wwPDB X-ray Structure Validation Summary Report (i)

May 22, 2020 – 04:12 am BST

PDB ID	:	2D1G
Title	:	Structure of Francisella tularensis Acid Phosphatase A (AcpA) bound to orthovanadate
Authors	:	Felts, R.L.; Reilly, T.J.; Tanner, J.J.
Deposited on	:	2005-08-20
Resolution	:	1.75 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 (i) symbol.

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

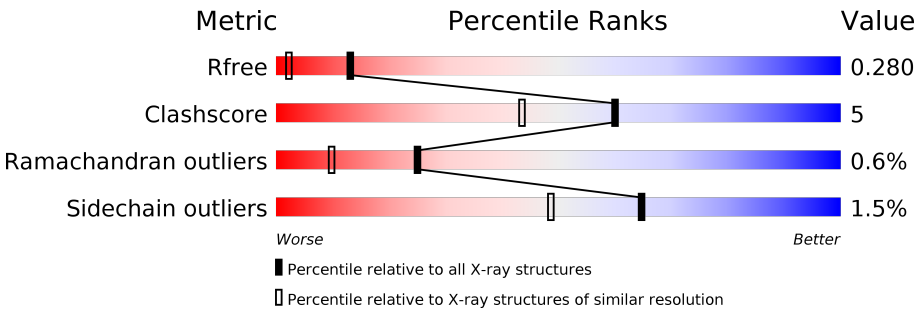
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.75 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	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	498	89% 7% • •
1	B	498	85% 9% • 5%

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

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	VO4	B	901	-	-	X	-
5	ETX	A	1006	-	-	X	-

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 8182 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.

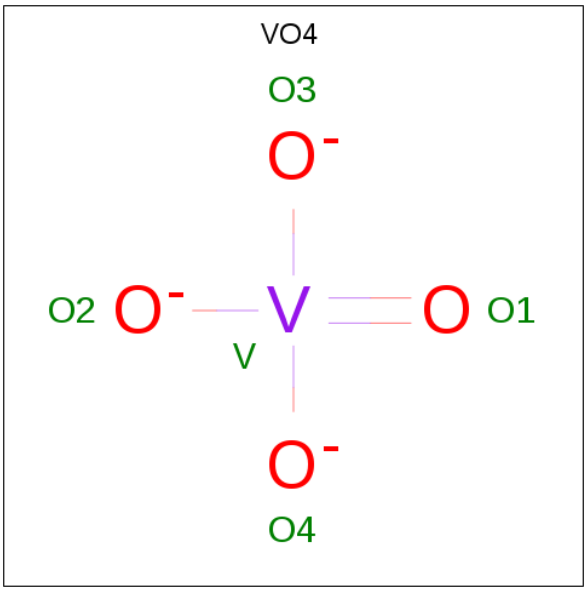
- Molecule 1 is a protein called acid phosphatase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	481	3817	2413	644	746	14	0	5	0
1	B	472	3725	2352	630	730	13	0	1	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	INITIATING METHIONINE	UNP Q2A5P7
A	375	ASP	ALA	SEE REMARK 999	UNP Q2A5P7
A	376	TRP	GLY	SEE REMARK 999	UNP Q2A5P7
A	377	ASP	ILE	SEE REMARK 999	UNP Q2A5P7
A	378	SER	ALA	SEE REMARK 999	UNP Q2A5P7
A	379	THR	HIS	SEE REMARK 999	UNP Q2A5P7
A	494	HIS	-	EXPRESSION TAG	UNP Q2A5P7
A	495	HIS	-	EXPRESSION TAG	UNP Q2A5P7
A	496	HIS	-	EXPRESSION TAG	UNP Q2A5P7
A	497	HIS	-	EXPRESSION TAG	UNP Q2A5P7
A	498	HIS	-	EXPRESSION TAG	UNP Q2A5P7
B	1	MET	-	INITIATING METHIONINE	UNP Q2A5P7
B	375	ASP	ALA	SEE REMARK 999	UNP Q2A5P7
B	376	TRP	GLY	SEE REMARK 999	UNP Q2A5P7
B	377	ASP	ILE	SEE REMARK 999	UNP Q2A5P7
B	378	SER	ALA	SEE REMARK 999	UNP Q2A5P7
B	379	THR	HIS	SEE REMARK 999	UNP Q2A5P7
B	494	HIS	-	EXPRESSION TAG	UNP Q2A5P7
B	495	HIS	-	EXPRESSION TAG	UNP Q2A5P7
B	496	HIS	-	EXPRESSION TAG	UNP Q2A5P7
B	497	HIS	-	EXPRESSION TAG	UNP Q2A5P7
B	498	HIS	-	EXPRESSION TAG	UNP Q2A5P7

- Molecule 2 is VANADATE ION (three-letter code: VO4) (formula: O_4V).

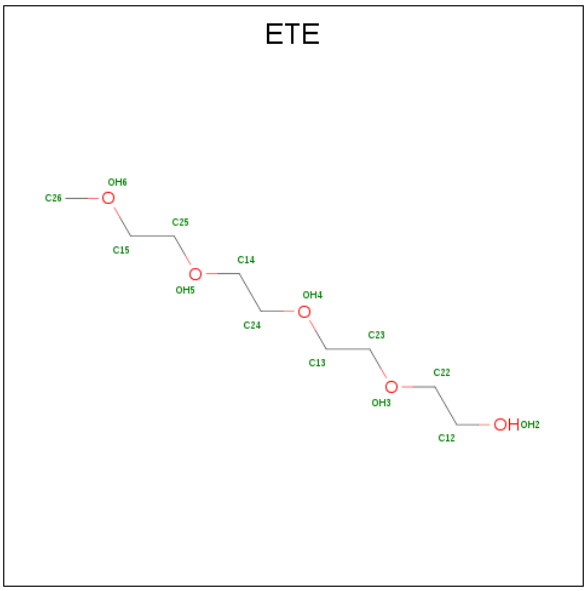


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O V 5 4 1	0	0
2	B	1	Total O V 5 4 1	0	0

- Molecule 3 is UNKNOWN ATOM OR ION (three-letter code: UNX) (formula: X).

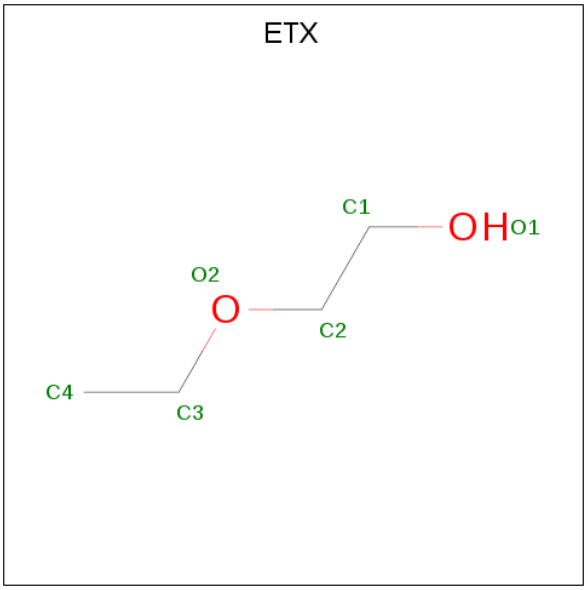
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total X 1 1	0	0
3	A	1	Total X 1 1	0	0

- Molecule 4 is 2-{2-[2-2-(METHOXY-ETHOXY)-ETHOXY]-ETHOXY}-ETHANOL (three-letter code: ETE) (formula: $C_9H_{20}O_5$).



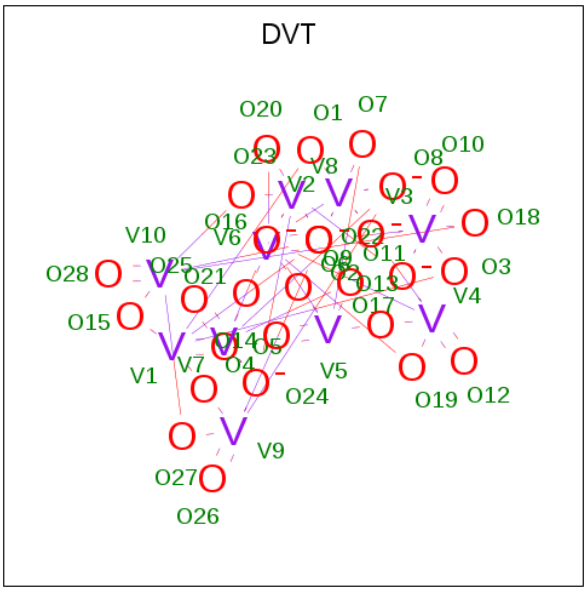
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	A	1	14	9	5	0	0

- Molecule 5 is 2-ETHOXYETHANOL (three-letter code: ETX) (formula: $C_4H_{10}O_2$).



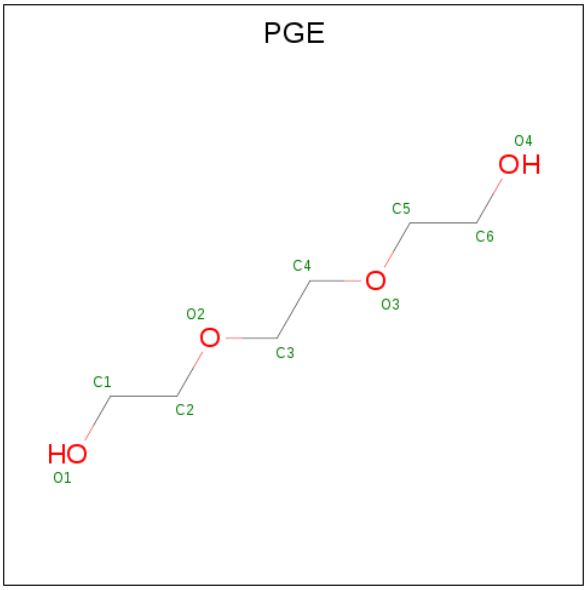
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
5	A	1	6	4	2	0	0

- Molecule 6 is DECAVANADATE (three-letter code: DVT) (formula: $O_{28}V_{10}$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	V		
6	B	1	38	28	10	0	0

- Molecule 7 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
7	B	1	10	6	4	0	0
			Total	C	O		
7	B	1	10	6	4	0	0

- Molecule 8 is water.

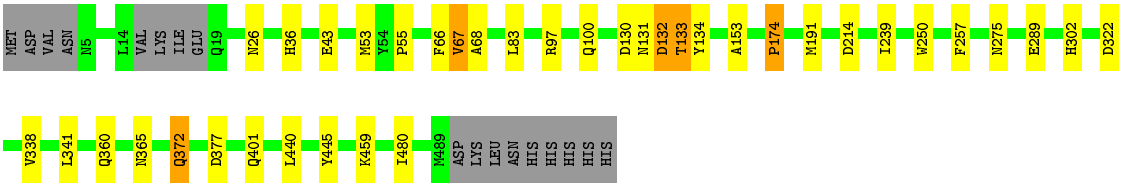
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
8	A	345	345	345	0	0
			Total	O		
8	B	205	205	205	0	0

3 Residue-property plots (i)

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. 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. 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.

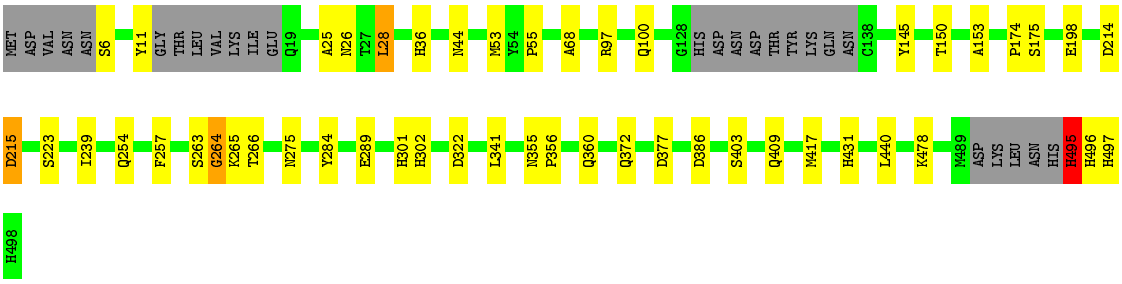
- Molecule 1: acid phosphatase

Chain A:



- Molecule 1: acid phosphatase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	112.08Å 144.40Å 123.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.28 – 1.75 44.23 – 1.75	Depositor EDS
% Data completeness (in resolution range)	100.0 (44.28-1.75) 97.2 (44.23-1.75)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.79 (at 1.75Å)	Xtriage
Refinement program	CNS, REFMAC	Depositor
R, R_{free}	0.198 , 0.231 0.263 , 0.280	Depositor DCC
R_{free} test set	4923 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	19.7	Xtriage
Anisotropy	0.140	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 56.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	8182	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.56% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: PGE, UNX, ETX, VO4, ETE, DVT

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	0.56	0/3941	0.64	1/5362 (0.0%)
1	B	0.68	9/3836 (0.2%)	0.61	1/5224 (0.0%)
All	All	0.62	9/7777 (0.1%)	0.62	2/10586 (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	4
1	B	0	1
All	All	0	5

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	223	SER	CB-OG	8.94	1.53	1.42
1	B	11	TYR	CE1-CZ	8.81	1.50	1.38
1	B	11	TYR	CG-CD2	8.79	1.50	1.39
1	B	215	ASP	CG-OD2	7.13	1.41	1.25
1	B	215	ASP	CG-OD1	6.73	1.40	1.25

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	67	VAL	N-CA-C	7.03	129.99	111.00
1	B	215	ASP	CB-CG-OD2	-5.28	113.55	118.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	130	ASP	Peptide
1	A	132	ASP	Peptide
1	A	43	GLU	Peptide
1	A	66	PHE	Peptide
1	B	495	HIS	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	3817	0	3537	31	0
1	B	3725	0	3413	37	0
2	A	5	0	0	1	0
2	B	5	0	0	2	0
3	A	1	0	0	0	0
3	B	1	0	0	1	0
4	A	14	0	20	1	0
5	A	6	0	10	4	0
6	B	38	0	0	1	0
7	B	20	0	28	0	0
8	A	345	0	0	3	0
8	B	205	0	0	4	0
All	All	8182	0	7008	67	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 5.

The worst 5 of 67 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:B:175:SER:OG	1:B:386:ASP:OD2	1.61	1.14
1:B:254:GLN:HG2	8:B:2162:HOH:O	1.74	0.85
1:A:401:GLN:HE22	1:B:431:HIS:H	1.27	0.82
1:B:26:ASN:HD21	1:B:36:HIS:HE1	1.30	0.78
1:B:214:ASP:OD1	1:B:302:HIS:HE1	1.71	0.71

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	482/498 (97%)	468 (97%)	12 (2%)	2 (0%)	34 17
1	B	465/498 (93%)	449 (97%)	12 (3%)	4 (1%)	17 5
All	All	947/996 (95%)	917 (97%)	24 (2%)	6 (1%)	25 10

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	133	THR
1	B	264	GLY
1	B	496	HIS
1	A	174	PRO
1	B	265	LYS

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	411/436 (94%)	405 (98%)	6 (2%)	65 49
1	B	399/436 (92%)	392 (98%)	7 (2%)	59 40
All	All	810/872 (93%)	797 (98%)	13 (2%)	65 45

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	377	ASP
1	B	28	LEU
1	B	417[A]	MET
1	A	372	GLN
1	B	377	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	360	GLN
1	A	372	GLN
1	B	360	GLN
1	A	365	ASN
1	A	401	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 carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 2 are unknown - leaving 7 for Mogul analysis.

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	$\#\|Z\| > 2$
7	PGE	B	1004	-	9,9,9	0.47	0	8,8,8	0.41	0
7	PGE	B	1002	-	9,9,9	0.41	0	8,8,8	0.29	0
5	ETX	A	1006	-	5,5,5	0.45	0	4,4,4	0.43	0
2	VO4	B	901	-	1,4,4	4.79	1 (100%)	-		
6	DVT	B	2001	-	60,60,60	0.90	3 (5%)	-		
4	ETE	A	1001	-	13,13,13	0.53	0	12,12,12	0.20	0
2	VO4	A	801	-	1,4,4	5.06	1 (100%)	-		

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
7	PGE	B	1004	-	-	0/7/7/7	-
4	ETE	A	1001	-	-	1/11/11/11	-
5	ETX	A	1006	-	-	1/3/3/3	-
7	PGE	B	1002	-	-	1/7/7/7	-

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	801	VO4	O1-V	5.06	1.92	1.63
2	B	901	VO4	O1-V	4.79	1.91	1.63
6	B	2001	DVT	O12-V4	2.44	1.63	1.60
6	B	2001	DVT	O2-V10	-2.34	2.26	2.33
6	B	2001	DVT	O16-V5	-2.22	2.26	2.33

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	1006	ETX	O1-C1-C2-O2
7	B	1002	PGE	C3-C4-O3-C5
4	A	1001	ETE	OH4-C13-C23-OH3

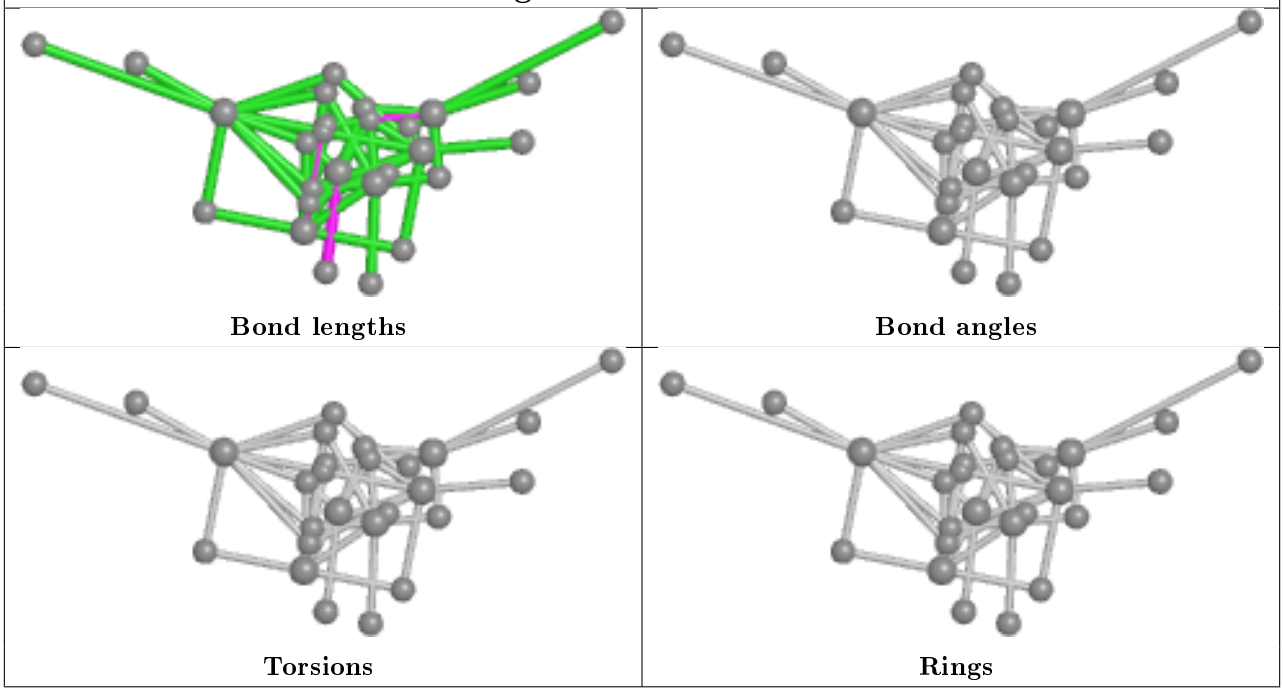
There are no ring outliers.

5 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1006	ETX	4	0
2	B	901	VO4	2	0
6	B	2001	DVT	1	0
4	A	1001	ETE	1	0
2	A	801	VO4	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand DVT B 2001



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

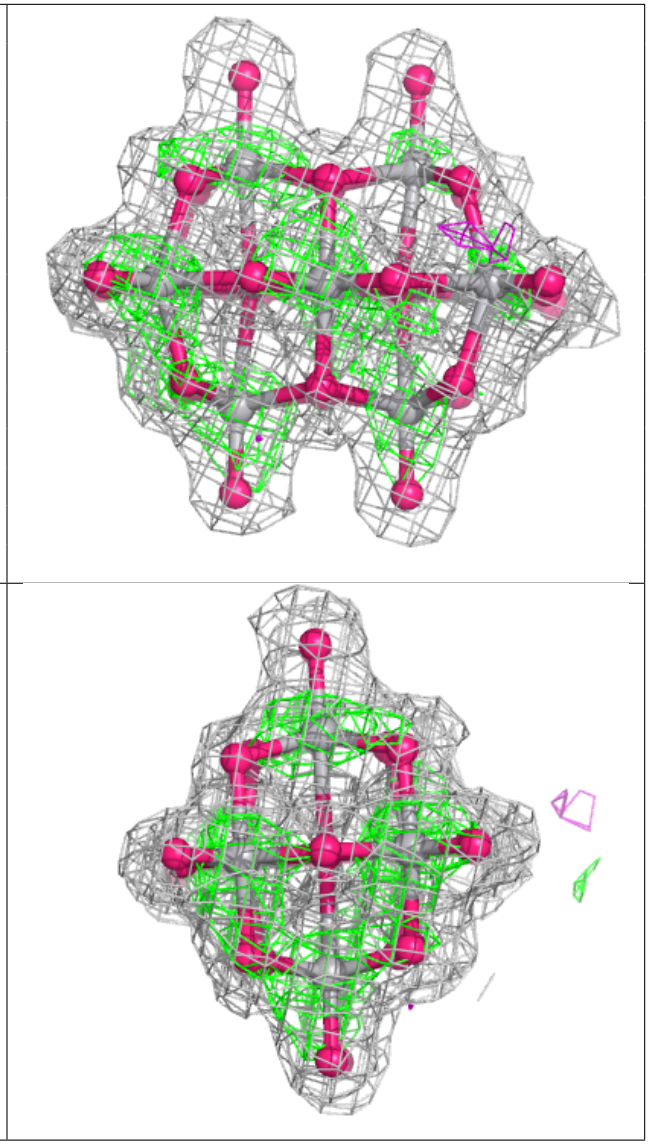
6.4 Ligands (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around DVT B 2001:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.