



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 2, 2021 – 07:29 AM EDT

PDB ID : 1GC4  
Title : THERMUS THERMOPHILUS ASPARTATE AMINOTRANSFERASE  
TETRA MUTANT 2 COMPLEXED WITH ASPARTATE  
Authors : Ura, H.; Nakai, T.; Hirotsu, K.; Kuramitsu, S.  
Deposited on : 2000-07-19  
Resolution : 3.30 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 : 2.23.2  
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) : 2.23.2

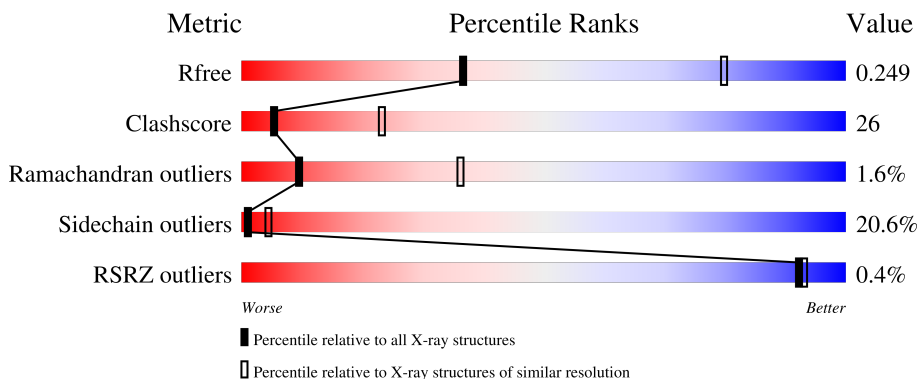
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

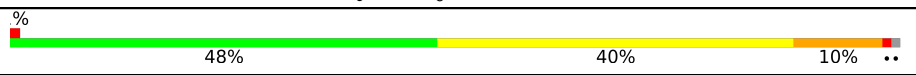



The reported resolution of this entry is 3.30 Å.

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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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 of 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	385	 48% 40% 10% ..
1	B	385	 49% 40% 10% ..
1	C	385	 47% 40% 11% .
1	D	385	 47% 41% 10% ..

## 2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 11888 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.

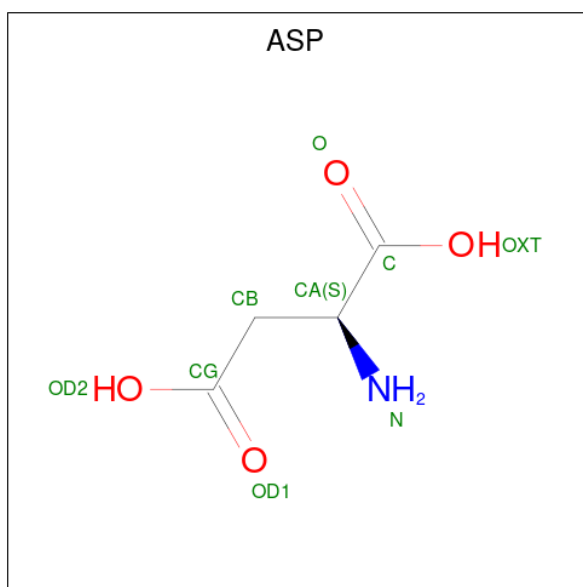
- Molecule 1 is a protein called ASPARTATE AMINOTRANSFERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	382	2948	1866	523	551	8	0	0	0
1	B	382	2948	1866	523	551	8	0	0	0
1	C	382	2948	1866	523	551	8	0	0	0
1	D	382	2948	1866	523	551	8	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

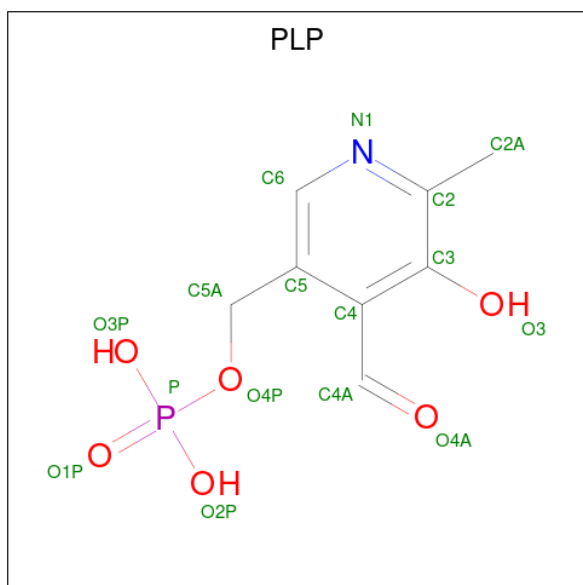
Chain	Residue	Modelled	Actual	Comment	Reference
A	14	ASP	SER	engineered mutation	UNP Q56232
A	16	VAL	THR	engineered mutation	UNP Q56232
A	101	SER	LYS	engineered mutation	UNP Q56232
A	261	ARG	SER	engineered mutation	UNP Q56232
B	514	ASP	SER	engineered mutation	UNP Q56232
B	516	VAL	THR	engineered mutation	UNP Q56232
B	601	SER	LYS	engineered mutation	UNP Q56232
B	761	ARG	SER	engineered mutation	UNP Q56232
C	1014	ASP	SER	engineered mutation	UNP Q56232
C	1016	VAL	THR	engineered mutation	UNP Q56232
C	1101	SER	LYS	engineered mutation	UNP Q56232
C	1261	ARG	SER	engineered mutation	UNP Q56232
D	1514	ASP	SER	engineered mutation	UNP Q56232
D	1516	VAL	THR	engineered mutation	UNP Q56232
D	1601	SER	LYS	engineered mutation	UNP Q56232
D	1761	ARG	SER	engineered mutation	UNP Q56232

- Molecule 2 is ASPARTIC ACID (three-letter code: ASP) (formula: C<sub>4</sub>H<sub>7</sub>NO<sub>4</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	9	4	1	4	0	0
2	B	1	9	4	1	4	0	0
2	C	1	9	4	1	4	0	0
2	D	1	9	4	1	4	0	0

- Molecule 3 is PYRIDOXAL-5'-PHOSPHATE (three-letter code: PLP) (formula: C<sub>8</sub>H<sub>10</sub>NO<sub>6</sub>P).

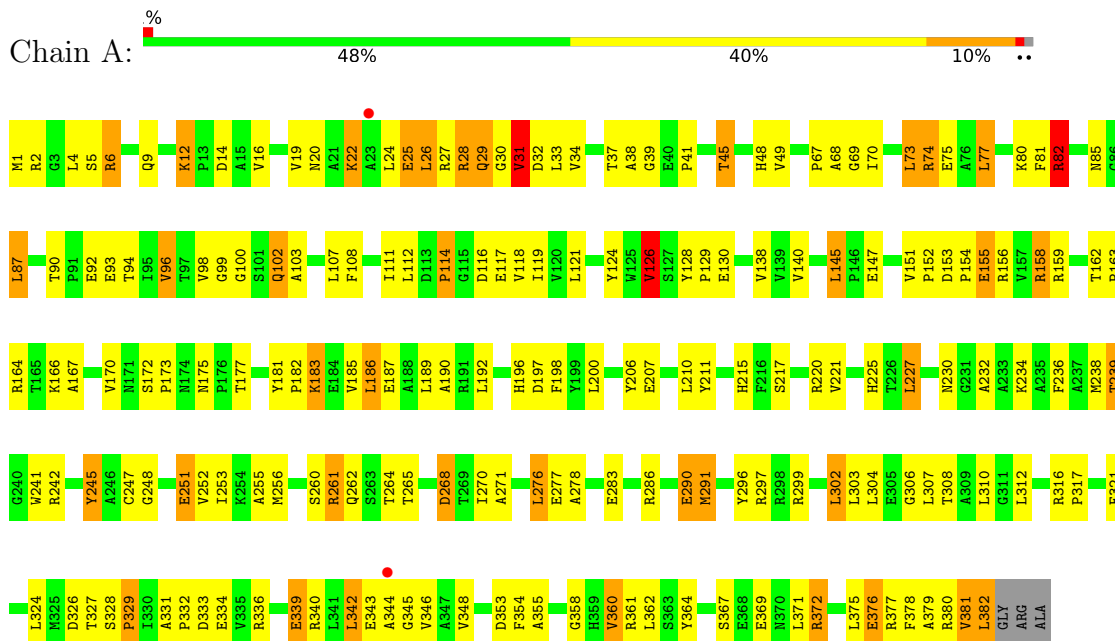


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
3	A	1	Total 15	8	1	5	1	0	0
3	B	1	Total 15	8	1	5	1	0	0
3	C	1	Total 15	8	1	5	1	0	0
3	D	1	Total 15	8	1	5	1	0	0

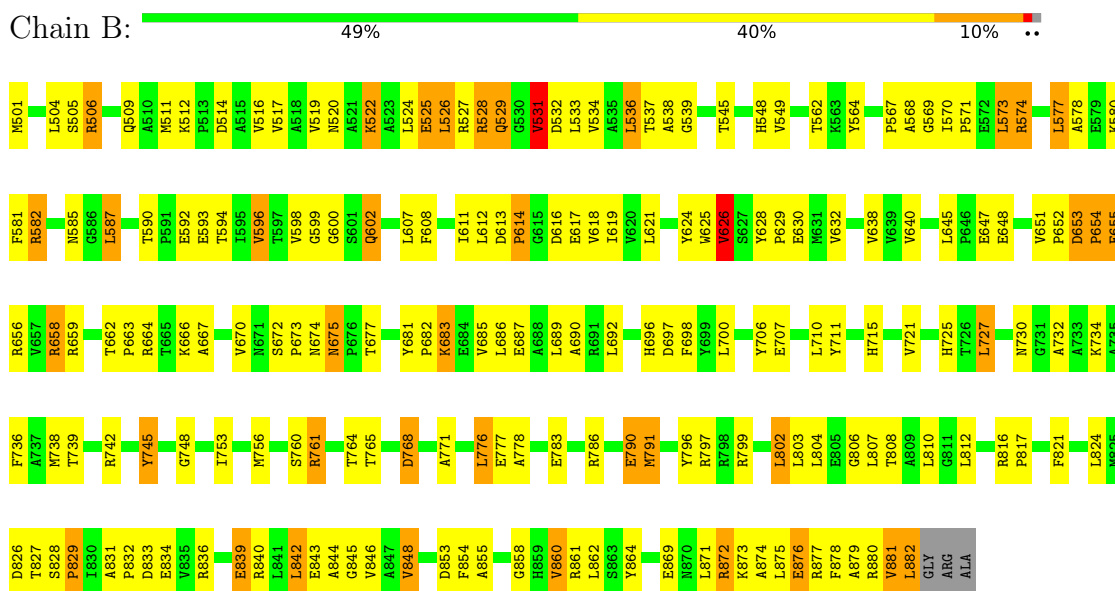
### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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 ( $RSRZ > 2$ ). 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.

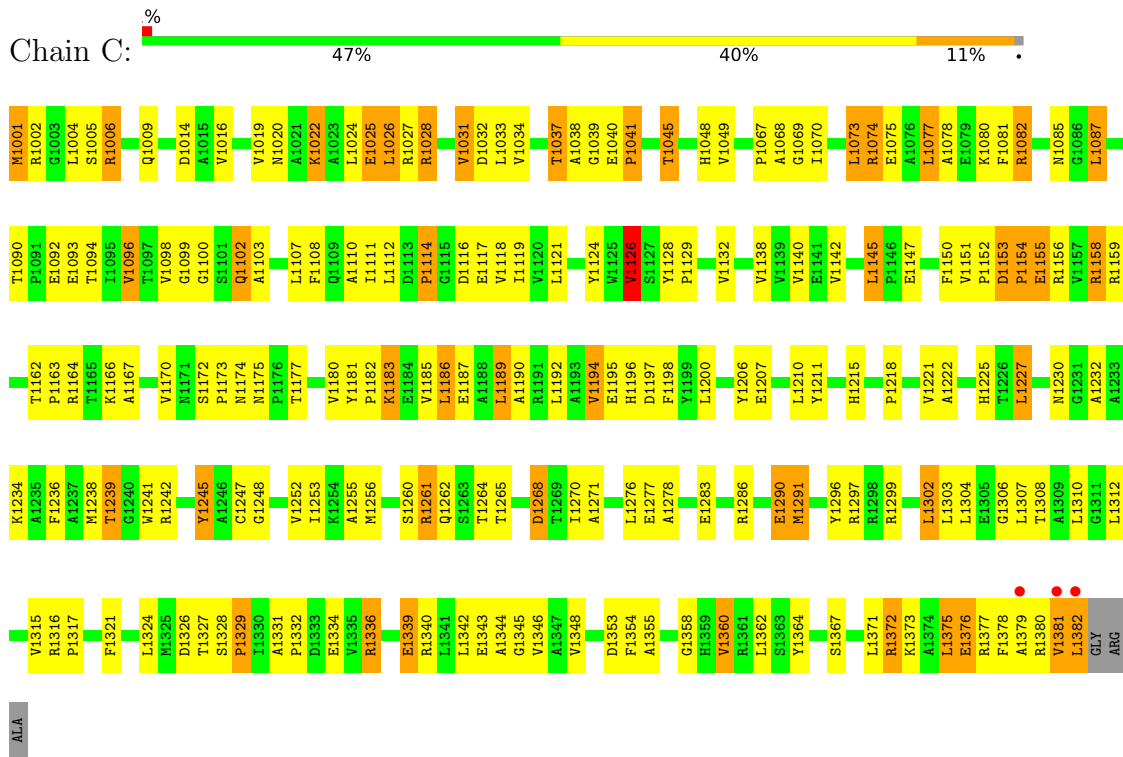
#### • Molecule 1: ASPARTATE AMINOTRANSFERASE



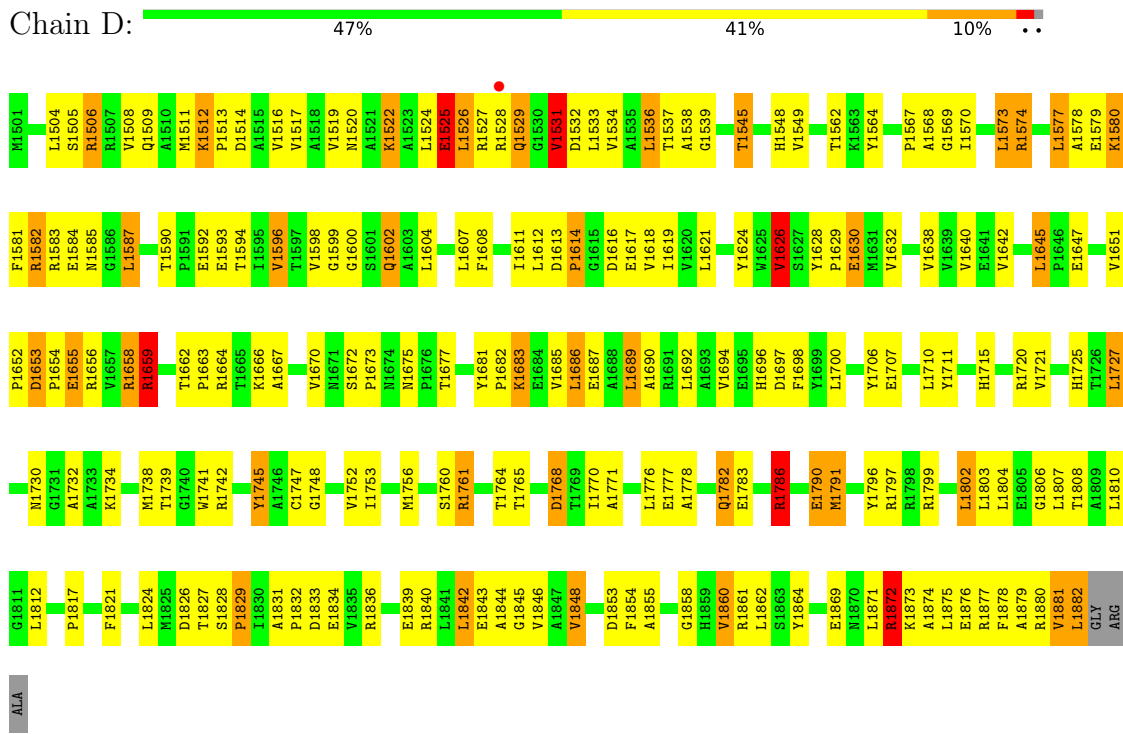
#### • Molecule 1: ASPARTATE AMINOTRANSFERASE



• Molecule 1: ASPARTATE AMINOTRANSFERASE



• Molecule 1: ASPARTATE AMINOTRANSFERASE



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	82.67Å 102.34Å 100.41Å 90.00° 112.14° 90.00°	Depositor
Resolution (Å)	8.00 – 3.30 45.29 – 3.30	Depositor EDS
% Data completeness (in resolution range)	97.0 (8.00-3.30) 97.1 (45.29-3.30)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.78 (at 3.32Å)	Xtrriage
Refinement program	X-PLOR 3.851	Depositor
R, $R_{free}$	0.217 , 0.277 0.201 , 0.249	Depositor DCC
$R_{free}$ test set	2263 reflections (9.94%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.9	Xtrriage
Anisotropy	0.526	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 57.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.41$ , $\langle L^2 \rangle = 0.23$	Xtrriage
Estimated twinning fraction	0.050 for h,-k,-h-l	Xtrriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	11888	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.94% 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: 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	0.43	0/3009	0.89	7/4092 (0.2%)
1	B	0.44	0/3009	0.87	5/4092 (0.1%)
1	C	0.42	0/3009	0.84	4/4092 (0.1%)
1	D	0.46	1/3009 (0.0%)	0.85	4/4092 (0.1%)
All	All	0.44	1/12036 (0.0%)	0.86	20/16368 (0.1%)

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	1
1	B	0	2
1	C	0	1
1	D	0	4
All	All	0	8

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	1525	GLU	CG-CD	-7.27	1.41	1.51

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	159	ARG	NE-CZ-NH2	-10.47	115.06	120.30
1	A	361	ARG	NE-CZ-NH1	-10.21	115.20	120.30
1	A	159	ARG	NE-CZ-NH1	9.66	125.13	120.30
1	B	659	ARG	NE-CZ-NH1	6.91	123.76	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	659	ARG	NE-CZ-NH2	-6.84	116.88	120.30

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	82	ARG	Sidechain
1	B	528	ARG	Sidechain
1	B	872	ARG	Sidechain
1	C	1159	ARG	Sidechain
1	D	1659	ARG	Sidechain

## 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	2948	0	2965	167	0
1	B	2948	0	2962	161	1
1	C	2948	0	2962	175	0
1	D	2948	0	2962	162	1
2	A	9	0	3	0	0
2	B	9	0	3	3	0
2	C	9	0	3	0	0
2	D	9	0	3	0	0
3	A	15	0	6	1	0
3	B	15	0	6	2	0
3	C	15	0	6	1	0
3	D	15	0	6	1	0
All	All	11888	0	11887	622	1

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 26.

The worst 5 of 622 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:D:1653:ASP:HB3	1:D:1656:ARG:HH21	1.28	0.97
1:B:653:ASP:HB3	1:B:656:ARG:HH21	1.32	0.94
1:A:372:ARG:HG3	1:A:372:ARG:HH11	1.31	0.94
1:C:1372:ARG:HG3	1:C:1372:ARG:HH11	1.33	0.92
1:D:1653:ASP:HB3	1:D:1656:ARG:NH2	1.86	0.89

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:528:ARG:O	1:D:1525:GLU:OE1[1_445]	2.19	0.01

## 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	380/385 (99%)	329 (87%)	45 (12%)	6 (2%)	9	36
1	B	380/385 (99%)	329 (87%)	46 (12%)	5 (1%)	12	40
1	C	380/385 (99%)	330 (87%)	44 (12%)	6 (2%)	9	36
1	D	380/385 (99%)	329 (87%)	44 (12%)	7 (2%)	8	35
All	All	1520/1540 (99%)	1317 (87%)	179 (12%)	24 (2%)	9	36

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	31	VAL
1	B	531	VAL
1	C	1031	VAL
1	D	1531	VAL
1	D	1626	VAL

### 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	305/306 (100%)	242 (79%)	63 (21%)	1	4
1	B	305/306 (100%)	245 (80%)	60 (20%)	1	5
1	C	305/306 (100%)	241 (79%)	64 (21%)	1	4
1	D	305/306 (100%)	241 (79%)	64 (21%)	1	4
All	All	1220/1224 (100%)	969 (79%)	251 (21%)	1	4

5 of 251 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	848	VAL
1	D	1686	LEU
1	C	1114	PRO
1	D	1664	ARG
1	D	1802	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	1196	HIS
1	D	1559	GLN
1	D	1696	HIS
1	D	1671	ASN
1	B	529	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides 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	ASP	A	414	3	2,8,8	0.33	0	1,10,10	5.49	1 (100%)
3	PLP	D	1913	2	15,15,16	2.72	4 (26%)	20,22,23	1.95	6 (30%)
3	PLP	A	413	2	15,15,16	2.02	4 (26%)	20,22,23	2.04	5 (25%)
3	PLP	B	913	2	15,15,16	2.44	6 (40%)	20,22,23	2.16	6 (30%)
3	PLP	C	1413	2	15,15,16	1.71	4 (26%)	20,22,23	2.21	7 (35%)
2	ASP	B	914	3	2,8,8	0.29	0	1,10,10	3.26	1 (100%)
2	ASP	D	1914	3	2,8,8	1.40	0	1,10,10	3.74	1 (100%)
2	ASP	C	1414	3	2,8,8	0.70	0	1,10,10	1.22	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	ASP	A	414	3	-	1/2/8/8	-
3	PLP	D	1913	2	-	2/6/6/8	0/1/1/1
3	PLP	A	413	2	-	0/6/6/8	0/1/1/1
3	PLP	B	913	2	-	4/6/6/8	0/1/1/1
3	PLP	C	1413	2	-	0/6/6/8	0/1/1/1
2	ASP	B	914	3	-	1/2/8/8	-
2	ASP	D	1914	3	-	1/2/8/8	-
2	ASP	C	1414	3	-	1/2/8/8	-

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	1913	PLP	C4A-C4	-8.28	1.34	1.51
3	B	913	PLP	C5-C4	5.90	1.47	1.40
3	A	413	PLP	C4A-C4	-5.16	1.40	1.51
3	B	913	PLP	O3-C3	-3.92	1.27	1.37
3	D	1913	PLP	C5-C4	3.45	1.44	1.40

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	413	PLP	O4P-C5A-C5	6.72	122.15	109.35
3	C	1413	PLP	O4P-C5A-C5	6.40	121.55	109.35
3	D	1913	PLP	O4P-C5A-C5	5.53	119.88	109.35
2	A	414	ASP	CB-CA-C	-5.49	101.79	110.69
3	B	913	PLP	C3-C4-C5	-5.11	113.22	118.74

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	414	ASP	C-CA-CB-CG
2	B	914	ASP	C-CA-CB-CG
2	C	1414	ASP	C-CA-CB-CG
2	D	1914	ASP	C-CA-CB-CG
3	B	913	PLP	C5A-O4P-P-O1P

There are no ring outliers.

5 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	1913	PLP	1	0
3	A	413	PLP	1	0
3	B	913	PLP	2	0
3	C	1413	PLP	1	0
2	B	914	ASP	3	0

## 5.7 Other polymers [\(i\)](#)

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	382/385 (99%)	-0.37	2 (0%) 91   91	2, 12, 35, 52	0
1	B	382/385 (99%)	-0.47	0 100   100	2, 12, 35, 57	0
1	C	382/385 (99%)	-0.32	3 (0%) 86   86	2, 14, 38, 52	0
1	D	382/385 (99%)	-0.35	1 (0%) 94   94	2, 13, 35, 55	0
All	All	1528/1540 (99%)	-0.38	6 (0%) 92   93	2, 13, 36, 57	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	1381	VAL	3.0
1	C	1382	LEU	2.4
1	C	1379	ALA	2.4
1	D	1528	ARG	2.3
1	A	23	ALA	2.1

### 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 monosaccharides 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	ASP	B	914	9/9	0.83	0.29	15,22,26,27	0
2	ASP	A	414	9/9	0.86	0.28	2,7,24,29	0
2	ASP	C	1414	9/9	0.87	0.21	17,24,29,29	0
2	ASP	D	1914	9/9	0.90	0.25	18,30,36,39	0
3	PLP	B	913	15/16	0.95	0.16	2,5,11,14	0
3	PLP	C	1413	15/16	0.96	0.15	2,6,10,11	0
3	PLP	A	413	15/16	0.97	0.15	3,8,13,17	0
3	PLP	D	1913	15/16	0.97	0.14	6,13,21,22	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.