



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 04:26 am GMT

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 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : recalc28949

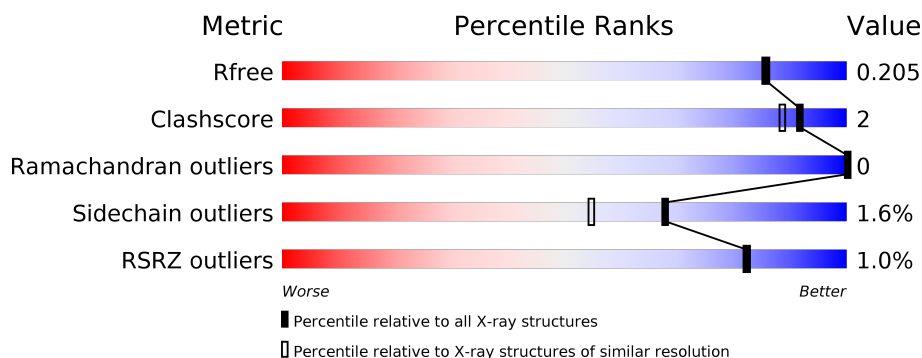
# 1 Overall quality at a glance

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))
$R_{free}$	100719	1923 (1.86-1.86)
Clashscore	112137	2083 (1.86-1.86)
Ramachandran outliers	110173	2060 (1.86-1.86)
Sidechain outliers	110143	2060 (1.86-1.86)
RSRZ outliers	101464	1932 (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. 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 style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 1%, green 84%, grey 9%);"></div> <div style="display: flex; justify-content: space-between; margin-top: 2px;"> <span>%</span> <span>84%</span> <span>6%</span> <span>9%</span> </div> </div>
1	B	705	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 1%, green 85%, grey 9%);"></div> <div style="display: flex; justify-content: space-between; margin-top: 2px;"> <span>%</span> <span>85%</span> <span>5%</span> <span>9%</span> </div> </div>
1	C	705	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 1%, green 84%, grey 9%);"></div> <div style="display: flex; justify-content: space-between; margin-top: 2px;"> <span>%</span> <span>84%</span> <span>7%</span> <span>9%</span> </div> </div>
1	D	705	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 1%, green 86%, grey 9%);"></div> <div style="display: flex; justify-content: space-between; margin-top: 2px;"> <span>%</span> <span>86%</span> <span>5%</span> <span>9%</span> </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...*

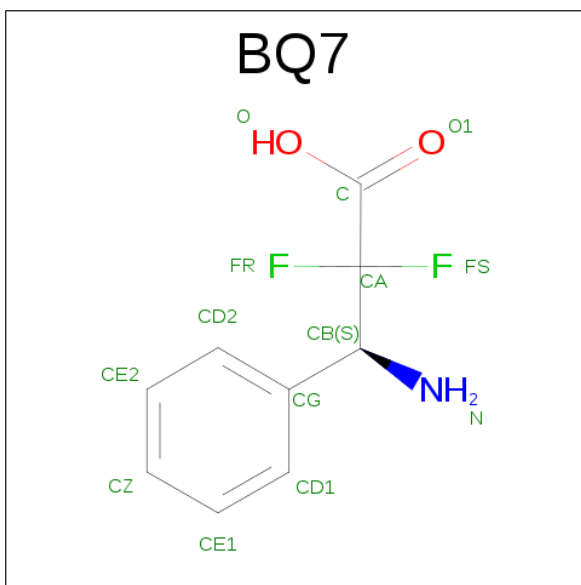
Chain	Residue	Modelled	Actual	Comment	Reference
A	175	MDO	SER	CHROMOPHORE	UNP Q68G84
A	175	MDO	GLY	CHROMOPHORE	UNP Q68G84
A	80	ALA	TYR	ENGINEERED MUTATION	UNP P42212
B	-19	MET	-	EXPRESSION TAG	UNP Q68G84
B	-18	GLY	-	EXPRESSION TAG	UNP Q68G84
B	-17	SER	-	EXPRESSION TAG	UNP Q68G84
B	-16	SER	-	EXPRESSION 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		

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

Chain A:

Amino Acid	Percentage
Met	84%
Gly	84%
Ser	84%
His	84%
His	84%
His	84%
His	84%
His	84%
His	84%
Ser	84%
Ser	84%
Ser	84%
Gly	84%
Leu	84%
Val	84%
Pro	84%
Arg	84%
Gly	84%
Ser	84%
His	84%
His	84%
Met	84%
Gly	84%
Phe	84%
Ala	84%
Val	84%
Ser	84%
Ser	84%
Arg	84%
S9	6%
K25	9%
K26	9%
I27	9%
T28	9%
V29	9%
L43	9%
H47	9%
R59	9%
E63	9%
W68	9%
R92	9%
R93	9%
T94	9%
S98	9%
E99	9%
E102	9%
C107	9%
F113	9%
Thr	9%
Lys	9%

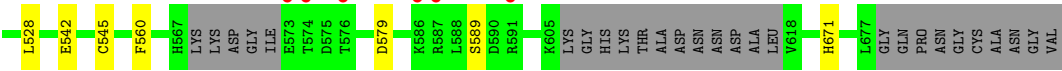
Chain B:

Sequence logo for Chain B. The y-axis represents information content in bits (0.00 to 0.25). The x-axis lists amino acids. A green bar at the top indicates a high conservation score of 85% (green), 5% (yellow), and 9% (grey).

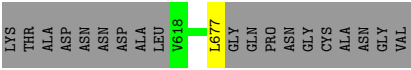
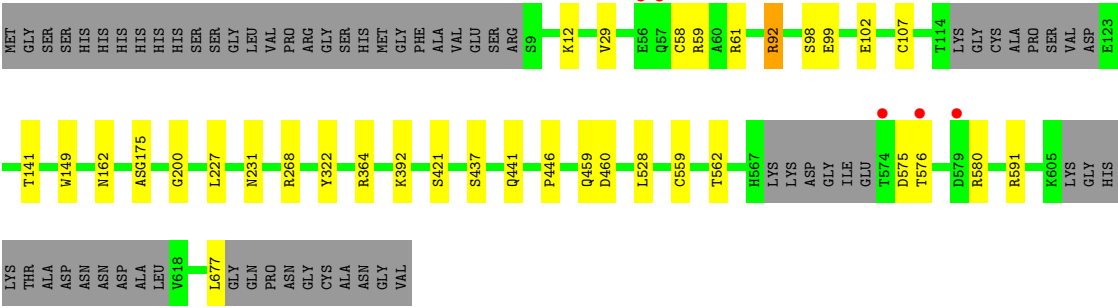
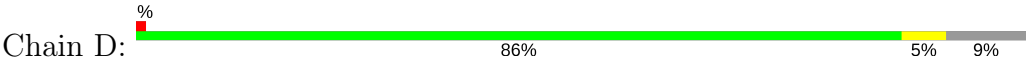
Position	Amino Acid	Information Content (bits)
1	Met	0.00
2	Gly	0.00
3	Ser	0.00
4	Ser	0.00
5	His	0.00
6	His	0.00
7	His	0.00
8	His	0.00
9	His	0.00
10	His	0.00
11	Ser	0.00
12	Ser	0.00
13	Gly	0.00
14	Leu	0.00
15	Val	0.00
16	Pro	0.00
17	Arg	0.00
18	Gly	0.00
19	Ser	0.00
20	His	0.00
21	Met	0.00
22	Gly	0.00
23	Phe	0.00
24	Ala	0.00
25	Val	0.00
26	Glu	0.00
27	Ser	0.00
28	Arg	0.00
29	Ser	0.00
30	Val	0.00

Chain C:

Amino Acid	Frequency (%)
MET	1
GLY	1
SER	1
HIS	1
HIS	1
HIS	1
HIS	1
HIS	1
HIS	1
SER	1
SER	1
GLY	1
LEU	1
VAL	1
PRO	1
ARG	1
GLY	1
SER	1
HIS	1
MET	1
GLY	1
PHE	1
ALA	1
VAL	1
GLU	1
SER	1
ARG	1
S9	1
I27	1
T28	1
V29	1
I43	1
V51	1
A55	1
E56	1
GLN	1
C58	1
W68	1
A88	1
R92	1
R96	1
E99	1
T114	1
LYS	1
GLY	1
CYS	1
ALA	1
PRO	1
SF9	1



● 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 (5.27%)	DCC
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

The worst 5 of 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

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.

5 of 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 [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	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.

The worst 5 of 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

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
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

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	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 [i](#)

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%)	70	58
1	B	542/591 (92%)	534 (98%)	8 (2%)	70	58

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	543/591 (92%)	534 (98%)	9 (2%)	66	52
1	D	542/591 (92%)	533 (98%)	9 (2%)	66	52
All	All	2167/2364 (92%)	2133 (98%)	34 (2%)	68	55

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	673	LEU
1	C	209	GLU
1	D	528	LEU
1	C	58	CYS
1	A	528	LEU

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 ⓘ

There are no RNA molecules in this entry.

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

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	A	175	1,2	12,13,14	2.68	5 (41%)	12,18,20	2.54	5 (41%)
1	MDO	B	175	1,2	12,13,14	2.58	5 (41%)	12,18,20	3.16	3 (25%)
1	MDO	C	175	1,2	12,13,14	2.75	6 (50%)	12,18,20	3.25	3 (25%)

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	2.72	5 (41%)	12,18,20	2.57	3 (25%)

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

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	175	MDO	C2-N3	-2.93	1.32	1.39
1	A	175	MDO	C2-N3	-2.74	1.33	1.39
1	C	175	MDO	CA2-N2	-2.68	1.33	1.39
1	B	175	MDO	C2-N3	-2.60	1.33	1.39
1	D	175	MDO	C2-N3	-2.52	1.33	1.39

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	175	MDO	O2-C2-CA2	-6.77	127.32	130.97
1	C	175	MDO	O2-C2-CA2	-5.94	127.77	130.97
1	A	175	MDO	O2-C2-CA2	-5.01	128.27	130.97
1	D	175	MDO	O2-C2-CA2	-4.97	128.29	130.97
1	B	175	MDO	O-C-CA3	-3.75	113.99	126.38

There are no chirality outliers.

There are no torsion outliers.

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	A	1678	1	9,14,14	1.98	3 (33%)	8,20,20	1.02	0
2	BQ7	B	1678	1	9,14,14	2.33	3 (33%)	8,20,20	0.80	0
2	BQ7	C	1678	1	9,14,14	2.51	3 (33%)	8,20,20	0.91	0
2	BQ7	D	1678	1	9,14,14	2.07	3 (33%)	8,20,20	0.87	0

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	A	1678	1	-	0/8/16/16	0/1/1/1
2	BQ7	B	1678	1	-	0/8/16/16	0/1/1/1
2	BQ7	C	1678	1	-	0/8/16/16	0/1/1/1
2	BQ7	D	1678	1	-	0/8/16/16	0/1/1/1

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1678	BQ7	FR-CA	-4.57	1.31	1.37
2	D	1678	BQ7	FR-CA	-4.40	1.31	1.37
2	C	1678	BQ7	FR-CA	-4.36	1.31	1.37
2	C	1678	BQ7	FS-CA	-4.24	1.31	1.37
2	A	1678	BQ7	FR-CA	-3.69	1.32	1.37

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	638/705 (90%)	-0.32	8 (1%) 77 78	9, 18, 35, 52	0
1	B	640/705 (90%)	-0.35	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	26 (1%) 82 82	8, 17, 35, 52	0

The worst 5 of 26 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	3.1
1	B	202	ASP	3.1
1	D	576	THR	3.1

### 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. 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(Å <sup>2</sup> )	Q<0.9
1	MDO	D	175	13/14	0.82	0.15	-	19,22,23,23	0
1	MDO	A	175	13/14	0.82	0.16	-	19,23,25,25	0
1	MDO	C	175	13/14	0.88	0.12	-	19,21,22,23	0
1	MDO	B	175	13/14	0.81	0.14	-	18,22,23,23	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. 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
2	BQ7	C	1678	14/14	0.93	0.09	0.63	18,21,25,25	0
2	BQ7	D	1678	14/14	0.96	0.09	0.30	17,18,21,22	0
2	BQ7	A	1678	14/14	0.95	0.09	0.08	19,20,24,24	0
2	BQ7	B	1678	14/14	0.96	0.08	-0.11	19,20,22,23	0

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

There are no such residues in this entry.