Full wwPDB X-ray Structure Validation Report (i)

Mar 5, 2026 – 07:38 AM UTC

PDB ID	:	1VIA / pdb_00001via
Title	:	Crystal structure of shikimate kinase
Authors	:	Structural GenomiX
Deposited on	:	2003-12-01
Resolution	:	1.57 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

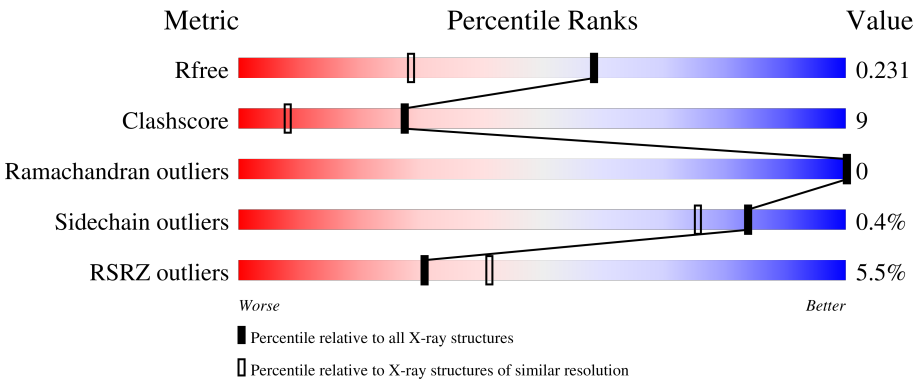
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

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 1.57 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}	180053	1094 (1.58-1.58)
Clashscore	190562	1105 (1.58-1.58)
Ramachandran outliers	187476	1082 (1.58-1.58)
Sidechain outliers	187428	1081 (1.58-1.58)
RSRZ outliers	180081	1094 (1.58-1.58)

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	175	6% 72% 19% • 8%
1	B	175	4% 78% 10% 12%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2935 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 shikimate kinase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	161	1321	851	215	248	4	3	0	5	0
1	B	154	1249	804	206	234	3	2	0	2	0

There are 30 discrepancies between the modelled and reference sequences:

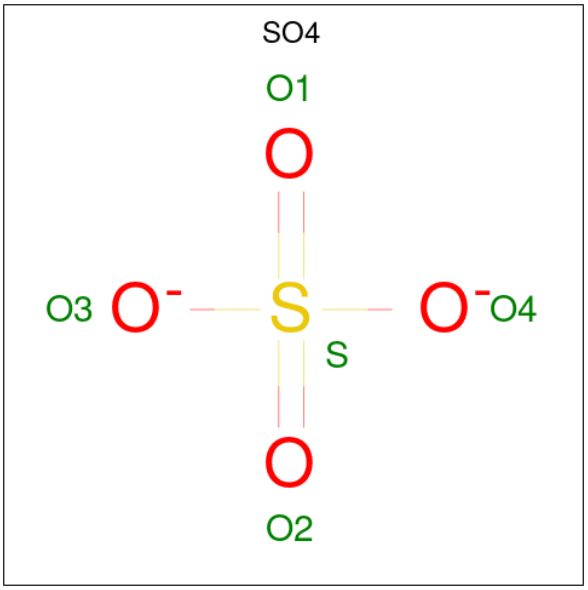
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	-	cloning artifact	UNP Q9PIB5
A	2	SER	-	cloning artifact	UNP Q9PIB5
A	3	LEU	-	cloning artifact	UNP Q9PIB5
A	13	MSE	MET	modified residue	UNP Q9PIB5
A	66	MSE	MET	modified residue	UNP Q9PIB5
A	166	GLU	-	cloning artifact	UNP Q9PIB5
A	167	GLY	-	cloning artifact	UNP Q9PIB5
A	168	GLY	-	cloning artifact	UNP Q9PIB5
A	169	SER	-	cloning artifact	UNP Q9PIB5
A	170	HIS	-	cloning artifact	UNP Q9PIB5
A	171	HIS	-	cloning artifact	UNP Q9PIB5
A	172	HIS	-	cloning artifact	UNP Q9PIB5
A	173	HIS	-	cloning artifact	UNP Q9PIB5
A	174	HIS	-	cloning artifact	UNP Q9PIB5
A	175	HIS	-	cloning artifact	UNP Q9PIB5
B	1	MSE	-	cloning artifact	UNP Q9PIB5
B	2	SER	-	cloning artifact	UNP Q9PIB5
B	3	LEU	-	cloning artifact	UNP Q9PIB5
B	13	MSE	MET	modified residue	UNP Q9PIB5
B	66	MSE	MET	modified residue	UNP Q9PIB5
B	166	GLU	-	cloning artifact	UNP Q9PIB5
B	167	GLY	-	cloning artifact	UNP Q9PIB5
B	168	GLY	-	cloning artifact	UNP Q9PIB5
B	169	SER	-	cloning artifact	UNP Q9PIB5
B	170	HIS	-	cloning artifact	UNP Q9PIB5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	171	HIS	-	cloning artifact	UNP Q9PIB5
B	172	HIS	-	cloning artifact	UNP Q9PIB5
B	173	HIS	-	cloning artifact	UNP Q9PIB5
B	174	HIS	-	cloning artifact	UNP Q9PIB5
B	175	HIS	-	cloning artifact	UNP Q9PIB5

- Molecule 2 is SULFATE ION (CCD ID: SO4) (formula: O_4S).



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

- Molecule 3 is water.

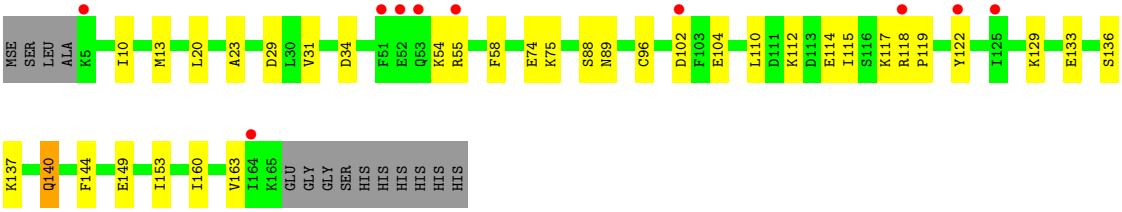
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	182	182	182	0	0
			Total	O		
3	B	168	168	168	0	0

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.

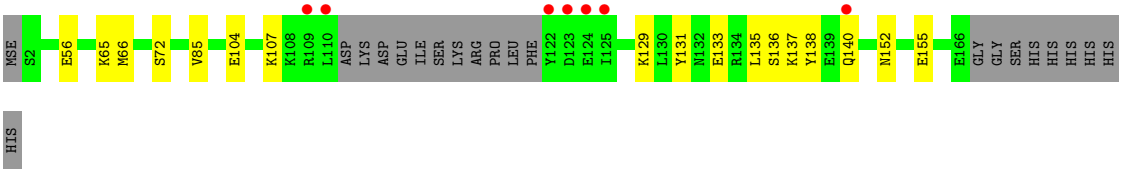
- Molecule 1: shikimate kinase

Chain A:



- Molecule 1: shikimate kinase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	41.93Å 43.23Å 52.61Å 81.51° 76.58° 77.87°	Depositor
Resolution (Å)	24.56 – 1.57 24.56 – 1.57	Depositor EDS
% Data completeness (in resolution range)	(Not available) (24.56-1.57) 99.4 (24.56-1.57)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.14 (at 1.57Å)	Xtriage
Refinement program	REFMAC 4.0	Depositor
R, R_{free}	0.198 , 0.228 (Not available) , 0.231	Depositor DCC
R_{free} test set	2427 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	17.6	Xtriage
Anisotropy	0.093	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 48.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	2935	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.43% 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: SO4

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	0.74	0/1358	1.13	3/1807 (0.2%)
1	B	0.73	0/1268	1.10	1/1684 (0.1%)
All	All	0.74	0/2626	1.11	4/3491 (0.1%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	88[A]	SER	N-CA-C	6.29	119.12	111.82
1	A	88[B]	SER	N-CA-C	6.29	119.12	111.82
1	A	34	ASP	CA-CB-CG	5.80	118.40	112.60
1	B	56	GLU	N-CA-C	5.38	116.83	111.07

There are no chirality outliers.

There are no planarity outliers.

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	1321	0	1327	36	0
1	B	1249	0	1265	16	0
2	A	5	0	0	0	0
2	B	10	0	0	0	0
3	A	182	0	0	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	168	0	0	1	0
All	All	2935	0	2592	47	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 9.

All (47) 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:A:31:VAL:HG22	1:B:66:MSE:HE1	1.25	1.12
1:B:85[A]:VAL:HG11	1:B:138:TYR:CD1	2.10	0.86
1:A:114:GLU:HB3	1:A:118:ARG:HH12	1.39	0.85
1:A:13[B]:MSE:HE1	1:A:117:LYS:HB2	1.61	0.82
1:A:31:VAL:HG22	1:B:66:MSE:CE	2.11	0.79
1:A:114:GLU:HB3	1:A:118:ARG:NH1	2.00	0.77
1:B:66:MSE:HE3	3:B:300:HOH:O	1.87	0.73
1:B:104:GLU:HA	1:B:107:LYS:HE3	1.74	0.70
1:A:102:ASP:HB3	3:A:343:HOH:O	1.94	0.67
1:A:29:ASP:OD2	1:B:65:LYS:HE3	1.99	0.61
1:A:74:GLU:HG2	1:B:72:SER:HB2	1.83	0.60
1:A:13[B]:MSE:HE3	1:A:118:ARG:HD2	1.87	0.57
1:B:85[A]:VAL:HG11	1:B:138:TYR:CE1	2.40	0.56
1:A:13[B]:MSE:SE	1:A:118:ARG:NH1	2.88	0.56
1:B:129:LYS:O	1:B:133:GLU:HG3	2.07	0.54
1:A:118:ARG:HB2	1:A:119:PRO:HD2	1.88	0.54
1:A:13[A]:MSE:HE1	1:A:122:TYR:CD2	2.43	0.54
1:A:10:ILE:HD13	1:A:96[B]:CYS:SG	2.48	0.54
1:A:74:GLU:HG2	1:B:72:SER:CB	2.39	0.53
1:A:137:LYS:HA	1:A:140:GLN:HG2	1.90	0.53
1:A:112:LYS:HA	1:A:115:ILE:HD13	1.91	0.52
1:A:54:LYS:O	1:A:55:ARG:HB2	2.10	0.52
1:A:23:ALA:HB2	1:A:153:ILE:HD11	1.92	0.51
1:A:20:LEU:HD21	1:A:160:ILE:HD11	1.93	0.50
1:A:75:LYS:HE2	3:A:313:HOH:O	2.11	0.50
1:A:112:LYS:O	1:A:115:ILE:HD13	2.12	0.48
1:B:85[A]:VAL:HG11	1:B:138:TYR:HD1	1.69	0.47
1:A:23:ALA:CB	1:A:153:ILE:HD11	2.45	0.47
1:B:152:ASN:OD1	1:B:155:GLU:HG3	2.14	0.47
1:A:114:GLU:CB	1:A:118:ARG:HH12	2.19	0.46
1:A:110:LEU:HD12	1:A:115:ILE:HD12	1.96	0.46
1:A:118:ARG:HB2	1:A:119:PRO:CD	2.45	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:137:LYS:HA	1:B:140:GLN:HG2	1.98	0.45
1:B:131:TYR:CZ	1:B:135:LEU:HD21	2.52	0.45
1:A:13[B]:MSE:CE	1:A:118:ARG:HD2	2.48	0.44
1:A:13[A]:MSE:HE3	1:A:118:ARG:NH2	2.34	0.43
1:A:54:LYS:HD2	1:A:58:PHE:CE1	2.54	0.42
1:A:136:SER:O	1:A:140:GLN:HG2	2.19	0.42
1:A:102:ASP:OD2	1:A:149:GLU:OE2	2.38	0.42
1:A:129:LYS:O	1:A:133:GLU:HG3	2.20	0.41
1:A:140:GLN:HE21	1:A:140:GLN:HB2	1.67	0.41
1:B:136:SER:O	1:B:140:GLN:HG2	2.20	0.41
1:A:144:PHE:CZ	1:A:163:VAL:HG21	2.55	0.41
1:A:89:ASN:HB2	3:A:242:HOH:O	2.20	0.41
1:A:104:GLU:HG3	3:A:202:HOH:O	2.20	0.40
1:B:85[A]:VAL:CG1	1:B:138:TYR:CE1	3.04	0.40
1:A:118:ARG:CB	1:A:119:PRO:CD	2.99	0.40

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	163/175 (93%)	159 (98%)	4 (2%)	0	100	100
1	B	151/175 (86%)	149 (99%)	2 (1%)	0	100	100
All	All	314/350 (90%)	308 (98%)	6 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	145/152 (95%)	144 (99%)	1 (1%)	76 61
1	B	135/152 (89%)	135 (100%)	0	100 100
All	All	280/304 (92%)	279 (100%)	1 (0%)	84 75

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

Mol	Chain	Res	Type
1	A	140	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	89	ASN
1	A	140	GLN
1	B	86	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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

3 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	SO4	B	176	-	4,4,4	0.73	0	6,6,6	0.58	0
2	SO4	A	176	-	4,4,4	0.70	0	6,6,6	0.13	0
2	SO4	B	177	-	4,4,4	0.65	0	6,6,6	0.12	0

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 (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)

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	159/175 (90%)	0.23	10 (6%)	26	35	11, 22, 40, 59	4 (2%)
1	B	152/175 (86%)	0.14	7 (4%)	37	49	12, 20, 36, 59	2 (1%)
All	All	311/350 (88%)	0.19	17 (5%)	30	41	11, 21, 39, 59	6 (1%)

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	110	LEU	4.8
1	B	122	TYR	4.8
1	A	55	ARG	2.9
1	A	5	LYS	2.7
1	B	123	ASP	2.6
1	A	164	ILE	2.6
1	B	140	GLN	2.6
1	A	102	ASP	2.6
1	A	53	GLN	2.5
1	A	122	TYR	2.5
1	A	125	ILE	2.5
1	B	109	ARG	2.3
1	B	124	GLU	2.2
1	A	118	ARG	2.1
1	B	125	ILE	2.1
1	A	52	GLU	2.0
1	A	51	PHE	2.0

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 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	SO4	A	176	5/5	0.86	0.14	57,57,58,58	0
2	SO4	B	177	5/5	0.86	0.13	56,56,57,57	0
2	SO4	B	176	5/5	0.96	0.08	28,28,30,32	0

6.5 Other polymers (i)

There are no such residues in this entry.