	175	MDO	N2-C1-CA-CB
1	A	175	MDO	N2-C1-CA-CB
1	C	175	MDO	N2-C1-CA-CB
1	B	175	MDO	N2-C1-CA-CB
1	D	175	MDO	N3-C1-CA-CB
1	A	175	MDO	N3-C1-CA-CB
1	C	175	MDO	N3-C1-CA-CB
1	B	175	MDO	N3-C1-CA-CB

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	175	MDO	4	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

4 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	BQ7	C	1678	1	10,14,14	2.49	3 (30%)	11,20,20	2.16	3 (27%)
2	BQ7	D	1678	1	10,14,14	2.06	3 (30%)	11,20,20	2.23	2 (18%)
2	BQ7	A	1678	1	10,14,14	1.98	3 (30%)	11,20,20	2.32	2 (18%)
2	BQ7	B	1678	1	10,14,14	2.30	3 (30%)	11,20,20	1.87	2 (18%)

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	BQ7	C	1678	1	-	2/9/16/16	0/1/1/1
2	BQ7	D	1678	1	-	1/9/16/16	0/1/1/1
2	BQ7	A	1678	1	-	1/9/16/16	0/1/1/1
2	BQ7	B	1678	1	-	2/9/16/16	0/1/1/1

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1678	BQ7	FR-CA	-4.96	1.31	1.37
2	D	1678	BQ7	FR-CA	-4.78	1.31	1.37

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1678	BQ7	FR-CA	-4.74	1.31	1.37
2	C	1678	BQ7	FS-CA	-4.61	1.31	1.37
2	A	1678	BQ7	FR-CA	-4.01	1.32	1.37
2	B	1678	BQ7	CG-CB	3.92	1.57	1.52
2	C	1678	BQ7	CG-CB	3.73	1.56	1.52
2	A	1678	BQ7	FS-CA	-3.57	1.33	1.37
2	B	1678	BQ7	FS-CA	-3.22	1.33	1.37
2	D	1678	BQ7	CG-CB	3.00	1.56	1.52
2	D	1678	BQ7	FS-CA	-2.72	1.34	1.37
2	A	1678	BQ7	CG-CB	2.29	1.55	1.52

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	1678	BQ7	FR-CA-CB	-6.18	102.19	109.43
2	A	1678	BQ7	FR-CA-CB	-5.22	103.31	109.43
2	C	1678	BQ7	FR-CA-CB	-4.89	103.70	109.43
2	B	1678	BQ7	FR-CA-CB	-4.80	103.81	109.43
2	A	1678	BQ7	FS-CA-CB	4.71	114.95	109.43
2	C	1678	BQ7	FS-CA-CB	4.30	114.47	109.43
2	D	1678	BQ7	FS-CA-CB	2.93	112.86	109.43
2	B	1678	BQ7	FS-CA-CB	2.89	112.82	109.43
2	C	1678	BQ7	CD1-CG-CB	-2.06	117.12	120.84

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	1678	BQ7	FS-CA-CB-N
2	A	1678	BQ7	FS-CA-CB-N
2	B	1678	BQ7	FS-CA-CB-N
2	D	1678	BQ7	FS-CA-CB-N
2	C	1678	BQ7	FR-CA-CB-N
2	B	1678	BQ7	FR-CA-CB-N

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	174:SER	C	175:MDO	N	1.79
1	A	175:MDO	C	178:ASP	N	1.62

## 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	638/705 (90%)	-0.32	7 (1%) 80 81	9, 18, 35, 52	0
1	B	640/705 (90%)	-0.36	5 (0%) 86 86	9, 17, 32, 51	0
1	C	641/705 (90%)	-0.34	8 (1%) 79 79	8, 18, 34, 51	0
1	D	640/705 (90%)	-0.39	5 (0%) 86 86	8, 16, 37, 51	0
All	All	2559/2820 (90%)	-0.35	25 (0%) 82 82	8, 17, 35, 52	0

All (25) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	57	GLN	3.6
1	B	56	GLU	3.2
1	A	202	ASP	3.2
1	B	202	ASP	3.1
1	D	576	THR	3.1
1	A	523	ASP	2.9
1	A	576	THR	2.8
1	A	567	HIS	2.8
1	D	56	GLU	2.7
1	A	522	ASN	2.6
1	D	579	ASP	2.6
1	C	586	LYS	2.5
1	B	567	HIS	2.5
1	C	522	ASN	2.3
1	C	591	ARG	2.3
1	A	217	ARG	2.3
1	D	57	GLN	2.2
1	C	573	GLU	2.2
1	C	587	ARG	2.2
1	C	576	THR	2.1
1	C	574	THR	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	577	LEU	2.1
1	B	590	ASP	2.1
1	C	217	ARG	2.0
1	D	574	THR	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [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
1	MDO	B	175	13/14	0.81	0.14	18,22,23,23	0
1	MDO	A	175	13/14	0.82	0.16	19,23,25,25	0
1	MDO	D	175	13/14	0.82	0.15	19,22,23,23	0
1	MDO	C	175	13/14	0.88	0.12	19,21,22,23	0

## 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	BQ7	C	1678	14/14	0.93	0.09	18,21,25,25	0
2	BQ7	A	1678	14/14	0.94	0.09	19,20,24,24	0
2	BQ7	D	1678	14/14	0.96	0.09	17,18,21,22	0
2	BQ7	B	1678	14/14	0.96	0.08	19,20,22,23	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.