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 ⓘ

Feb 12, 2017 – 09:29 pm GMT

PDB ID	:	4GCH
Title	:	STRUCTURE AND ACTIVITY OF TWO PHOTOREVERSIBLE CINNA-MATES BOUND TO CHYMOTRYPSIN
Authors	:	Stoddard, B.L.; Ringe, D.; Petsko, G.A.
Deposited on	:	1989-09-25
Resolution	:	1.90 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

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

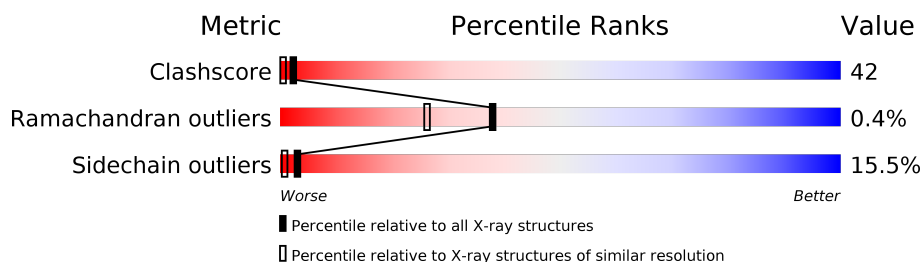|		
MolProbity	:	4.02b-467
Mogul	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

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 1.90 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
Clashscore		42
Ramachandran outliers		0.4%
Sidechain outliers		15.5%

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))
Clashscore	112137	5731 (1.90-1.90)
Ramachandran outliers	110173	5669 (1.90-1.90)
Sidechain outliers	110143	5670 (1.90-1.90)

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	E	13	15% 46% 15% 8% 15%
2	F	131	37% 48% 11% •
3	G	97	40% 33% 22% • •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	DMC	G	246	-	-	X	-

WORLDWIDE
wwPDB
PROTEIN DATA BANK

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 1826 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 GAMMA-CHYMOTRYPSIN A.

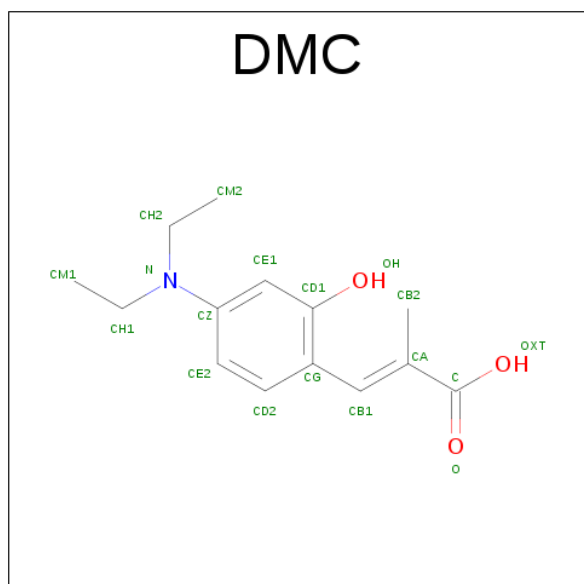| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	11	Total	C	N	O	S	0	0	1
			69	45	12	11	1			

- Molecule 2 is a protein called GAMMA-CHYMOTRYPSIN A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	131	Total	C	N	O	S	0	0	0
			979	618	162	195	4			

- Molecule 3 is a protein called GAMMA-CHYMOTRYPSIN A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	95	Total	C	N	O	S	0	0	0
			688	429	120	132	7			

- Molecule 4 is 3-(4-DIETHYLAMINO-2-HYDROXY-PHENYL)-2-METHYL-PROPIONIC ACID (three-letter code: DMC) (formula: $C_{14}H_{19}NO_3$).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	G	1	Total 17	C 14	N 1	O 2	0	0

- Molecule 5 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	E	4	Total 4	O 4	0	0
5	F	50	Total 50	O 50	0	0
5	G	19	Total 19	O 19	0	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 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.
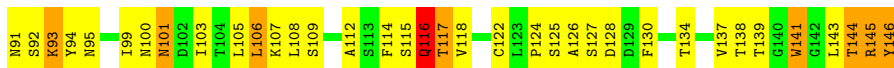
Note EDS was not executed.
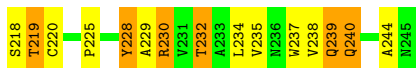
• Molecule 1: GAMMA-CHYMOTRYPSIN A

Chain E:
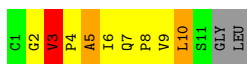
• Molecule 2: GAMMA-CHYMOTRYPSIN A

Chain F:

• Molecule 3: GAMMA-CHYMOTRYPSIN A

Chain G:

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, α, β, γ	69.70Å　　69.70Å　　97.50Å 90.00°　　90.00°　　90.00°	Depositor
Resolution (Å)	45.00　–　1.90	Depositor
% Data completeness (in resolution range)	(Not available) (45.00-1.90)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	PROLSQ	Depositor
R, R_{free}	0.209　,　(Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1826	wwPDB-VP
Average B, all atoms (Å^2)	8.0	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: DMC

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	E	1.39	0/70	2.37	7/97 (7.2%)
2	F	1.23	3/999 (0.3%)	2.22	40/1361 (2.9%)
3	G	1.10	1/701 (0.1%)	2.03	18/955 (1.9%)
All	All	1.19	4/1770 (0.2%)	2.15	65/2413 (2.7%)

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	E	0	1
2	F	0	1
3	G	0	1
All	All	0	3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	20	GLU	CD-OE2	8.01	1.34	1.25
2	F	49	GLU	CD-OE2	7.96	1.34	1.25
2	F	70	GLU	CD-OE2	7.08	1.33	1.25
3	G	214	SER	CB-OG	-6.85	1.33	1.42

The worst 5 of 65 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	145	ARG	NE-CZ-NH2	13.71	127.16	120.30
2	F	64	ASP	CB-CG-OD1	13.23	130.21	118.30
2	F	145	ARG	NE-CZ-NH1	-10.33	115.14	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	94	TYR	CB-CG-CD1	-8.99	115.60	121.00
2	F	146	TYR	CB-CG-CD1	-8.62	115.83	121.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	10	LEU	Mainchain
2	F	145	ARG	Sidechain
3	G	154	ARG	Sidechain

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	E	69	0	76	8	1
2	F	979	0	948	83	1
3	G	688	0	686	71	7
4	G	17	0	15	29	0
5	E	4	0	0	1	5
5	F	50	0	0	15	5
5	G	19	0	0	4	1
All	All	1826	0	1725	145	11

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 42.

The worst 5 of 145 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)
4:G:246:DMC:CM2	4:G:246:DMC:CH2	1.77	1.55
3:G:217:SER:O	4:G:246:DMC:CH1	1.73	1.36
4:G:246:DMC:CH2	4:G:246:DMC:N	1.86	1.35
4:G:246:DMC:HE1	4:G:246:DMC:CM2	1.63	1.26
4:G:246:DMC:HE1	4:G:246:DMC:HM23	1.35	1.06

The worst 5 of 11 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)
5:F:274:HOH:O	5:F:286:HOH:O[2_675]	0.35	1.85
3:G:240:GLN:CG	5:E:251:HOH:O[5_646]	1.08	1.12
3:G:240:GLN:CD	5:E:251:HOH:O[5_646]	1.11	1.09
3:G:240:GLN:CB	5:E:251:HOH:O[5_646]	1.66	0.54
5:F:308:HOH:O	5:F:308:HOH:O[8_777]	1.67	0.53

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	E	9/13 (69%)	9 (100%)	0	0	100	100
2	F	129/131 (98%)	119 (92%)	9 (7%)	1 (1%)	22	11
3	G	93/97 (96%)	84 (90%)	9 (10%)	0	100	100
All	All	231/241 (96%)	212 (92%)	18 (8%)	1 (0%)	38	26

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	48	ASN

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	E	8/10 (80%)	7 (88%)	1 (12%)	5	2
2	F	109/109 (100%)	99 (91%)	10 (9%)	11	4
3	G	76/77 (99%)	57 (75%)	19 (25%)	1	0
All	All	193/196 (98%)	163 (84%)	30 (16%)	3	1

5 of 30 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	G	160	LEU
3	G	166	THR
3	G	232	THR
3	G	165	ASN
3	G	170	LYS

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

Mol	Chain	Res	Type
2	F	100	ASN
3	G	239	GLN
2	F	116	GLN
2	F	48	ASN
3	G	165	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry (i)

1 ligand 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	#$|Z| > 2$
4	DMC	G	246	-	17,17,18	9.07	10 (58%)	20,22,24	4.65	10 (50%)

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	DMC	G	246	-	-	0/14/14/16	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	G	246	DMC	OH-CD1	-14.62	1.06	1.36
4	G	246	DMC	CZ-N	-5.97	1.22	1.38
4	G	246	DMC	CM1-CH1	-3.60	1.28	1.49
4	G	246	DMC	CB2-CA	-3.51	1.43	1.50
4	G	246	DMC	CD2-CE2	-2.41	1.34	1.38

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	G	246	DMC	CH2-N-CZ	-8.92	107.77	121.39
4	G	246	DMC	CB2-CA-C	-5.63	111.31	117.27
4	G	246	DMC	CE1-CD1-CG	-3.36	117.77	120.76
4	G	246	DMC	OH-CD1-CE1	-2.58	112.59	119.47
4	G	246	DMC	CM2-CH2-N	2.69	125.70	112.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 29 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	G	246	DMC	29	0

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.