W O R L D W I D E

wwPDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

May 18, 2020 – 12:36 am BST

|		
PDB ID	:	1BG1
Title	:	TRANSCRIPTION FACTOR STAT3B/DNA COMPLEX
Authors	:	Becker, S.; Groner, B.; Muller, C.W.
Deposited on	:	1998-06-03
Resolution	:	2.25 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 2.25 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.
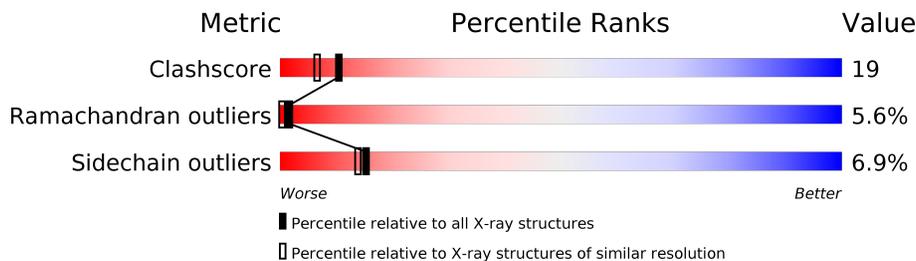
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	B	18	72% 28%
2	A	596	59% 28% 6% 6%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5018 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 DNA (5'-D(*TP*GP*CP*AP*TP*TP*TP*CP*CP*CP* GP*TP*AP*AP*AP*TP*CP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	18	Total 360	C 175	N 59	O 109	P 17	0	0	0

- Molecule 2 is a protein called PROTEIN (TRANSCRIPTION FACTOR STAT3B).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	A	559	Total 4504	C 2869	N 766	O 838	P 1	S 30	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	631	SER	LYS	CONFLICT	UNP P42227
A	705	PTR	TYR	MODIFIED RESIDUE	UNP P42227

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	25	Total 25	O 25	0	0
3	A	129	Total 129	O 129	0	0

3 Residue-property plots ⓘ

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.

Note EDS was not executed.

• Molecule 1: DNA (5'-D(*TP*GP*CP*AP*TP*TP*TP*CP*CP*CP*GP*TP*AP*AP*AP*TP *CP*T)-3')

Chain B:
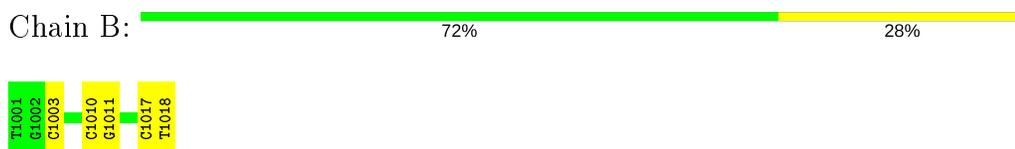
• Molecule 2: PROTEIN (TRANSCRIPTION FACTOR STAT3B)

Chain A:
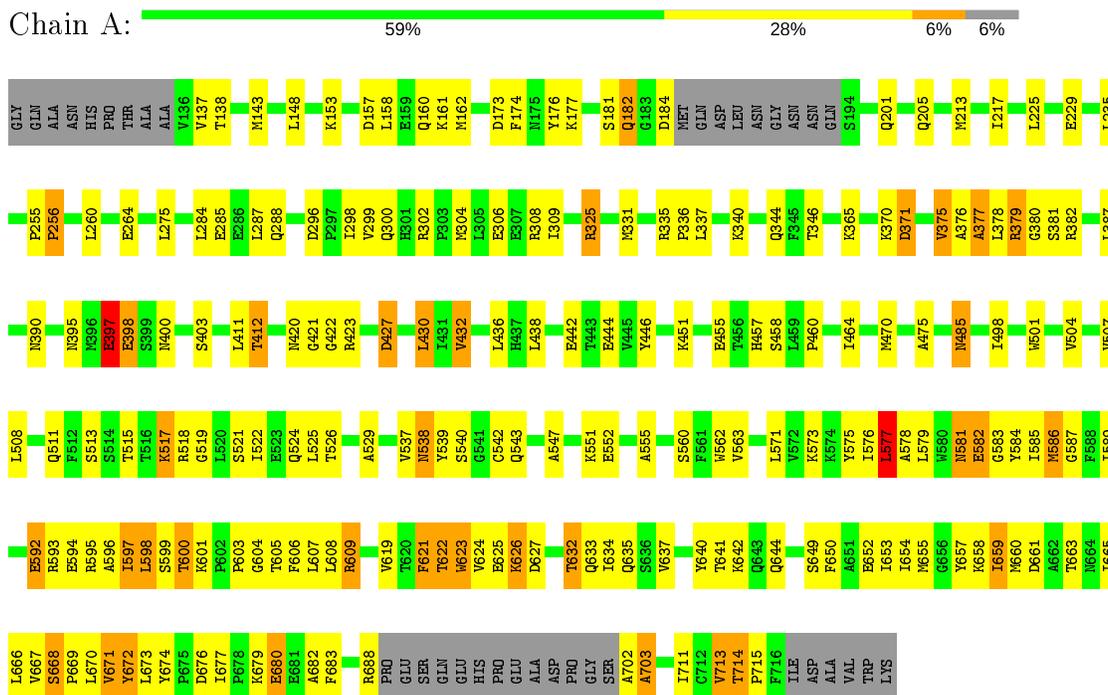
4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	I 41	Depositor
Cell constants a, b, c, α, β, γ	174.00Å 174.00Å 79.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.25	Depositor
% Data completeness (in resolution range)	94.4 (20.00-2.25)	Depositor
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR 3.1	Depositor
R, R_{free}	0.246 , 0.301	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5018	wwPDB-VP
Average B, all atoms (Å^2)	60.0	wwPDB-VP

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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	B	0.82	1/401 (0.2%)	0.92	0/616
2	A	0.62	0/4573	0.78	1/6174 (0.0%)
All	All	0.64	1/4974 (0.0%)	0.80	1/6790 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1003	DC	N1-C2	5.11	1.45	1.40

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	668	SER	N-CA-C	5.27	125.24	111.00

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	B	360	0	207	2	0
2	A	4504	0	4555	177	0
3	A	129	0	0	3	0
3	B	25	0	0	0	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	5018	0	4762	179	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 19.

The worst 5 of 179 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)
2:A:598:LEU:HD13	2:A:624:VAL:HG13	1.32	1.03
2:A:641:THR:HG22	2:A:644:GLN:HG2	1.40	1.02
2:A:597:ILE:HD12	2:A:622:THR:HG21	1.46	0.95
2:A:659:ILE:H	2:A:713:VAL:HG13	1.34	0.93
2:A:598:LEU:HD22	2:A:624:VAL:HG22	1.60	0.83

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
2	A	552/596 (93%)	476 (86%)	45 (8%)	31 (6%)	2 0

5 of 31 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	182	GLN
2	A	256	PRO
2	A	371	ASP
2	A	377	ALA
2	A	555	ALA

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	
2	A	504/533 (95%)	469 (93%)	35 (7%)	15	14

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	420	ASN
2	A	485	ASN
2	A	661	ASP
2	A	427	ASP
2	A	432	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
2	A	390	ASN
2	A	400	ASN
2	A	485	ASN
2	A	301	HIS
2	A	457	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	PTR	A	705	2	15,16,17	1.57	3 (20%)	19,22,24	0.85	1 (5%)

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
2	PTR	A	705	2	-	1/10/11/13	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	705	PTR	CE2-CZ	2.81	1.44	1.38
2	A	705	PTR	CE2-CD2	2.78	1.43	1.38
2	A	705	PTR	CD2-CG	2.02	1.43	1.38

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	705	PTR	O3P-P-OH	2.01	111.54	105.24

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	705	PTR	CA-CB-CG-CD2

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.