wwPDB X-ray Structure Validation Summary Report (i)

Jun 11, 2024 – 06:20 PM EDT

PDB ID	:	1EYG
Title	:	Crystal structure of chymotryptic fragment of E. coli ssb bound to two 35-mer single strand DNAS
Authors	:	Raghunathan, S.; Waksman, G.
Deposited on	:	2000-05-06
Resolution	:	2.80 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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:

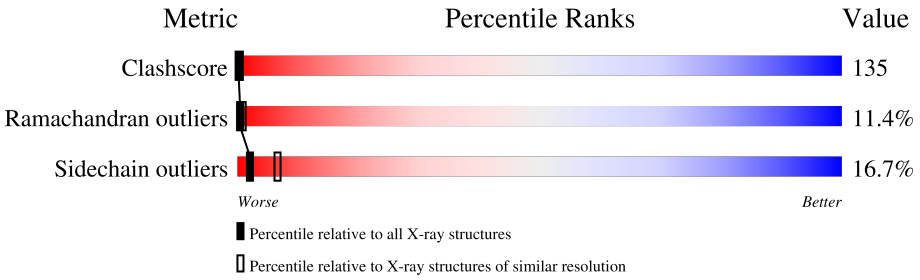
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 2.80 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	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)

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 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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	Q	35	34% 34% 11% 20%
1	R	35	43% 23% 34%
2	A	116	30% 48% 16% • •
2	B	116	28% 47% 13% • 10%
2	C	116	28% 49% 19% • •
2	D	116	25% 54% 16% • •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4244 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 SINGLE STRANDED 28-MER OF D(C).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Q	28	Total	C	N	O	P	0	0	0
			529	252	84	166	27			
1	R	23	Total	C	N	O	P	0	0	0
			431	207	69	134	21			

- Molecule 2 is a protein called SINGLE-STRAND DNA-BINDING PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	112	Total	C	N	O	S	0	0	0
			833	523	147	159	4			
2	B	104	Total	C	N	O	S	0	0	0
			765	479	134	149	3			
2	C	112	Total	C	N	O	S	0	0	0
			819	511	143	161	4			
2	D	112	Total	C	N	O	S	0	0	0
			828	517	148	160	3			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	Q	12	Total	O	0	0
			12	12		
3	R	8	Total	O	0	0
			8	8		
3	A	4	Total	O	0	0
			4	4		
3	B	2	Total	O	0	0
			2	2		
3	C	6	Total	O	0	0
			6	6		
3	D	7	Total	O	0	0
			7	7		

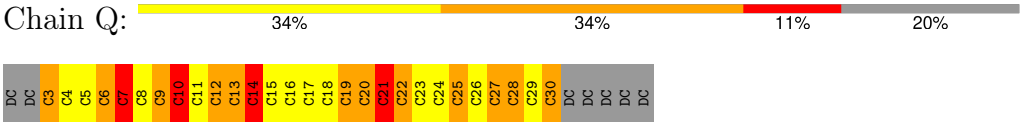
3 Residue-property plots (i)

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. 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.

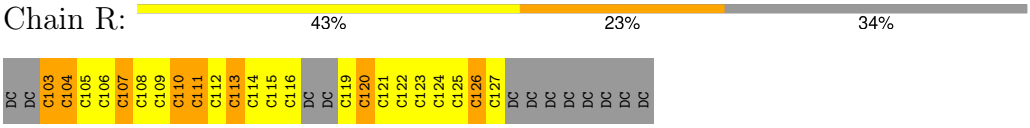
• Molecule 1: SINGLE STRANDED 28-MER OF D(C)

Chain Q:



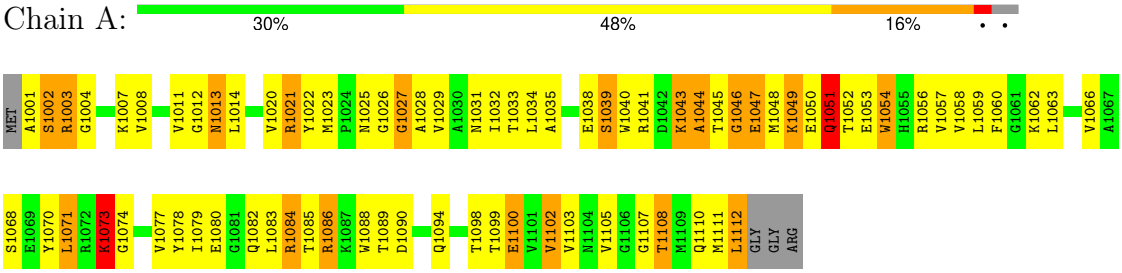
• Molecule 1: SINGLE STRANDED 28-MER OF D(C)

Chain R:



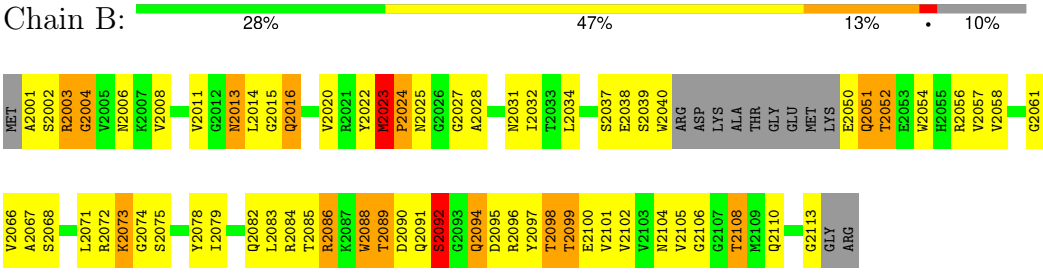
• Molecule 2: SINGLE-STRAND DNA-BINDING PROTEIN

Chain A:



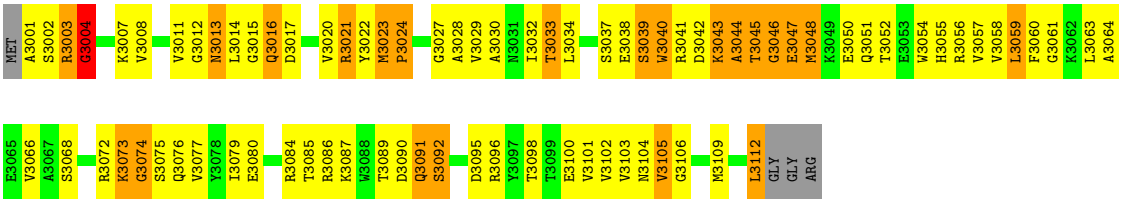
• Molecule 2: SINGLE-STRAND DNA-BINDING PROTEIN

Chain B:



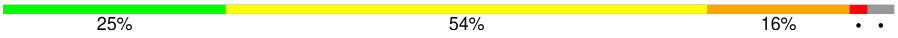
• Molecule 2: SINGLE-STRAND DNA-BINDING PROTEIN

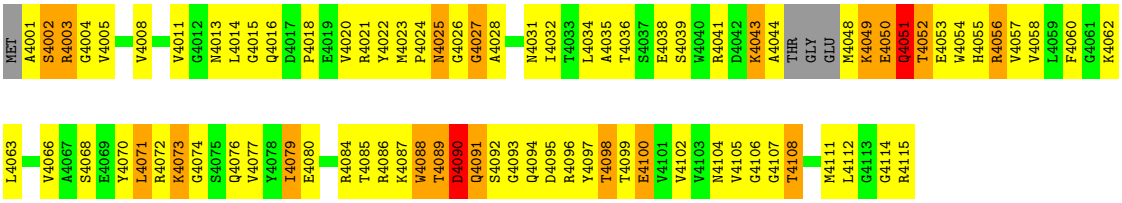
Chain C:



• Molecule 2: SINGLE-STRAND DNA-BINDING PROTEIN

Chain D:





4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	98.69Å 71.08Å 79.16Å 90.00° 91.93° 90.00°	Depositor
Resolution (Å)	30.00 – 2.80	Depositor
% Data completeness (in resolution range)	91.3 (30.00-2.80)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 0.5	Depositor
R, R_{free}	0.256 , 0.298	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4244	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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	Q	1.01	2/584 (0.3%)	1.68	21/888 (2.4%)
1	R	1.10	4/475 (0.8%)	2.16	16/720 (2.2%)
2	A	0.83	2/846 (0.2%)	1.04	6/1146 (0.5%)
2	B	0.80	0/776	0.96	0/1054
2	C	0.77	0/831	0.97	1/1129 (0.1%)
2	D	1.42	2/840 (0.2%)	1.11	3/1137 (0.3%)
All	All	1.01	10/4352 (0.2%)	1.32	47/6074 (0.8%)

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	Q	1	5
2	D	0	1
All	All	1	6

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	4091	GLN	N-CA	32.42	2.11	1.46
2	D	4089	THR	C-O	-11.59	1.01	1.23
2	A	1089	THR	C-O	-10.43	1.03	1.23
1	R	111	DC	C5'-C4'	7.79	1.59	1.51
1	Q	21	DC	O3'-P	6.11	1.68	1.61

The worst 5 of 47 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	R	104	DC	O5'-P-OP1	-27.93	77.19	110.70
1	R	103	DC	OP1-P-O3'	-19.76	61.73	105.20
1	R	104	DC	O5'-P-OP2	-14.09	93.02	105.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	R	103	DC	OP2-P-O3'	-12.18	78.40	105.20
1	Q	21	DC	O4'-C1'-N1	10.84	115.59	108.00

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	Q	22	DC	C3'

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	Q	10	DC	Sidechain
1	Q	14	DC	Sidechain
1	Q	19	DC	Sidechain
1	Q	21	DC	Sidechain
1	Q	7	DC	Sidechain

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	Q	529	0	310	369	0
1	R	431	0	257	318	0
2	A	833	0	802	185	0
2	B	765	0	726	206	1
2	C	819	0	770	258	1
2	D	828	0	784	254	0
3	A	4	0	0	2	0
3	B	2	0	0	0	0
3	C	6	0	0	2	0
3	D	7	0	0	2	0
3	Q	12	0	0	4	0
3	R	8	0	0	4	0
All	All	4244	0	3649	1062	1

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 135.

The worst 5 of 1062 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:D:4111:MET:SD	2:D:4111:MET:CE	2.02	1.47
1:R:109:DC:H2'	1:R:110:DC:C6	1.51	1.46
1:R:111:DC:C5'	2:B:2102:VAL:HG11	1.55	1.36
1:R:125:DC:H5'	2:A:1056:ARG:NH2	1.42	1.34
1:R:121:DC:O5'	2:A:1054:TRP:HZ2	1.08	1.28

All (1) 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)
2:B:2086:ARG:NH2	2:C:3089:THR:O[3_445]	2.17	0.03

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	110/116 (95%)	87 (79%)	11 (10%)	12 (11%)	0	1
2	B	100/116 (86%)	82 (82%)	8 (8%)	10 (10%)	0	1
2	C	110/116 (95%)	82 (74%)	11 (10%)	17 (16%)	0	0
2	D	108/116 (93%)	85 (79%)	13 (12%)	10 (9%)	0	1
All	All	428/464 (92%)	336 (78%)	43 (10%)	49 (11%)	0	1

5 of 49 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	1027	GLY
2	A	1046	GLY
2	A	1047	GLU
2	A	1105	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	2004	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
2	A	82/95 (86%)	67 (82%)	15 (18%)	1 5
2	B	76/95 (80%)	61 (80%)	15 (20%)	1 4
2	C	80/95 (84%)	68 (85%)	12 (15%)	3 9
2	D	80/95 (84%)	69 (86%)	11 (14%)	3 11
All	All	318/380 (84%)	265 (83%)	53 (17%)	2 6

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	2099	THR
2	C	3033	THR
2	D	4088	TRP
2	B	2108	THR
2	C	3016	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	4031	ASN
2	D	4110	GLN
2	D	4055	HIS
2	C	3013	ASN
2	C	3094	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

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

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.