Full wwPDB X-ray Structure Validation Report (i)

Jan 5, 2026 – 03:38 pm GMT

PDB ID	:	9HRB / pdb_00009hrb
Title	:	A canonical homodimer of type III polyketide synthase from Aspergillus thesauricus IBT 34227
Authors	:	Zhang, L.; Groves, M.R.
Deposited on	:	2024-12-17
Resolution	:	1.90 Å(reported)

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

|		
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	?.? (???), CSD ??CSD?? (????)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.47

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 1.90 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	Percentile Ranks	Value
Rfree		0.207
Clashscore		4
Ramachandran outliers		0
Sidechain outliers		1.6%
RSRZ outliers		1.0%

Worse ··· *Better*

▮ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution
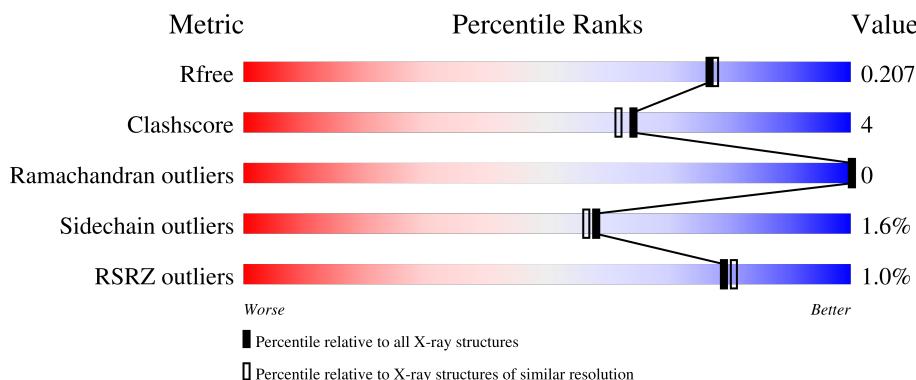
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	164625	7293 (1.90-1.90)
Clashscore	180529	8090 (1.90-1.90)
Ramachandran outliers	177936	8022 (1.90-1.90)
Sidechain outliers	177891	8022 (1.90-1.90)
RSRZ outliers	164620	7292 (1.90-1.90)

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	431	2%　79%　10%　10%
1	B	431	80%　8%　·　11%
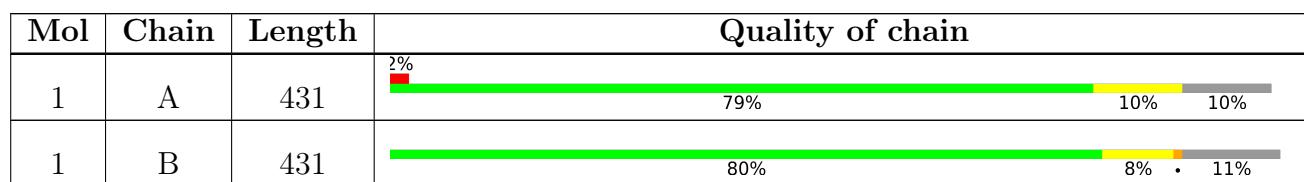
2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 12096 atoms, of which 5994 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 type III polyketide synthase from Aspergillus thesauricus IBT 34227.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	386	Total	C	H	N	O	S	98	5	0
			5931	1843	2996	525	550	17			
1	B	385	Total	C	H	N	O	S	98	4	0
			5919	1841	2988	525	549	16			

- Molecule 2 is DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula: $C_4H_{10}O_3$).
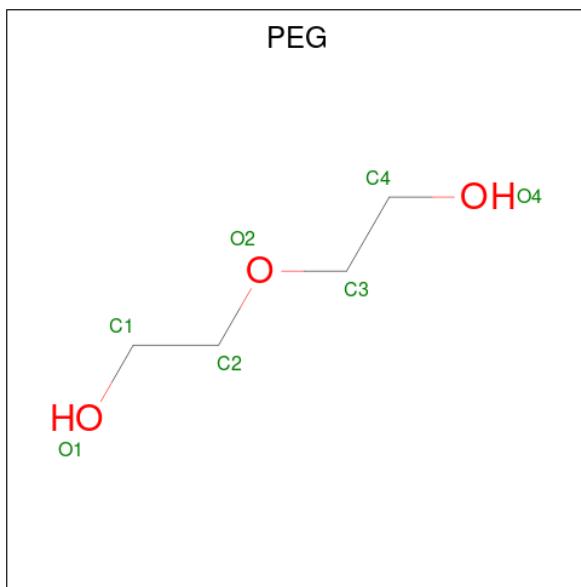
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	H	O	2	0
			17	4	10	3		

- Molecule 3 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Mg	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	109	Total 109	O 109	0	0
4	B	119	Total 119	O 119	0	0

3 Residue-property plots (i)

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: type III polyketide synthase from Aspergillus thesauricus IBT 34227
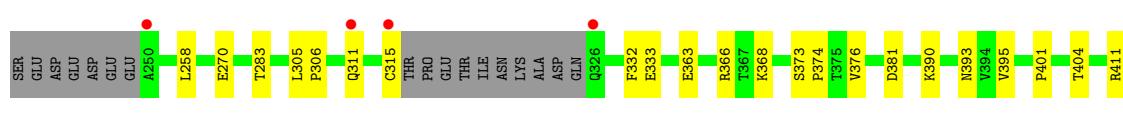
- Molecule 1: type III polyketide synthase from Aspergillus thesauricus IBT 34227
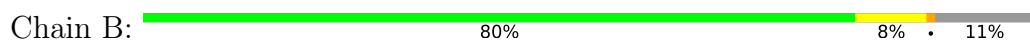
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	71.39Å 89.97Å 141.24Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.03 – 1.90 45.03 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.8 (45.03-1.90) 99.8 (45.03-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.34 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, R_{free}	0.170 , 0.205 0.172 , 0.207	Depositor DCC
R_{free} test set	3657 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	31.7	Xtriage
Anisotropy	0.061	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 27.2	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	12096	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.47% 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: PEG, CSD, 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	0.68	0/2997	1.12	5/4068 (0.1%)
1	B	0.66	0/2990	1.12	7/4058 (0.2%)
All	All	0.67	0/5987	1.12	12/8126 (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
All	All	0	3

There are no bond length outliers.

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	218	ASP	CA-CB-CG	6.56	119.16	112.60
1	B	404	THR	CA-CB-OG1	-6.27	100.20	109.60
1	B	205	ASP	CA-CB-CG	6.20	118.80	112.60
1	B	64	THR	N-CA-CB	-5.82	102.03	110.70
1	B	395	VAL	N-CA-CB	5.67	118.96	111.25
1	A	404	THR	CA-CB-OG1	-5.58	101.23	109.60
1	A	270	GLU	CB-CG-CD	5.38	121.75	112.60
1	B	199	GLU	CB-CA-C	-5.11	102.83	110.90
1	A	74	THR	CA-CB-OG1	-5.10	101.94	109.60
1	A	233	GLU	N-CA-CB	5.06	118.01	110.22
1	B	233	GLU	N-CA-CB	5.05	117.55	110.12
1	B	286	THR	CA-CB-OG1	-5.01	102.08	109.60

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	38	ARG	Sidechain
1	B	175	ARG	Sidechain
1	B	56	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	2935	2996	2971	25	0
1	B	2931	2988	2972	25	0
2	A	7	10	10	0	0
3	B	1	0	0	0	0
4	A	109	0	0	2	0
4	B	119	0	0	4	0
All	All	6102	5994	5953	45	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 4.

All (45) 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:393:ASN:HD21	1:B:411:ARG:HH21	1.33	0.75
1:B:199:GLU:OE1	1:B:282:HIS:HD2	1.72	0.73
1:B:41:LEU:O	1:B:45:ILE:HD13	1.92	0.69
1:A:376[B]:VAL:HG23	4:A:685:HOH:O	1.96	0.64
1:B:14:HIS:HD2	1:B:381:ASP:OD2	1.82	0.61
1:A:174:ALA:HB2	1:B:174:ALA:HB2	1.82	0.61
1:A:27:ASP:OD2	1:A:49:HIS:HE1	1.85	0.59
1:A:401:PRO:HG2	1:B:121:CYS:HB3	1.85	0.59
1:B:27:ASP:OD2	1:B:49:HIS:HE1	1.86	0.59
1:A:14:HIS:HE1	4:A:664:HOH:O	1.85	0.59
1:A:145:LEU:O	1:B:262:HIS:HE1	1.87	0.57
1:B:282:HIS:HE1	4:B:644:HOH:O	1.91	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:376[A]:VAL:HG22	4:B:679:HOH:O	2.09	0.52
1:B:14:HIS:HE1	4:B:626:HOH:O	1.93	0.51
1:A:393:ASN:HD21	1:A:411:ARG:HH11	1.57	0.51
1:B:393:ASN:HD21	1:B:411:ARG:NH2	2.07	0.50
1:B:393:ASN:ND2	1:B:411:ARG:HH21	2.07	0.49
1:A:130:LYS:HD2	1:B:269:ILE:HD11	1.94	0.49
1:B:264:VAL:HG22	1:B:404:THR:HG22	1.95	0.48
1:B:54:GLU:OE1	1:B:366:ARG:NH1	2.45	0.47
1:A:16:TYR:CD2	1:A:368:LYS:HE2	2.50	0.46
1:A:212:VAL:HG21	1:A:283:THR:HG21	1.97	0.46
1:A:54:GLU:OE1	1:A:366:ARG:NH1	2.44	0.46
1:B:110:HIS:HB2	1:B:181:ILE:HG12	1.98	0.45
1:A:190:MET:N	1:A:191:PRO:CD	2.79	0.45
1:A:258:LEU:O	1:A:311:GLN:OE1	2.33	0.45
1:A:121:CYS:HB3	1:B:401:PRO:HG2	1.99	0.45
1:B:34:HIS:HE1	4:B:668:HOH:O	1.99	0.44
1:A:127[A]:LEU:HD22	1:A:127[A]:LEU:N	2.33	0.44
1:A:14:HIS:HD2	1:A:381:ASP:OD2	2.01	0.43
1:A:332:PHE:CD2	1:A:395:VAL:HG23	2.53	0.43
1:A:373:SER:N	1:A:374:PRO:HD2	2.34	0.43
1:A:209:ILE:O	1:A:209:ILE:HG13	2.19	0.43
1:A:333:GLU:OE2	1:A:390:LYS:HB2	2.19	0.43
1:B:187:GLU:HG3	1:B:372:SER:HB3	2.01	0.43
1:B:363:GLU:CD	1:B:366:ARG:HH21	2.27	0.42
1:B:266:PRO:O	1:B:268:THR:HG23	2.19	0.42
1:B:271:HIS:HD2	1:B:286:THR:OG1	2.03	0.41
1:A:17:PRO:HB3	1:A:59:ILE:HD12	2.01	0.41
1:A:305:LEU:N	1:A:306:PRO:CD	2.83	0.41
1:A:17:PRO:HB2	1:A:59:ILE:HA	2.03	0.41
1:B:402:GLY:N	1:B:403:MET:HA	2.35	0.41
1:A:363:GLU:CD	1:A:366:ARG:HH21	2.29	0.40
1:A:4:GLN:HA	1:A:5:PRO:HD3	1.99	0.40
1:B:199:GLU:OE1	1:B:282:HIS:CD2	2.62	0.40

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	384/431 (89%)	376 (98%)	8 (2%)	0	100	100
1	B	382/431 (89%)	378 (99%)	4 (1%)	0	100	100
All	All	766/862 (89%)	754 (98%)	12 (2%)	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	315/347 (91%)	311 (99%)	4 (1%)	65	65
1	B	313/347 (90%)	306 (98%)	7 (2%)	47	43
All	All	628/694 (90%)	617 (98%)	11 (2%)	58	52

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	45	ILE
1	A	59	ILE
1	A	102	ARG
1	A	315	CYS
1	B	59	ILE
1	B	64	THR
1	B	119	GLN
1	B	304	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	376[A]	VAL
1	B	376[B]	VAL
1	B	401	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (21) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	14	HIS
1	A	34	HIS
1	A	46	GLN
1	A	49	HIS
1	A	80	HIS
1	A	118	ASN
1	A	119	GLN
1	A	167	GLN
1	A	271	HIS
1	A	292	HIS
1	A	313	GLN
1	A	393	ASN
1	B	4	GLN
1	B	14	HIS
1	B	34	HIS
1	B	49	HIS
1	B	118	ASN
1	B	262	HIS
1	B	271	HIS
1	B	282	HIS
1	B	393	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues are modelled in this entry.

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.

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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	385/431 (89%)	-0.34	7 (1%)	67	70	15, 34, 61, 86	4 (1%)
1	B	384/431 (89%)	-0.43	1 (0%)	90	91	14, 34, 56, 84	4 (1%)
All	All	769/862 (89%)	-0.39	8 (1%)	79	81	14, 34, 59, 86	8 (1%)

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	250	ALA	2.7
1	A	315	CYS	2.5
1	A	235	ARG	2.3
1	B	80[A]	HIS	2.3
1	A	80	HIS	2.2
1	A	311	GLN	2.1
1	A	326	GLN	2.1
1	A	210	SER	2.0

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

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
1	CSD	B	155	8/9	0.97	0.07	28,31,44,50	1
1	CSD	A	155	8/9	0.98	0.06	27,29,38,42	1

6.3 Carbohydrates ⓘ

There are no oligosaccharides 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	PEG	A	500	7/7	0.87	0.16	54,63,68,68	2
3	MG	B	501	1/1	0.98	0.04	29,29,29,29	0

6.5 Other polymers (i)

There are no such residues in this entry.