Full wwPDB X-ray Structure Validation Report (i)

Jan 31, 2016 – 10:23 PM GMT

PDB ID	:	1TGB
Title	:	CRYSTAL STRUCTURE OF BOVINE TRYPSINOGEN AT 1.8 ANGSTROMS RESOLUTION. II. CRYSTALLOGRAPHIC REFINEMENT, REFINED CRYSTAL STRUCTURE AND COMPARISON WITH BOVINE TRYPSIN
Authors	:	Bode, W.; Fehlhammer, H.; Huber, R.
Deposited on	:	1979-03-07
Resolution	:	1.80 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

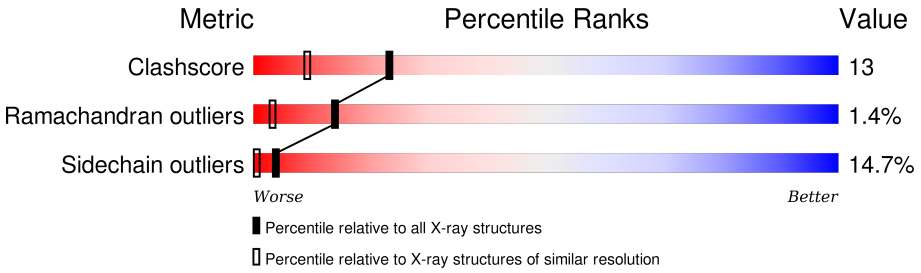
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 1.80 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	102246	5383 (1.80-1.80)
Ramachandran outliers	100387	5320 (1.80-1.80)
Sidechain outliers	100360	5319 (1.80-1.80)

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. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	229	57% 32% 7%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 1750 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 TRYPSINOGEN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	223	1629	1012	279	324	14	143	0	0

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	120	120	120	25	0

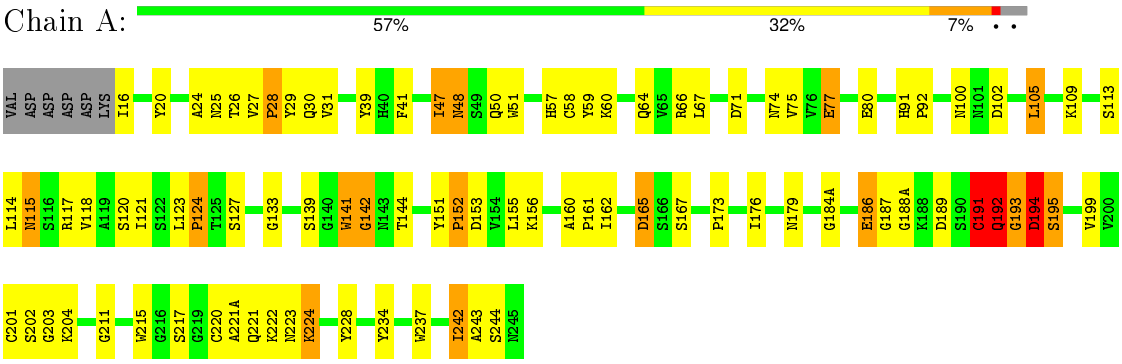
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

Note EDS was not executed.

- Molecule 1: TRYPSINOGEN

Chain A:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	55.10Å 55.10Å 109.40Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	(Not available) – 1.80	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-1.80)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	unknown	Depositor
R, R_{free}	(Not available) , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1750	wwPDB-VP
Average B, all atoms (Å^2)	1.0	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: CA

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	1.43	12/1660 (0.7%)	1.66	49/2250 (2.2%)

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	3

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	191	CYS	N-CA	-26.96	0.92	1.46
1	A	188(A)	GLY	N-CA	13.69	1.66	1.46
1	A	186	GLU	CA-CB	-11.25	1.29	1.53
1	A	191	CYS	CA-CB	9.62	1.75	1.53
1	A	215	TRP	NE1-CE2	-7.54	1.27	1.37
1	A	51	TRP	NE1-CE2	-7.45	1.27	1.37
1	A	141	TRP	NE1-CE2	-7.36	1.27	1.37
1	A	237	TRP	NE1-CE2	-7.28	1.28	1.37
1	A	165	ASP	CG-OD2	6.12	1.39	1.25
1	A	192	GLN	CA-C	-5.63	1.38	1.52
1	A	151	TYR	CA-CB	5.40	1.65	1.53
1	A	77	GLU	CD-OE1	-5.36	1.19	1.25

All (49) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	193	GLY	N-CA-C	-12.01	83.08	113.10

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	192	GLN	CB-CA-C	11.84	134.08	110.40
1	A	192	GLN	CA-C-N	9.65	135.50	116.20
1	A	192	GLN	CA-C-O	-9.62	99.89	120.10
1	A	28	PRO	CA-CB-CG	-9.33	86.27	104.00
1	A	191	CYS	CA-CB-SG	8.98	130.16	114.00
1	A	173	PRO	CB-CA-C	8.88	134.21	112.00
1	A	221(A)	ALA	N-CA-CB	-8.86	97.70	110.10
1	A	189	ASP	CB-CG-OD1	7.64	125.17	118.30
1	A	191	CYS	N-CA-C	-7.38	91.08	111.00
1	A	165	ASP	CB-CG-OD1	7.33	124.90	118.30
1	A	124	PRO	CA-N-CD	-7.30	101.28	111.50
1	A	153	ASP	CB-CG-OD1	7.30	124.87	118.30
1	A	102	ASP	CB-CG-OD1	7.27	124.84	118.30
1	A	71	ASP	CB-CG-OD1	7.21	124.79	118.30
1	A	162	ILE	N-CA-C	-7.14	91.72	111.00
1	A	195	SER	CB-CA-C	6.89	123.19	110.10
1	A	142	GLY	N-CA-C	6.80	130.10	113.10
1	A	133	GLY	N-CA-C	6.78	130.04	113.10
1	A	187	GLY	C-N-CA	6.41	135.76	122.30
1	A	188(A)	GLY	N-CA-C	6.39	129.08	113.10
1	A	211	GLY	N-CA-C	6.29	128.84	113.10
1	A	244	SER	N-CA-C	6.20	127.73	111.00
1	A	58	CYS	CB-CA-C	-6.17	98.07	110.40
1	A	28	PRO	CA-N-CD	-6.12	102.94	111.50
1	A	194	ASP	CB-CG-OD1	6.07	123.77	118.30
1	A	186	GLU	N-CA-CB	5.99	121.39	110.60
1	A	28	PRO	CB-CA-C	5.99	126.97	112.00
1	A	114	LEU	N-CA-C	-5.92	95.01	111.00
1	A	39	TYR	N-CA-C	-5.92	95.02	111.00
1	A	220	CYS	N-CA-C	-5.91	95.03	111.00
1	A	124	PRO	N-CA-C	5.91	127.47	112.10
1	A	16	ILE	CA-CB-CG2	-5.81	99.27	110.90
1	A	47	ILE	CB-CA-C	-5.66	100.28	111.60
1	A	77	GLU	OE1-CD-OE2	5.65	130.08	123.30
1	A	51	TRP	N-CA-C	5.64	126.24	111.00
1	A	176	ILE	CA-CB-CG2	-5.58	99.75	110.90
1	A	20	TYR	CB-CG-CD1	-5.53	117.68	121.00
1	A	60	LYS	N-CA-C	-5.38	96.48	111.00
1	A	234	TYR	CB-CG-CD1	-5.35	117.79	121.00
1	A	59	TYR	CB-CG-CD1	-5.32	117.81	121.00
1	A	244	SER	CB-CA-C	-5.31	100.02	110.10
1	A	41	PHE	N-CA-C	5.24	125.15	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	203	GLY	N-CA-C	5.24	126.19	113.10
1	A	74	ASN	N-CA-C	5.23	125.12	111.00
1	A	28	PRO	N-CA-C	5.20	125.63	112.10
1	A	77	GLU	CG-CD-OE2	-5.14	108.03	118.30
1	A	151	TYR	CB-CG-CD1	-5.12	117.93	121.00
1	A	29	TYR	CB-CG-CD1	-5.05	117.97	121.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	192	GLN	Peptide
1	A	193	GLY	Peptide
1	A	47	ILE	Peptide

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	1629	0	1590	39	0
2	A	1	0	0	0	0
3	A	120	0	0	2	0
All	All	1750	0	1590	39	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 13.

All (39) 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:223:ASN:O	1:A:224:LYS:HD3	1.41	1.20
1:A:109:LYS:HD2	3:A:833:HOH:O	1.47	1.11
1:A:57:HIS:NE2	1:A:195:SER:HB3	1.72	1.04
1:A:109:LYS:CD	3:A:833:HOH:O	2.09	0.85
1:A:115:ASN:HD22	1:A:117:ARG:H	1.24	0.83

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:31:VAL:CG2	1:A:67:LEU:HD23	2.11	0.80
1:A:115:ASN:ND2	1:A:117:ARG:H	1.90	0.69
1:A:91:HIS:CG	1:A:92:PRO:HD2	2.31	0.66
1:A:26:THR:CG2	1:A:26:THR:O	2.47	0.63
1:A:31:VAL:HG22	1:A:67:LEU:HD23	1.81	0.62
1:A:100:ASN:HD21	1:A:179:ASN:HD22	1.48	0.61
1:A:57:HIS:NE2	1:A:195:SER:CB	2.58	0.61
1:A:31:VAL:HG23	1:A:67:LEU:HD23	1.82	0.60
1:A:48:ASN:HD22	1:A:48:ASN:C	2.04	0.59
1:A:64:GLN:HE21	1:A:66:ARG:HE	1.49	0.58
1:A:26:THR:O	1:A:26:THR:HG22	2.03	0.57
1:A:48:ASN:HD22	1:A:50:GLN:H	1.52	0.57
1:A:64:GLN:NE2	1:A:66:ARG:HE	2.04	0.56
1:A:57:HIS:CE1	1:A:195:SER:HB3	2.40	0.55
1:A:24:ALA:O	1:A:25:ASN:HB2	2.06	0.55
1:A:223:ASN:O	1:A:224:LYS:CD	2.34	0.54
1:A:31:VAL:CG2	1:A:67:LEU:CD2	2.87	0.53
1:A:27:VAL:HG12	1:A:30:GLN:HB2	1.91	0.52
1:A:30:GLN:NE2	1:A:139:SER:OG	2.44	0.51
1:A:121:ILE:O	1:A:121:ILE:HG23	2.12	0.48
1:A:242:ILE:HG22	1:A:243:ALA:N	2.26	0.48
1:A:67:LEU:O	1:A:80:GLU:HA	2.14	0.48
1:A:144:THR:HG23	1:A:152:PRO:HD3	1.95	0.47
1:A:64:GLN:HE22	1:A:66:ARG:HH21	1.63	0.47
1:A:141:TRP:O	1:A:152:PRO:HD2	2.15	0.46
1:A:48:ASN:ND2	1:A:48:ASN:C	2.69	0.46
1:A:194:ASP:O	1:A:195:SER:C	2.54	0.45
1:A:160:ALA:HB1	1:A:184(A):GLY:HA2	1.99	0.44
1:A:199:VAL:HG21	1:A:228:TYR:CD1	2.53	0.43
1:A:48:ASN:ND2	1:A:50:GLN:H	2.15	0.43
1:A:115:ASN:C	1:A:115:ASN:HD22	2.23	0.42
1:A:75:VAL:HG22	1:A:77:GLU:CG	2.50	0.42
1:A:115:ASN:ND2	1:A:118:VAL:H	2.19	0.41
1:A:105:LEU:HD12	1:A:105:LEU:HA	1.86	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	221/229 (96%)	202 (91%)	16 (7%)	3 (1%)	14 3

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	142	GLY
1	A	191	CYS
1	A	194	ASP

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	184/190 (97%)	157 (85%)	27 (15%)	4 0

All (27) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	28	PRO
1	A	48	ASN
1	A	105	LEU
1	A	113	SER
1	A	115	ASN
1	A	120	SER
1	A	123	LEU
1	A	124	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	127	SER
1	A	152	PRO
1	A	155	LEU
1	A	156	LYS
1	A	161	PRO
1	A	165	ASP
1	A	167	SER
1	A	186	GLU
1	A	191	CYS
1	A	192	GLN
1	A	194	ASP
1	A	201	CYS
1	A	202	SER
1	A	204	LYS
1	A	217	SER
1	A	221	GLN
1	A	222	LYS
1	A	224	LYS
1	A	242	ILE

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

Mol	Chain	Res	Type
1	A	30	GLN
1	A	48	ASN
1	A	64	GLN
1	A	100	ASN
1	A	101	ASN
1	A	115	ASN
1	A	210	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

There are no carbohydrates in this entry.

5.6 Ligand geometry

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 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

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

6.4 Ligands

EDS was not executed - this section will therefore be empty.

6.5 Other polymers

EDS was not executed - this section will therefore be empty.