



wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 02:12 am BST

PDB ID : 5AA6
Title : Homohexameric Structure of the second Vanadate-Dependent Bromoperoxidase (AnII) from *Ascophyllum nodosum*
Authors : Radlow, M.; Jeudy, A.; Dabin, J.; Delage, L.; Leblanc, C.; Hartung, J.; Czjzek, M.
Deposited on : 2015-07-23
Resolution : 2.26 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

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

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

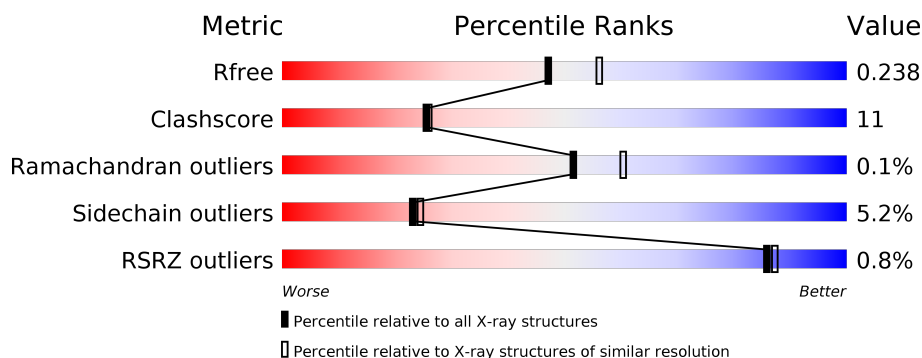
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.26 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	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)
RSRZ outliers	127900	1356 (2.26-2.26)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 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	597	<div> <div style="width: 83%;"></div> <div style="width: 15%;"></div> <div style="width: 2%;"></div> </div> <div>83% 15% ..</div>
1	B	597	<div> <div style="width: 79%;"></div> <div style="width: 19%;"></div> <div style="width: 2%;"></div> </div> <div>79% 19% .</div>
1	C	597	<div> <div style="width: 78%;"></div> <div style="width: 18%;"></div> <div style="width: 4%;"></div> </div> <div>78% 18% ..</div>
1	D	597	<div> <div style="width: 80%;"></div> <div style="width: 17%;"></div> <div style="width: 3%;"></div> </div> <div>80% 17% ..</div>
1	E	597	<div> <div style="width: 82%;"></div> <div style="width: 14%;"></div> <div style="width: 4%;"></div> </div> <div>82% 14% ..</div>
1	F	597	<div> <div style="width: 82%;"></div> <div style="width: 15%;"></div> <div style="width: 3%;"></div> </div> <div>82% 15% ..</div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 29755 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 VANADIUM-DEPENDENT BROMOPEROXIDASE 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	594	Total	C	N	O	S	0	11	0
			4601	2906	786	878	31			
1	B	595	Total	C	N	O	S	4	11	0
			4620	2912	791	885	32			
1	C	594	Total	C	N	O	S	0	17	0
			4640	2930	790	887	33			
1	D	594	Total	C	N	O	S	0	13	0
			4614	2912	786	883	33			
1	E	594	Total	C	N	O	S	0	16	0
			4625	2924	786	882	33			
1	F	594	Total	C	N	O	S	0	13	0
			4618	2914	785	888	31			

There are 36 discrepancies between the modelled and reference sequences:

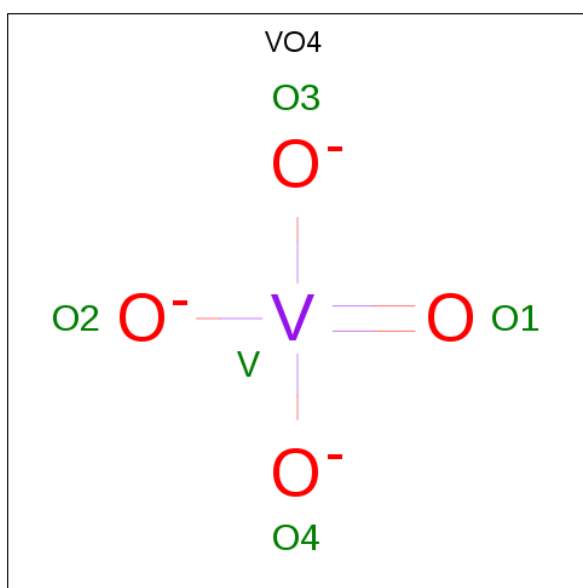
Chain	Residue	Modelled	Actual	Comment	Reference
A	68	VAL	ALA	conflict	UNP K7ZUA3
A	87	ILE	VAL	conflict	UNP K7ZUA3
A	103	MET	VAL	conflict	UNP K7ZUA3
A	203	VAL	ALA	conflict	UNP K7ZUA3
A	204	VAL	ALA	conflict	UNP K7ZUA3
A	327	MET	LEU	conflict	UNP K7ZUA3
B	68	VAL	ALA	conflict	UNP K7ZUA3
B	87	ILE	VAL	conflict	UNP K7ZUA3
B	103	MET	VAL	conflict	UNP K7ZUA3
B	203	VAL	ALA	conflict	UNP K7ZUA3
B	204	VAL	ALA	conflict	UNP K7ZUA3
B	327	MET	LEU	conflict	UNP K7ZUA3
C	68	VAL	ALA	conflict	UNP K7ZUA3
C	87	ILE	VAL	conflict	UNP K7ZUA3
C	103	MET	VAL	conflict	UNP K7ZUA3
C	203	VAL	ALA	conflict	UNP K7ZUA3
C	204	VAL	ALA	conflict	UNP K7ZUA3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	327	MET	LEU	conflict	UNP K7ZUA3
D	68	VAL	ALA	conflict	UNP K7ZUA3
D	87	ILE	VAL	conflict	UNP K7ZUA3
D	103	MET	VAL	conflict	UNP K7ZUA3
D	203	VAL	ALA	conflict	UNP K7ZUA3
D	204	VAL	ALA	conflict	UNP K7ZUA3
D	327	MET	LEU	conflict	UNP K7ZUA3
E	68	VAL	ALA	conflict	UNP K7ZUA3
E	87	ILE	VAL	conflict	UNP K7ZUA3
E	103	MET	VAL	conflict	UNP K7ZUA3
E	203	VAL	ALA	conflict	UNP K7ZUA3
E	204	VAL	ALA	conflict	UNP K7ZUA3
E	327	MET	LEU	conflict	UNP K7ZUA3
F	68	VAL	ALA	conflict	UNP K7ZUA3
F	87	ILE	VAL	conflict	UNP K7ZUA3
F	103	MET	VAL	conflict	UNP K7ZUA3
F	203	VAL	ALA	conflict	UNP K7ZUA3
F	204	VAL	ALA	conflict	UNP K7ZUA3
F	327	MET	LEU	conflict	UNP K7ZUA3

- Molecule 2 is VANADATE ION (three-letter code: VO4) (formula: O₄V).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	V	0	0
			5	4	1		
2	B	1	Total	O	V	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	C	1	Total	O	V	0	0
			5	4	1		
2	D	1	Total	O	V	0	0
			5	4	1		
2	E	1	Total	O	V	0	0
			5	4	1		
2	F	1	Total	O	V	0	0
			5	4	1		

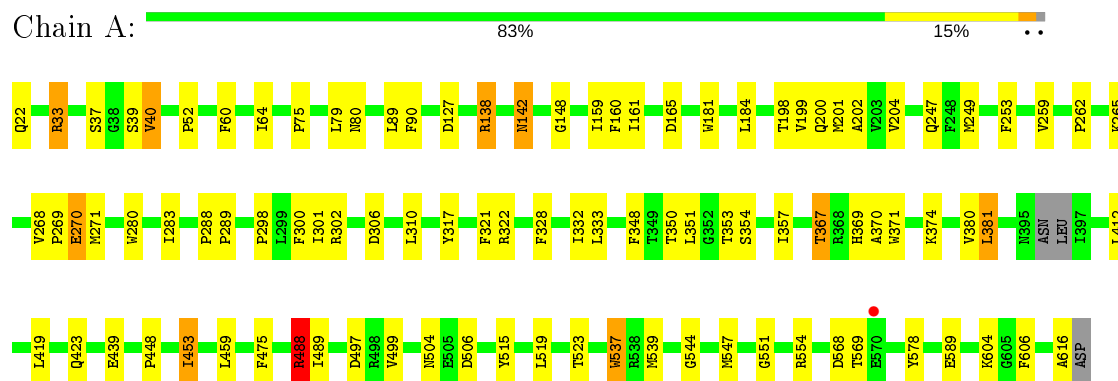
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	431	Total	O	0	0
			431	431		
3	B	338	Total	O	0	0
			338	338		
3	C	249	Total	O	0	0
			249	249		
3	D	355	Total	O	0	0
			355	355		
3	E	299	Total	O	0	0
			299	299		
3	F	335	Total	O	0	0
			335	335		

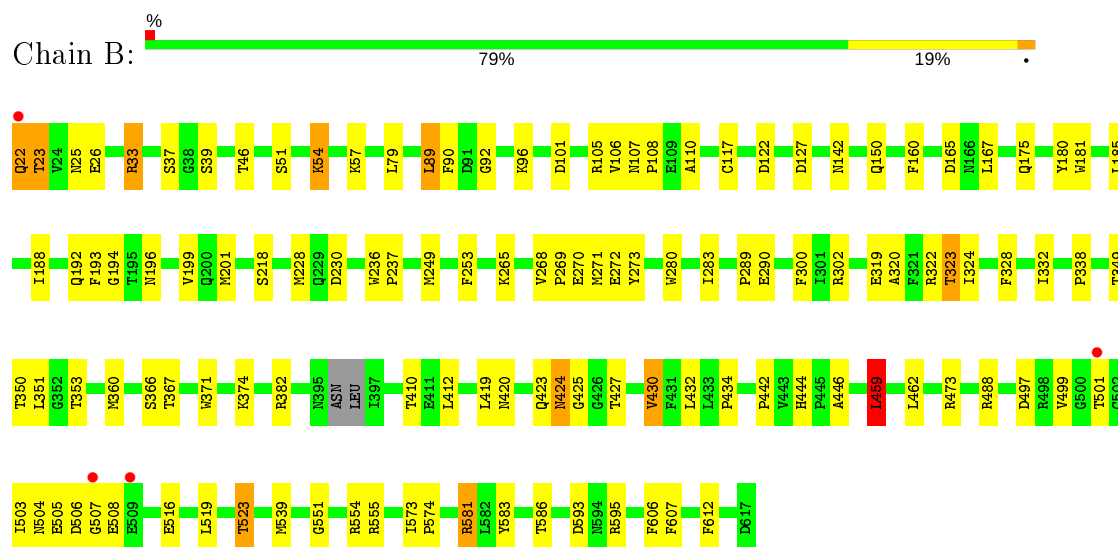
3 Residue-property plots

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

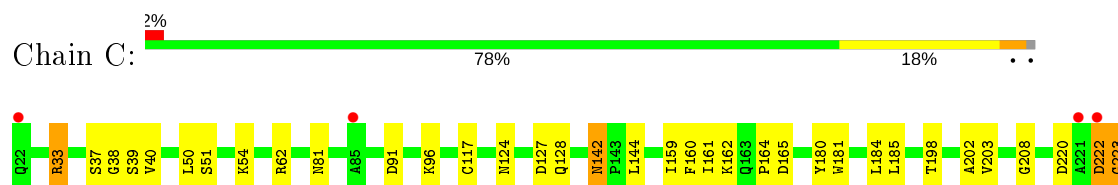
• Molecule 1: VANADIUM-DEPENDENT BROMOPEROXIDASE 2

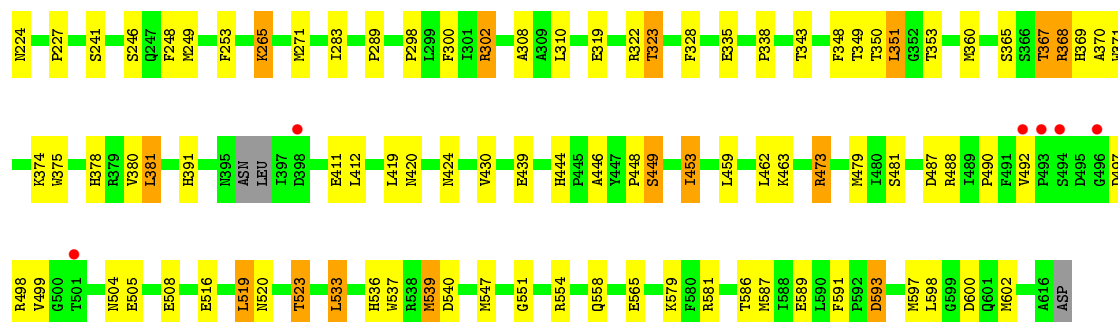


• Molecule 1: VANADIUM-DEPENDENT BROMOPEROXIDASE 2

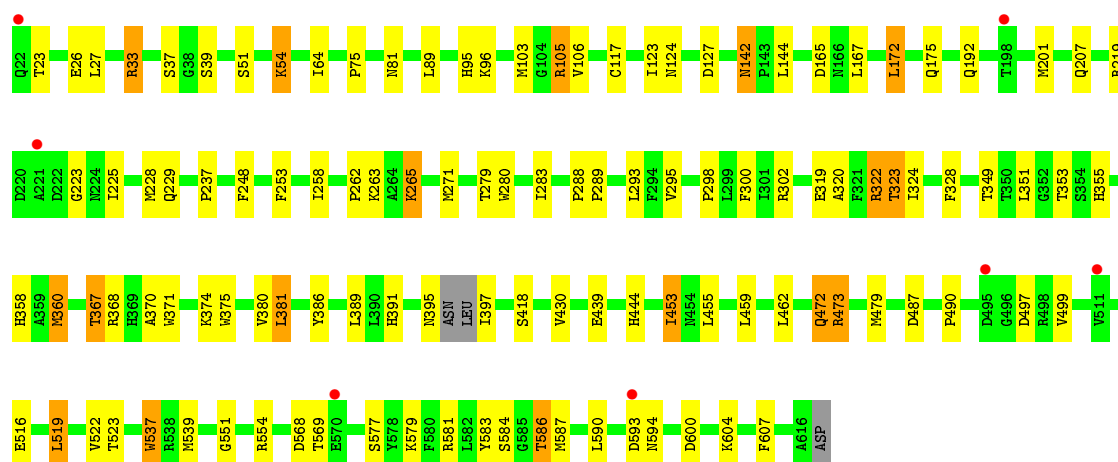
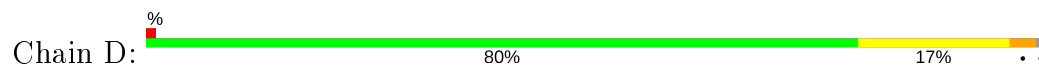


• Molecule 1: VANADIUM-DEPENDENT BROMOPEROXIDASE 2

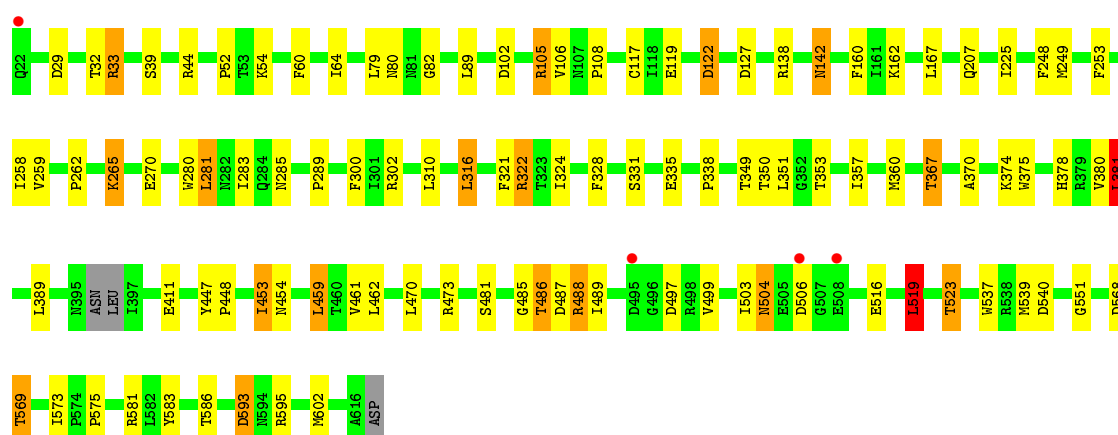
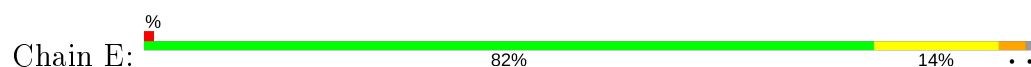




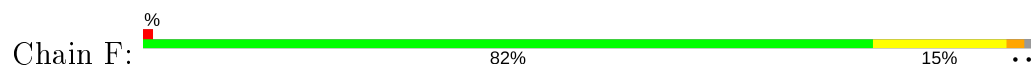
• Molecule 1: VANADIUM-DEPENDENT BROMOPEROXIDASE 2

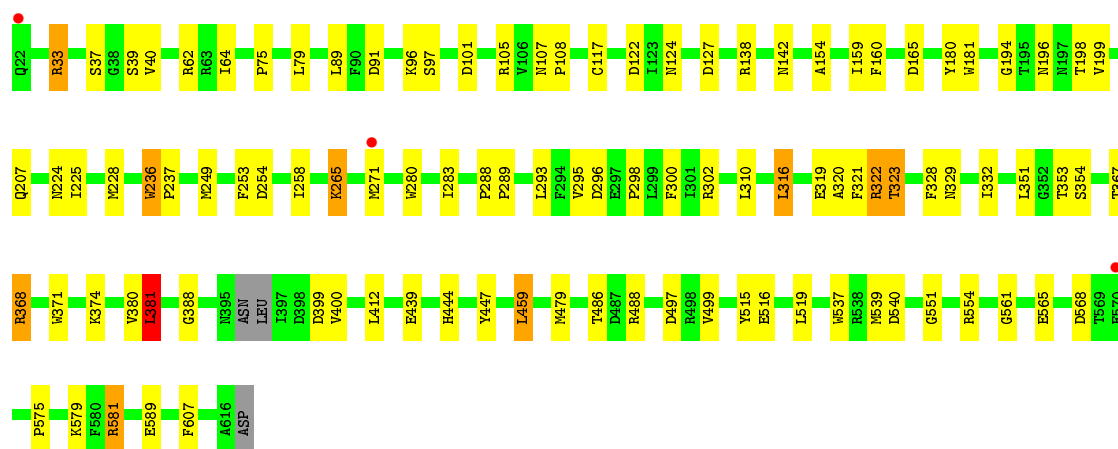


• Molecule 1: VANADIUM-DEPENDENT BROMOPEROXIDASE 2



• Molecule 1: VANADIUM-DEPENDENT BROMOPEROXIDASE 2





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	67.79Å 236.84Å 118.00Å 90.00° 93.74° 90.00°	Depositor
Resolution (Å)	70.00 – 2.26 47.20 – 2.26	Depositor EDS
% Data completeness (in resolution range)	98.6 (70.00-2.26) 98.6 (47.20-2.26)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.15 (at 2.27Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.170 , 0.236 0.172 , 0.238	Depositor DCC
R_{free} test set	8509 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	16.9	Xtriage
Anisotropy	0.109	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 43.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	29755	wwPDB-VP
Average B, all atoms (Å ²)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.03% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: VO4

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	0.61	2/4734 (0.0%)	0.75	2/6428 (0.0%)
1	B	0.61	5/4750 (0.1%)	0.73	1/6448 (0.0%)
1	C	0.58	1/4785 (0.0%)	0.71	1/6497 (0.0%)
1	D	0.60	4/4753 (0.1%)	0.75	3/6454 (0.0%)
1	E	0.60	2/4773 (0.0%)	0.74	6/6481 (0.1%)
1	F	0.62	3/4757 (0.1%)	0.77	8/6462 (0.1%)
All	All	0.60	17/28552 (0.1%)	0.74	21/38770 (0.1%)

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	505	GLU	CB-CG	-8.07	1.36	1.52
1	B	371	TRP	CD2-CE2	6.39	1.49	1.41
1	F	371	TRP	CD2-CE2	6.01	1.48	1.41
1	E	375	TRP	CD2-CE2	5.48	1.48	1.41
1	D	280	TRP	CD2-CE2	5.36	1.47	1.41

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	459	LEU	CB-CG-CD2	-8.34	96.82	111.00
1	F	316	LEU	CA-CB-CG	-6.66	99.97	115.30
1	E	316	LEU	CA-CB-CG	-6.14	101.17	115.30
1	F	122	ASP	CB-CG-OD1	6.11	123.80	118.30
1	E	322	ARG	NE-CZ-NH2	-5.64	117.48	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4601	0	4471	102	0
1	B	4620	0	4475	122	0
1	C	4640	0	4514	117	0
1	D	4614	0	4479	112	0
1	E	4625	0	4506	89	0
1	F	4618	0	4478	95	0
2	A	5	0	0	1	0
2	B	5	0	0	0	0
2	C	5	0	0	1	0
2	D	5	0	0	0	0
2	E	5	0	0	1	0
2	F	5	0	0	0	0
3	A	431	0	0	18	0
3	B	338	0	0	17	0
3	C	249	0	0	12	0
3	D	355	0	0	15	0
3	E	299	0	0	6	0
3	F	335	0	0	8	0
All	All	29755	0	26923	579	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 11.

The worst 5 of 579 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:E:33:ARG:HH11	1:E:33:ARG:CG	1.45	1.30
1:D:33:ARG:HH11	1:D:33:ARG:CG	1.40	1.30
1:A:551:GLY:HA3	3:A:2331:HOH:O	1.18	1.29
1:F:33:ARG:HH11	1:F:33:ARG:CG	1.49	1.25
1:C:33:ARG:HH11	1:C:33:ARG:CG	1.50	1.23

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	601/597 (101%)	586 (98%)	15 (2%)	0	100	100
1	B	602/597 (101%)	584 (97%)	17 (3%)	1 (0%)	47	55
1	C	607/597 (102%)	580 (96%)	26 (4%)	1 (0%)	47	55
1	D	603/597 (101%)	571 (95%)	32 (5%)	0	100	100
1	E	606/597 (102%)	588 (97%)	18 (3%)	0	100	100
1	F	603/597 (101%)	586 (97%)	17 (3%)	0	100	100
All	All	3622/3582 (101%)	3495 (96%)	125 (4%)	2 (0%)	51	60

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	89	LEU
1	C	223	GLY

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	498/495 (101%)	482 (97%)	16 (3%)	39	47
1	B	500/495 (101%)	479 (96%)	21 (4%)	30	34
1	C	505/495 (102%)	464 (92%)	41 (8%)	11	10
1	D	501/495 (101%)	464 (93%)	37 (7%)	13	12
1	E	503/495 (102%)	465 (92%)	38 (8%)	13	12

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	F	502/495 (101%)	483 (96%)	19 (4%)	33	39
All	All	3009/2970 (101%)	2837 (94%)	172 (6%)	23	20

5 of 172 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	600	ASP
1	D	293	LEU
1	F	127	ASP
1	D	54	LYS
1	D	117	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	163	GLN
1	D	22	GLN
1	E	444	HIS
1	C	124	ASN
1	F	142	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](#)

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

6 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	VO4	E	800	-	1,4,4	4.67	1 (100%)	-		
2	VO4	A	800	-	1,4,4	4.83	1 (100%)	-		
2	VO4	C	800	-	1,4,4	5.16	1 (100%)	-		
2	VO4	D	800	-	1,4,4	5.05	1 (100%)	-		
2	VO4	F	800	-	1,4,4	4.96	1 (100%)	-		
2	VO4	B	800	-	1,4,4	5.00	1 (100%)	-		

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	800	VO4	O1-V	5.16	1.93	1.63
2	D	800	VO4	O1-V	5.05	1.92	1.63
2	B	800	VO4	O1-V	5.00	1.92	1.63
2	F	800	VO4	O1-V	4.96	1.92	1.63
2	A	800	VO4	O1-V	4.83	1.91	1.63

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	800	VO4	1	0
2	A	800	VO4	1	0
2	C	800	VO4	1	0

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, 95th 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 ²)	Q<0.9
1	A	594/597 (99%)	-0.53	1 (0%) 95 96	8, 15, 29, 56	13 (2%)
1	B	595/597 (99%)	-0.35	4 (0%) 87 88	11, 19, 34, 65	12 (2%)
1	C	594/597 (99%)	-0.25	10 (1%) 70 73	12, 21, 41, 74	7 (1%)
1	D	594/597 (99%)	-0.39	7 (1%) 79 81	10, 18, 33, 74	13 (2%)
1	E	594/597 (99%)	-0.45	4 (0%) 87 88	9, 17, 33, 62	6 (1%)
1	F	594/597 (99%)	-0.43	3 (0%) 91 91	8, 15, 29, 57	8 (1%)
All	All	3565/3582 (99%)	-0.40	29 (0%) 86 87	8, 18, 34, 74	59 (1%)

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	22	GLN	4.7
1	D	221	ALA	3.5
1	B	507	GLY	3.1
1	E	495	ASP	3.0
1	C	492	VAL	2.9

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

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th 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(Å ²)	Q<0.9
2	VO4	C	800	5/5	0.98	0.13	27,27,28,31	0
2	VO4	A	800	5/5	0.99	0.14	16,16,16,19	0
2	VO4	E	800	5/5	0.99	0.17	18,18,19,20	2
2	VO4	D	800	5/5	0.99	0.14	23,23,24,26	0
2	VO4	F	800	5/5	0.99	0.14	21,22,23,23	0
2	VO4	B	800	5/5	0.99	0.15	25,26,27,29	0

6.5 Other polymers [i](#)

There are no such residues in this entry.