Full wwPDB X-ray Structure Validation Report ⓘ

Jun 17, 2024 – 03:05 AM EDT

PDB ID	:	5HQW
Title	:	Crystal structure of a trans-AT PKS dehydratase domain of C0ZGQ6 from Brevibacillus brevis
Authors	:	Jakob, R.P.; Hauswirth, P.; Dilmi, J.; Herbst, D.A.; Maier, T.
Deposited on	:	2016-01-22
Resolution	:	2.39 Å(reported)

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

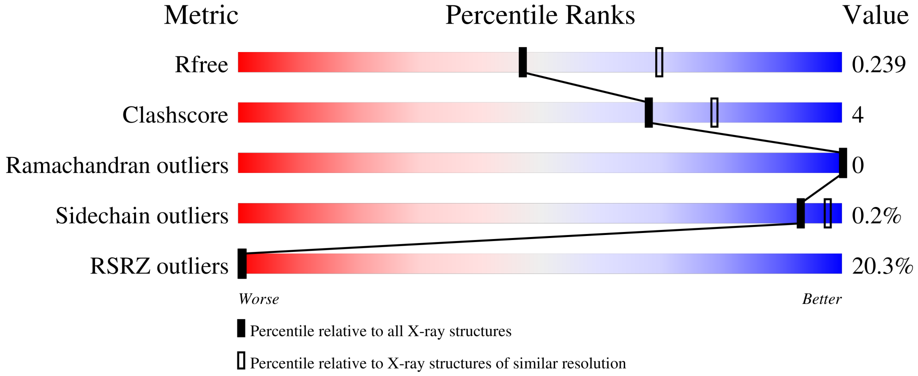
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance ⓘ

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

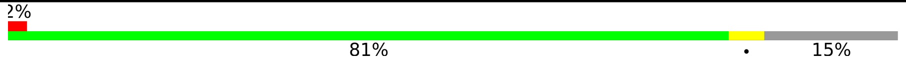
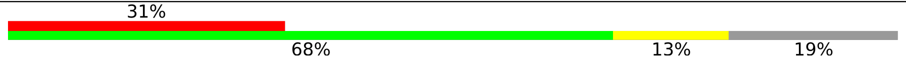
The reported resolution of this entry is 2.39 Å.

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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 >=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% 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	336	
1	B	336	

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 8667 atoms, of which 4236 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 Putative polyketide synthase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	287	4402	1405	2184	374	428	11	0	1	0
			Total	C	H	N	O	S			
1	B	271	4140	1323	2052	353	403	9	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2205	SER	-	expression tag	UNP C0ZGQ6
A	2206	MET	-	expression tag	UNP C0ZGQ6
B	2205	SER	-	expression tag	UNP C0ZGQ6
B	2206	MET	-	expression tag	UNP C0ZGQ6

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
2	A	1	1	1	0	0

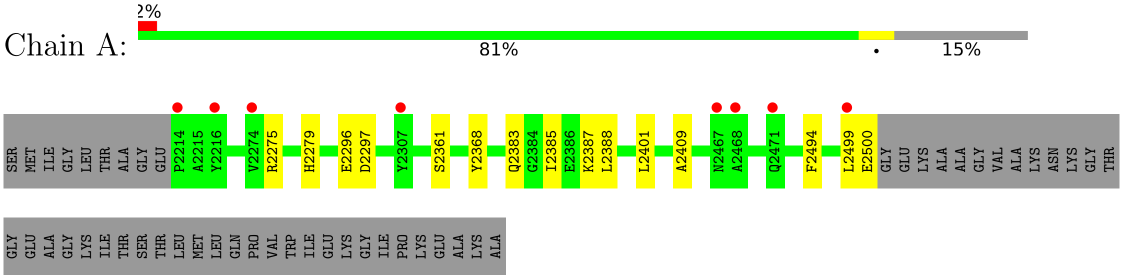
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	105	105	105	0	0
			Total	O		
3	B	19	19	19	0	0

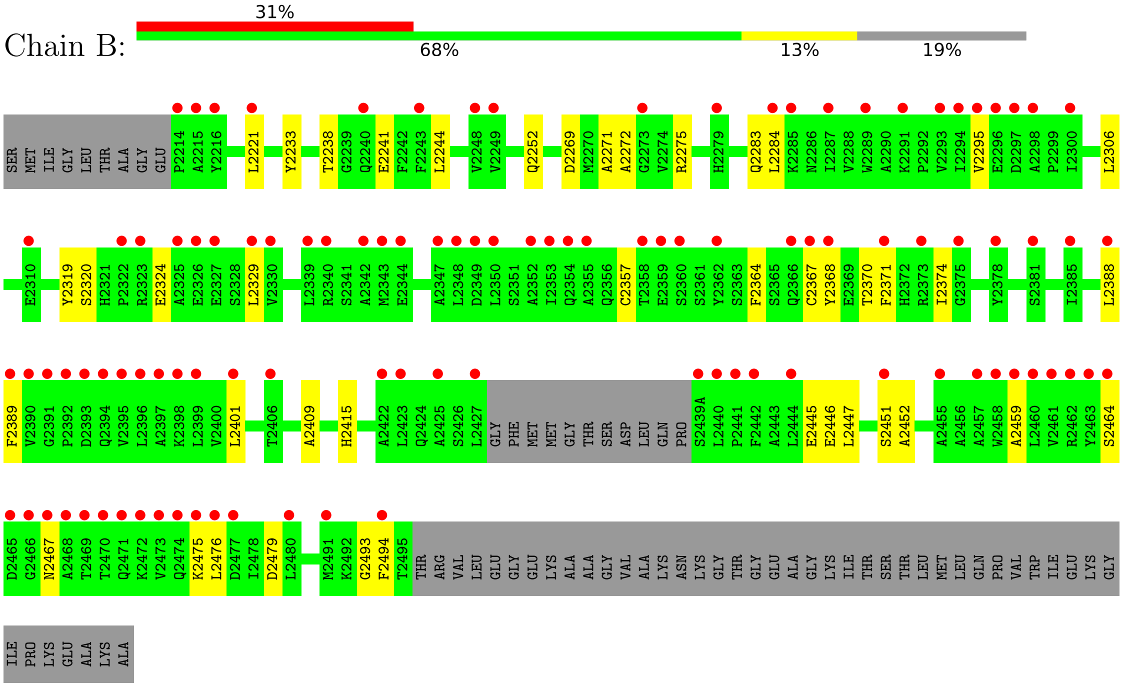
3 Residue-property plots ⓘ

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: Putative polyketide synthase



- Molecule 1: Putative polyketide synthase



4 Data and refinement statistics ⓘ

Property	**Value**	**Source**
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	41.01Å 63.28Å 124.14Å 90.00° 91.91° 90.00°	Depositor
Resolution (Å)	62.03 – 2.39 62.04 – 2.39	Depositor EDS
% Data completeness (in resolution range)	97.9 (62.03-2.39) 97.9 (62.04-2.39)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.65 (at 2.40Å)	Xtriage
Refinement program	PHENIX (1.10pre_2131: ???)	Depositor
R, R_{free}	0.210 , 0.236 0.212 , 0.239	Depositor DCC
R_{free} test set	1245 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	43.6	Xtriage
Anisotropy	0.659	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 64.4	EDS
L-test for twinning[2]	$< \|L\| >$ = 0.49, $< L^2 >$ = 0.33	Xtriage
Estimated twinning fraction	0.039 for h,-k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	8667	wwPDB-VP
Average B, all atoms (Å^2)	93.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.32% 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 ⓘ

5.1 Standard geometry ⓘ

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

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	0.27	0/2266	0.49	0/3081
1	B	0.26	0/2133	0.45	0/2900
All	All	0.26	0/4399	0.47	0/5981

There are no bond length outliers.

There are no bond angle outliers.

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	2218	2184	2183	7	0
1	B	2088	2052	2051	25	1
2	A	1	0	0	0	0
3	A	105	0	0	0	0
3	B	19	0	0	1	0
All	All	4431	4236	4234	32	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 4.

All (32) 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:2283:GLN:NE2	1:B:2446:GLU:OE2	2.21	0.73
1:B:2324:GLU:OE1	3:B:2601:HOH:O	2.09	0.69
1:B:2269:ASP:OD1	1:B:2275:ARG:NH1	2.32	0.63
1:B:2464:SER:HB2	1:B:2475:LYS:HE3	1.83	0.59
1:B:2244:LEU:HD13	1:B:2295:VAL:HG21	1.89	0.55
1:B:2464:SER:CB	1:B:2475:LYS:HE3	2.38	0.54
1:B:2370:THR:O	1:B:2374:ILE:N	2.33	0.54
1:B:2319:TYR:CD1	1:B:2329:LEU:HD11	2.44	0.53
1:B:2238:THR:O	1:B:2241:GLU:OE1	2.30	0.50
1:B:2367:CYS:SG	1:B:2388:LEU:HD22	2.53	0.49
1:B:2371:PHE:HA	1:B:2374:ILE:HG22	1.95	0.47
1:B:2271:ALA:CB	1:B:2306:LEU:HD12	2.46	0.46
1:A:2401:LEU:HD21	1:A:2409:ALA:HB2	1.96	0.46
1:A:2275:ARG:HG2	1:A:2279:HIS:HB2	1.99	0.45
1:A:2296:GLU:O	1:A:2297:ASP:HB2	2.16	0.44
1:A:2385:ILE:HG21	1:A:2388:LEU:HD13	2.00	0.44
1:B:2364:PHE:O	1:B:2368:TYR:N	2.44	0.44
1:B:2401:LEU:HD21	1:B:2409:ALA:HB2	2.00	0.43
1:B:2272:ALA:HB1	1:B:2275:ARG:HD3	2.02	0.42
1:B:2445:GLU:HB2	1:B:2493:GLY:HA3	2.01	0.42
1:B:2320:SER:O	1:B:2329:LEU:HD12	2.19	0.42
1:B:2409:ALA:HA	1:B:2415:HIS:CD2	2.54	0.42
1:B:2451:SER:OG	1:B:2452:ALA:N	2.52	0.42
1:A:2361:SER:OG	1:A:2387:LYS:HE2	2.19	0.42
1:B:2475:LYS:HA	1:B:2494:PHE:O	2.20	0.42
1:B:2475:LYS:O	1:B:2476:LEU:HD23	2.20	0.41
1:A:2368:TYR:CZ	1:A:2383:GLN:HA	2.55	0.41
1:B:2357:CYS:HB3	1:B:2389:PHE:C	2.41	0.41
1:A:2499:LEU:HD13	1:A:2500:GLU:N	2.35	0.41
1:B:2284:LEU:HD12	1:B:2447:LEU:HD23	2.03	0.41
1:B:2459:ALA:HA	1:B:2479:ASP:O	2.20	0.41
1:B:2221:LEU:HD21	1:B:2233:TYR:CD1	2.56	0.41

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:2252:GLN:NE2	1:B:2467:ASN:O[2_559]	2.18	0.02

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	286/336 (85%)	276 (96%)	10 (4%)	0	100	100
1	B	267/336 (80%)	253 (95%)	14 (5%)	0	100	100
All	All	553/672 (82%)	529 (96%)	24 (4%)	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	242/275 (88%)	241 (100%)	1 (0%)	91	96
1	B	227/275 (82%)	227 (100%)	0	100	100
All	All	469/550 (85%)	468 (100%)	1 (0%)	93	97

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2494	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

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

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, 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	287/336 (85%)	0.64	8 (2%) 53 51	33, 45, 83, 151	0
1	B	271/336 (80%)	2.12	105 (38%) 0 0	49, 123, 192, 228	0
All	All	558/672 (83%)	1.36	113 (20%) 1 0	33, 67, 173, 228	0

All (113) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	2439(A)	SER	10.6
1	B	2440	LEU	10.4
1	B	2389	PHE	9.6
1	B	2471	GLN	9.0
1	B	2460	LEU	8.8
1	B	2476	LEU	8.6
1	B	2352	ALA	8.0
1	B	2470	THR	8.0
1	B	2371	PHE	7.9
1	B	2388	LEU	7.7
1	B	2349	ASP	7.7
1	B	2395	VAL	7.5
1	B	2461	VAL	7.1
1	A	2214	PRO	6.6
1	B	2463	TYR	6.5
1	A	2467	ASN	6.3
1	B	2425	ALA	6.2
1	B	2396	LEU	6.0
1	B	2375	GLY	5.8
1	B	2469	THR	5.7
1	B	2474	GLN	5.7
1	B	2427	LEU	5.6
1	B	2393	ASP	5.5
1	A	2471	GLN	5.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	2350	LEU	5.4
1	B	2473	VAL	5.4
1	B	2480	LEU	5.3
1	B	2468	ALA	5.3
1	B	2323	ARG	5.2
1	B	2368	TYR	5.2
1	B	2466	GLY	5.1
1	B	2392	PRO	5.0
1	B	2397	ALA	4.9
1	B	2467	ASN	4.9
1	B	2348	LEU	4.7
1	B	2353	ILE	4.6
1	B	2360	SER	4.6
1	B	2472	LYS	4.6
1	B	2342	ALA	4.5
1	A	2499	LEU	4.5
1	B	2284	LEU	4.4
1	B	2310	GLU	4.4
1	B	2295	VAL	4.3
1	B	2296	GLU	4.2
1	B	2325	ALA	4.2
1	B	2398	LYS	4.2
1	B	2339	LEU	4.1
1	B	2422	ALA	4.0
1	B	2441	PRO	4.0
1	B	2358	THR	3.9
1	B	2354	GLN	3.9
1	B	2391	GLY	3.9
1	B	2248	VAL	3.9
1	B	2442	PHE	3.8
1	B	2390	VAL	3.8
1	B	2330	VAL	3.7
1	B	2459	ALA	3.7
1	B	2399	LEU	3.6
1	B	2297	ASP	3.5
1	B	2455	ALA	3.5
1	B	2343	MET	3.5
1	B	2322	PRO	3.5
1	B	2249	VAL	3.3
1	B	2367	CYS	3.3
1	B	2475	LYS	3.3
1	B	2401	LEU	3.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	2366	GLN	3.2
1	B	2327	GLU	3.1
1	B	2465	ASP	3.1
1	B	2477	ASP	3.1
1	B	2347	ALA	3.1
1	B	2423	LEU	3.1
1	B	2300	ILE	3.0
1	B	2444	LEU	2.9
1	B	2240	GLN	2.8
1	B	2298	ALA	2.8
1	B	2362	TYR	2.8
1	B	2340	ARG	2.8
1	B	2457	ALA	2.7
1	B	2243	PHE	2.7
1	B	2287	ILE	2.6
1	B	2279	HIS	2.6
1	B	2216	TYR	2.6
1	A	2468	ALA	2.6
1	B	2355	ALA	2.6
1	B	2458	TRP	2.6
1	B	2381	SER	2.5
1	B	2373	ARG	2.5
1	B	2326	GLU	2.5
1	B	2329	LEU	2.5
1	A	2274	VAL	2.5
1	B	2215	ALA	2.5
1	B	2291	LYS	2.5
1	B	2494	PHE	2.4
1	A	2216	TYR	2.4
1	B	2221	LEU	2.4
1	B	2464	SER	2.4
1	B	2289	TRP	2.3
1	B	2294	ILE	2.3
1	B	2214	PRO	2.3
1	B	2293	VAL	2.3
1	B	2285	LYS	2.2
1	B	2451	SER	2.2
1	B	2394	GLN	2.2
1	B	2273	GLY	2.2
1	B	2344	GLU	2.2
1	B	2406	THR	2.2
1	B	2491	MET	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	2307	TYR	2.1
1	B	2462	ARG	2.1
1	B	2378	TYR	2.1
1	B	2359	GLU	2.1
1	B	2385	ILE	2.0

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

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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^{th} 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	MG	A	2601	1/1	0.84	0.58	68,68,68,68	0

6.5 Other polymers ⓘ

There are no such residues in this entry.