W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 7, 2022 – 04:31 PM EDT

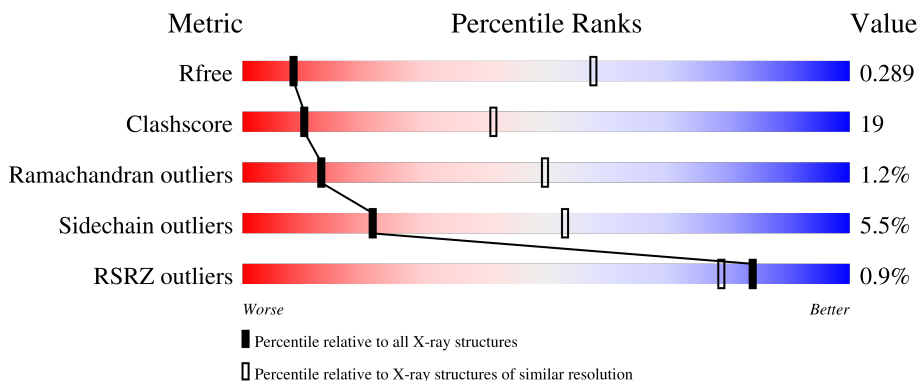|		
PDB ID	:	1Z1B
Title	:	Crystal structure of a lambda integrase dimer bound to a COC' core site
Authors	:	Biswas, T.; Aihara, H.; Radman-Livaja, M.; Filman, D.; Landy, A.; Ellenberger, T.
Deposited on	:	2005-03-03
Resolution	:	3.80 Å(reported)

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 3.80 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.
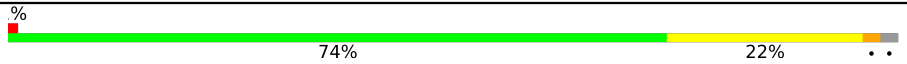
Metric	Percentile Ranks	Value
Rfree		0.289
Clashscore		19
Ramachandran outliers		1.2%
Sidechain outliers		5.5%
RSRZ outliers		0.9%

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$_{free}$	130704	1212 (4.00-3.60)
Clashscore	141614	1288 (4.00-3.60)
Ramachandran outliers	138981	1243 (4.00-3.60)
Sidechain outliers	138945	1237 (4.00-3.60)
RSRZ outliers	127900	1121 (4.00-3.60)

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	C	14	43% 36% 21%
2	D	18	6% 50% 44%
3	E	29	24% 38% 38%
4	F	26	27% 46% 23% •
5	G	26	19% 65% 12% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	A	356	% 74% 22%
6	B	356	% 70% 26%

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 7653 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 DNA chain called 5'-D(*CP*T*CP*GP*TP*TP*CP*AP*GP*CP*TP*TP* TP*TP*TP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	11	Total 219	C 107	N 34	O 68	P 10	0	0	0

- Molecule 2 is a DNA chain called 5'-D(*TP*TP*TP*AP*TP*AP*CP*TP*AP*AP*GP*TP *TP*GP*GP*CP*AP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	18	Total 366	C 178	N 62	O 109	P 17	0	0	0

- Molecule 3 is a DNA chain called 29-MER DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	29	Total 595	C 285	N 117	O 165	P 28	0	0	0

- Molecule 4 is a DNA chain called 26-MER DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	25	Total 506	C 243	N 96	O 143	P 24	0	0	0

- Molecule 5 is a DNA chain called 26-MER DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	G	25	Total 513	C 247	N 89	O 153	P 24	0	0	0

- Molecule 6 is a protein called Integrase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
6	A	349	Total 2741	C 1718	N 498	O 514	S 11		0	0	0
6	B	350	Total 2713	C 1700	N 491	O 510	P 1	S 11	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

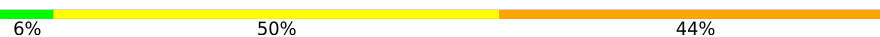| Chain	Residue	Modelled	Actual	Comment	Reference
A	174	LYS	GLU	engineered mutation	UNP P03700
A	342	PTR	TYR	modified residue	UNP P03700
B	174	LYS	GLU	engineered mutation	UNP P03700
B	342	PTR	TYR	modified residue	UNP P03700

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: 5'-D(*CP*T*CP*GP*TP*TP*CP*AP*GP*CP*TP*TP*TP*TP*TP*T)-3'
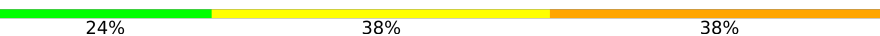
Chain C:
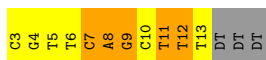
- Molecule 2: 5'-D(*TP*TP*TP*AP*TP*AP*CP*TP*AP*AP*GP*TP*TP*GP*GP*CP*AP*T)-3'
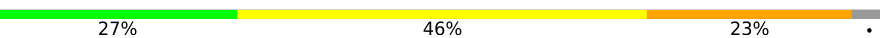
Chain D:
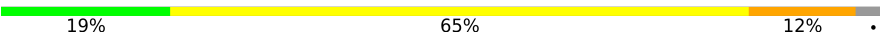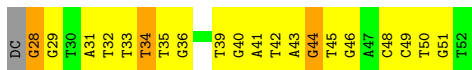

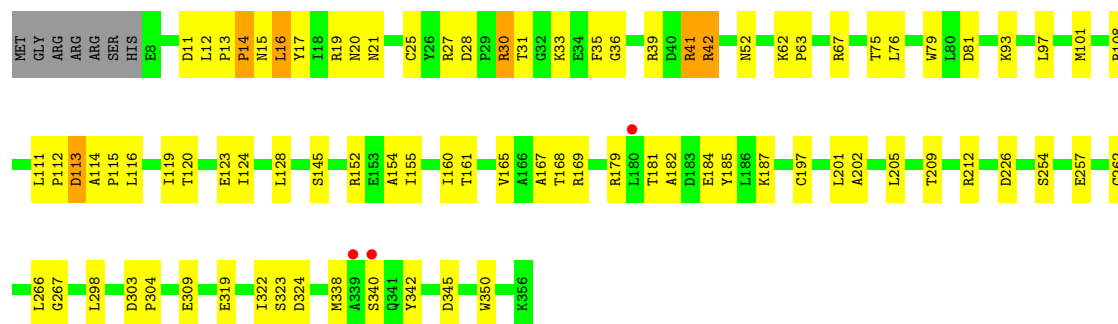- Molecule 3: 29-MER DNA

Chain E:

- Molecule 4: 26-MER DNA

Chain F:

- Molecule 5: 26-MER DNA

Chain G:

- Molecule 6: Integrase

Chain A:

• Molecule 6: Integrase

Chain B:
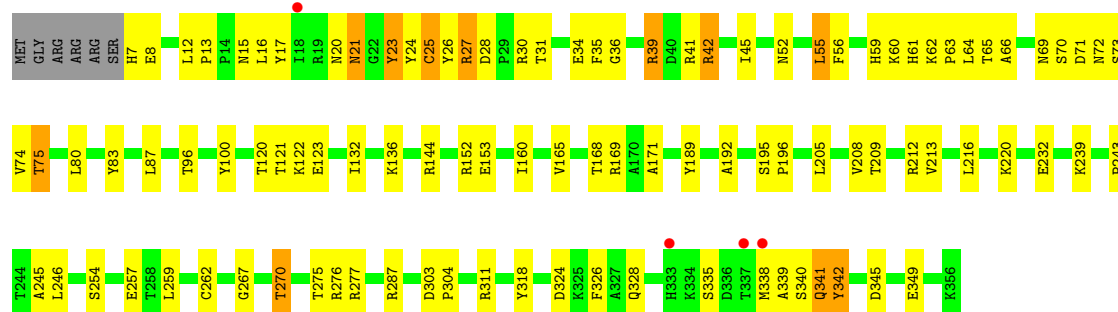
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	97.60Å 211.28Å 191.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.80 36.71 – 3.65	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-3.80) 99.0 (36.71-3.65)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.88 (at 3.66Å)	Xtriage
Refinement program	REFMAC 5.1	Depositor
R, R_{free}	0.248 , 0.296 0.245 , 0.289	Depositor DCC
R_{free} test set	1081 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å²)	140.7	Xtriage
Anisotropy	0.289	Xtriage
Bulk solvent k_{sol}(e/Å³), B_{sol}(Å²)	0.25 , 56.1	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	0.93	EDS
Total number of atoms	7653	wwPDB-VP
Average B, all atoms (Å²)	106.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.34% 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: PTR

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	C	1.17	0/243	1.89	6/373 (1.6%)
2	D	1.26	4/409 (1.0%)	1.86	13/630 (2.1%)
3	E	1.27	5/670 (0.7%)	1.78	19/1032 (1.8%)
4	F	0.71	0/568	1.40	7/873 (0.8%)
5	G	0.77	0/574	1.48	4/886 (0.5%)
6	A	0.46	1/2775 (0.0%)	0.69	1/3737 (0.0%)
6	B	0.51	1/2744 (0.0%)	0.74	2/3701 (0.1%)
All	All	0.71	11/7983 (0.1%)	1.13	52/11232 (0.5%)

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	49	DA	O3'-P	10.85	1.74	1.61
2	D	15	DT	O3'-P	6.65	1.69	1.61
6	B	55	LEU	C-N	6.45	1.48	1.34
3	E	49	DA	N9-C4	-6.01	1.34	1.37
6	A	14	PRO	N-CD	-5.75	1.39	1.47

The worst 5 of 52 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	8	DA	O4'-C1'-N9	-9.91	101.06	108.00
3	E	48	DT	P-O3'-C3'	9.85	131.52	119.70
5	G	44	DG	O4'-C1'-N9	-9.05	101.67	108.00
4	F	13	DC	O4'-C1'-N1	-8.89	101.77	108.00
5	G	34	DT	O4'-C1'-N1	-8.23	102.23	108.00

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	C	219	0	127	13	0
2	D	366	0	208	30	0
3	E	595	0	327	35	2
4	F	506	0	282	23	0
5	G	513	0	287	36	0
6	A	2741	0	2754	70	2
6	B	2713	0	2679	85	0
All	All	7653	0	6664	271	2

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 19.

The worst 5 of 271 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)
6:A:15:ASN:HD22	6:A:52:ASN:ND2	1.02	1.44
6:A:15:ASN:ND2	6:A:52:ASN:HD21	1.19	1.40
6:A:15:ASN:ND2	6:A:52:ASN:ND2	1.72	1.25
5:G:44:DG:H2"	5:G:45:DT:H5'	1.18	1.17
5:G:33:DT:H2"	5:G:34:DT:H5"	1.33	1.10

All (2) 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)
3:E:47:DG:OP2	6:A:342:PTR:OH[4_565]	2.12	0.08
3:E:47:DG:OP1	6:A:212:ARG:NE[4_565]	2.15	0.05

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	
6	A	346/356 (97%)	324 (94%)	20 (6%)	2 (1%)	25	62
6	B	347/356 (98%)	327 (94%)	14 (4%)	6 (2%)	9	43
All	All	693/712 (97%)	651 (94%)	34 (5%)	8 (1%)	13	50

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
6	B	59	HIS
6	B	71	ASP
6	A	20	ASN
6	A	267	GLY
6	B	267	GLY

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	
6	A	286/299 (96%)	275 (96%)	11 (4%)	33	61
6	B	275/299 (92%)	255 (93%)	20 (7%)	14	45
All	All	561/598 (94%)	530 (94%)	31 (6%)	21	53

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	B	27	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
6	B	169	ARG
6	B	41	ARG
6	B	324	ASP
6	B	75	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
6	B	20	ASN
6	B	21	ASN
6	B	52	ASN
6	A	333	HIS
6	A	52	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.

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	$\#|Z| > 2$
6	PTR	B	342	1,6	11,15,17	2.31	1 (9%)	12,19,24	0.27	0
6	PTR	A	342	6	11,12,17	0.46	0	12,15,24	0.61	0

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
6	PTR	B	342	1,6	-	2/7/10/13	0/1/1/1
6	PTR	A	342	6	-	0/5/6/13	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	342	PTR	OH-CZ	-7.48	1.29	1.40

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	342	PTR	N-CA-CB-CG
6	B	342	PTR	C-CA-CB-CG

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	342	PTR	2	0
6	A	342	PTR	1	1

5.5   Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6   Ligand geometry ⓘ

There are no ligands in this entry.

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	C	11/14 (78%)	-0.75	0	100	100	49, 66, 118, 135	0
2	D	18/18 (100%)	-0.65	0	100	100	40, 87, 177, 181	0
3	E	29/29 (100%)	-0.54	0	100	100	46, 80, 155, 180	0
4	F	25/26 (96%)	-0.14	0	100	100	90, 139, 174, 181	0
5	G	25/26 (96%)	-0.18	0	100	100	81, 135, 178, 182	0
6	A	348/356 (97%)	-0.29	3 (0%)	84	79	46, 101, 159, 175	0
6	B	349/356 (98%)	-0.23	4 (1%)	80	74	44, 96, 159, 175	0
All	All	805/825 (97%)	-0.28	7 (0%)	84	79	40, 100, 161, 182	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
6	B	337	THR	5.1
6	B	338	MET	3.5
6	A	339	ALA	3.1
6	B	333	HIS	2.6
6	A	180	LEU	2.3

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
6	PTR	A	342	12/17	0.91	0.43	131,150,152,152	0
6	PTR	B	342	15/17	0.91	0.27	48,62,115,118	0

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.