Full wwPDB X-ray Structure Validation Report (i)

Mar 6, 2026 – 10:37 AM UTC

PDB ID	:	2G5N / pdb_00002g5n
Title	:	Indole-amidine Complexes with Bovine Trypsin
Authors	:	Kline, A.D.; Briggs, S.L.; Subramaniam, S.
Deposited on	:	2006-02-23
Resolution	:	1.51 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 (1)) were used in the production of this report:

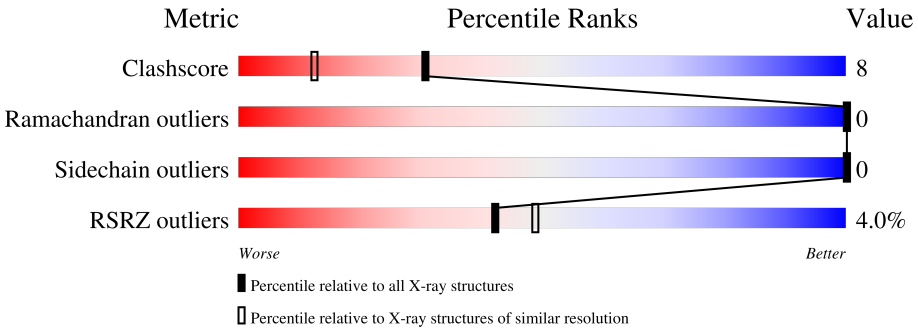
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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 1.51 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	190562	6116 (1.54-1.50)
Ramachandran outliers	187476	6002 (1.54-1.50)
Sidechain outliers	187428	5999 (1.54-1.50)
RSRZ outliers	180081	5891 (1.54-1.50)

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	223	4% 82% 17%

2 Entry composition

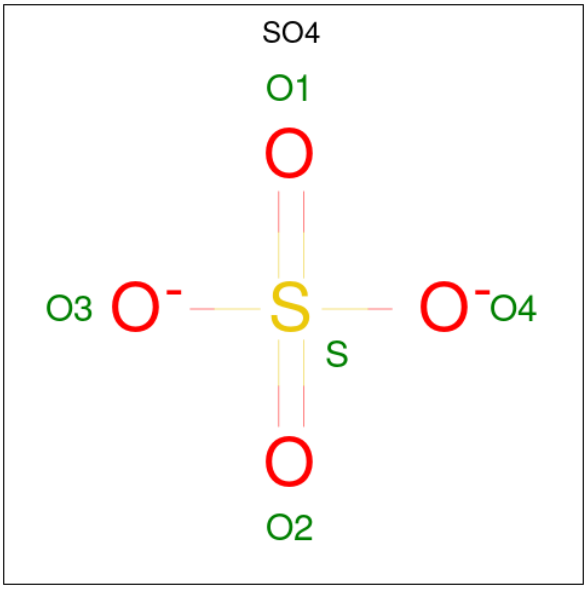
There are 5 unique types of molecules in this entry. The entry contains 1918 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Cationic trypsin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	223	1632	1015	279	324	14	0	1	0

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



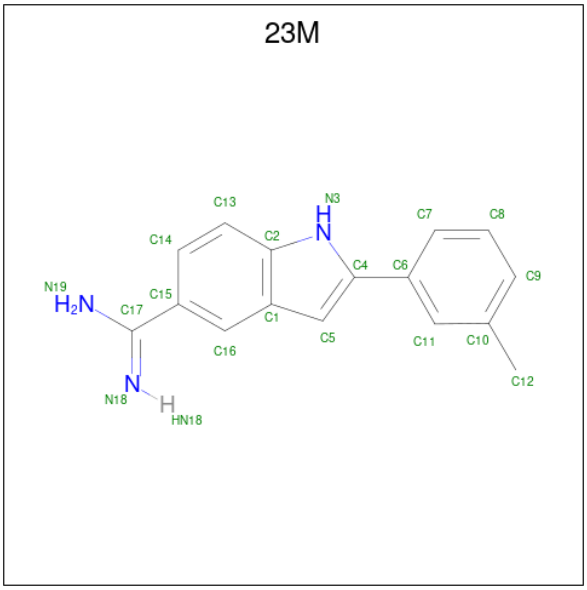
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
2	A	1	5	4	1	0	0

- Molecule 3 is CALCIUM ION (CCD ID: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
3	A	1	1	1	0	0

- Molecule 4 is 2-(3-METHYLPHENYL)-1H-INDOLE-5-CARBOXIMIDAMIDE (CCD ID:

23M) (formula: $C_{16}H_{15}N_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N 19 16 3	0	0

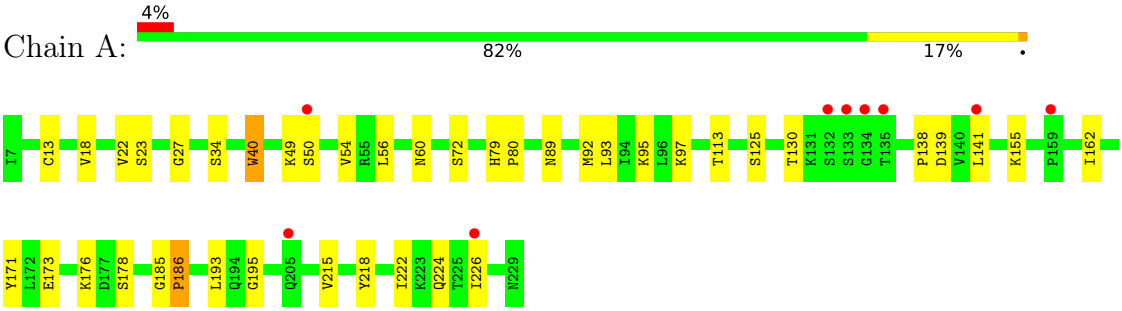
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	261	Total O 261 261	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: Cationic trypsin



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	54.34Å 54.34Å 107.66Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	43.12 – 1.51 43.12 – 1.51	Depositor EDS
% Data completeness (in resolution range)	98.8 (43.12-1.51) 99.0 (43.12-1.51)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.30 (at 1.51Å)	Xtriage
Refinement program	CNX 2002	Depositor
R, R_{free}	0.187 , 0.204 0.188 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	14.6	Xtriage
Anisotropy	0.041	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 37.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.038 for -h,-k,l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	1918	wwPDB-VP
Average B, all atoms (Å^2)	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.09% 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: SO4, CA, 23M

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	0.33	1/1668 (0.1%)	0.92	9/2261 (0.4%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	92	MET	SD-CE	-5.83	1.65	1.79

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	185	GLY	CA-C-N	5.93	126.43	120.14
1	A	185	GLY	C-N-CA	5.93	126.43	120.14
1	A	18	VAL	CA-C-N	5.67	125.35	119.56
1	A	18	VAL	C-N-CA	5.67	125.35	119.56
1	A	113	THR	N-CA-C	-5.60	106.49	113.38
1	A	186	PRO	N-CA-C	5.52	120.90	111.68
1	A	40	TRP	N-CA-C	5.16	117.86	109.24
1	A	27	GLY	N-CA-C	-5.11	107.49	114.64
1	A	178	SER	N-CA-C	-5.08	102.48	110.36

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	1632	0	1595	27	0
2	A	5	0	0	0	0
3	A	1	0	0	0	0
4	A	19	0	15	0	0
5	A	261	0	0	6	0
All	All	1918	0	1610	27	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 8.

All (27) 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:193:LEU:HG	5:A:477:HOH:O	1.71	0.88
1:A:215:VAL:HG21	5:A:477:HOH:O	1.91	0.70
1:A:22:VAL:HG11	1:A:54:VAL:HG13	1.74	0.69
1:A:13:CYS:SG	1:A:141[B]:LEU:HD13	2.43	0.59
1:A:195:GLY:N	5:A:477:HOH:O	2.36	0.58
1:A:93:LEU:HD13	1:A:226:ILE:CD1	2.35	0.57
1:A:40:TRP:CD2	1:A:226:ILE:HG13	2.43	0.53
1:A:34:SER:OG	1:A:186:PRO:HB3	2.10	0.51
1:A:22:VAL:CG1	1:A:54:VAL:HG13	2.40	0.51
1:A:49:LYS:HG2	1:A:50:SER:O	2.12	0.49
1:A:130:THR:HG23	1:A:138:PRO:HD3	1.95	0.48
1:A:93:LEU:HD22	1:A:226:ILE:HD11	1.95	0.48
1:A:60:ASN:HA	1:A:139:ASP:O	2.14	0.47
1:A:72:SER:HB2	1:A:97:LYS:HD3	1.95	0.47
1:A:155:LYS:HG2	1:A:162:ILE:HB	1.98	0.46
1:A:56:LEU:N	1:A:56:LEU:HD12	2.31	0.46
1:A:171:TYR:HB3	1:A:173:GLU:OE1	2.17	0.45
1:A:222:ILE:O	1:A:226:ILE:HD13	2.17	0.45
1:A:22:VAL:CG1	1:A:23:SER:N	2.79	0.45
1:A:224:GLN:HG3	5:A:504:HOH:O	2.16	0.44
1:A:125:SER:HB2	1:A:141[B]:LEU:HD11	1.99	0.43
1:A:176:LYS:NZ	5:A:407:HOH:O	2.50	0.42
1:A:95:LYS:HE2	5:A:443:HOH:O	2.20	0.42
1:A:49:LYS:HG2	1:A:50:SER:N	2.35	0.42
1:A:40:TRP:CG	1:A:226:ILE:HG13	2.56	0.41
1:A:79:HIS:CG	1:A:80:PRO:HD2	2.55	0.41
1:A:89:ASN:HA	1:A:218:TYR:OH	2.20	0.41

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	222/223 (100%)	217 (98%)	5 (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	185/184 (100%)	185 (100%)	0	100 100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	21	GLN
1	A	67	ASN
1	A	85	ASN
1	A	89	ASN
1	A	121	GLN
1	A	205	GLN
1	A	224	GLN

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)

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 for Mogul analysis.

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	$\#\|Z\| > 2$
4	23M	A	301	-	21,21,21	2.30	14 (66%)	27,30,30	1.77	5 (18%)
2	SO4	A	302	-	4,4,4	0.30	0	6,6,6	0.10	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
4	23M	A	301	-	-	2/8/8/8	0/3/3/3

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	301	23M	C17-N18	3.47	1.42	1.27
4	A	301	23M	C17-N19	3.08	1.41	1.33
4	A	301	23M	C11-C10	2.84	1.43	1.39
4	A	301	23M	C14-C15	2.82	1.43	1.39
4	A	301	23M	C4-N3	-2.75	1.35	1.38
4	A	301	23M	C11-C6	2.71	1.43	1.39
4	A	301	23M	C7-C6	2.68	1.43	1.39

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	301	23M	C16-C15	2.64	1.43	1.39
4	A	301	23M	C6-C4	2.59	1.51	1.47
4	A	301	23M	C2-N3	-2.39	1.34	1.38
4	A	301	23M	C8-C7	2.39	1.43	1.38
4	A	301	23M	C14-C13	2.37	1.42	1.38
4	A	301	23M	C13-C2	2.34	1.43	1.39
4	A	301	23M	C8-C9	2.12	1.42	1.38

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	301	23M	C16-C1-C2	6.70	123.27	119.13
4	A	301	23M	C1-C16-C15	-3.43	115.92	121.32
4	A	301	23M	C13-C2-C1	-2.50	119.86	122.13
4	A	301	23M	C14-C15-C16	2.42	122.04	119.25
4	A	301	23M	C15-C17-N19	2.04	121.11	118.01

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	301	23M	C14-C15-C17-N18
4	A	301	23M	C16-C15-C17-N18

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

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, 95^{th} 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	223/223 (100%)	0.32	9 (4%) 42 49	9, 15, 24, 36	1 (0%)

All (9) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	133	SER	4.6
1	A	132	SER	3.1
1	A	134	GLY	2.8
1	A	135	THR	2.8
1	A	226	ILE	2.4
1	A	205	GLN	2.2
1	A	159	PRO	2.2
1	A	50	SER	2.2
1	A	141[A]	LEU	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 oligosaccharides 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
4	23M	A	301	19/19	0.81	0.14	19,22,29,29	0
2	SO4	A	302	5/5	0.82	0.15	31,33,34,35	0
3	CA	A	303	1/1	1.00	0.06	11,11,11,11	0

6.5 Other polymers (i)

There are no such residues in this entry.