W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 12:02 am BST

PDB ID	:	6CHA
Title	:	STRUCTURE OF A TETRAHEDRAL TRANSITION STATE COMPLEX OF ALPHA-*CHYMOTRYPSIN AT 1.8-*ANGSTROMS RESOLUTION
Authors	:	Tulinsky, A.; Blevins, R.A.
Deposited on	:	1987-02-06
Resolution	:	1.80 Å(reported)

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

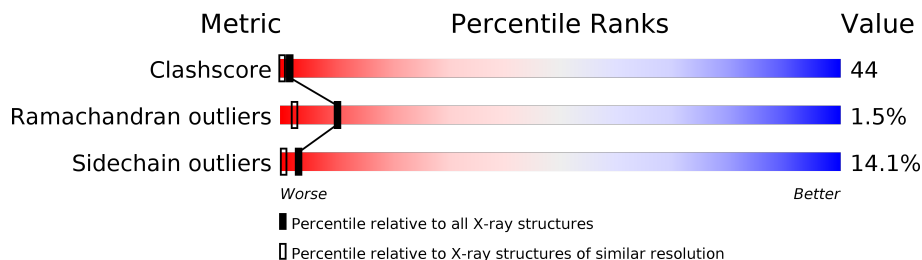|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 1.80 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	Percentile Ranks	Value
Clashscore		44
Ramachandran outliers		1.5%
Sidechain outliers		14.1%

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	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (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 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	A	13	23% / 46% / 31%
1	E	13	31% / 31% / 8% / 31%
2	B	131	27% / 49% / 18% / 6%
2	F	131	34% / 45% / 17% / •
3	C	97	33% / 39% / 23% / 5%
3	G	97	33% / 44% / 21% / •

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 3679 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 ALPHA-CHYMOTRYPSIN A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	9	Total	C	N	O	S	0	0	1
			54	34	10	9	1			
1	E	9	Total	C	N	O	S	0	0	1
			54	34	10	9	1			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	131	Total	C	N	O	S	0	0	0
			980	618	162	196	4			
2	F	131	Total	C	N	O	S	0	0	0
			980	618	162	196	4			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	97	Total	C	N	O	S	0	0	0
			702	436	123	136	7			
3	G	97	Total	C	N	O	S	0	0	0
			702	436	123	136	7			

- Molecule 4 is PHENYLETHANE BORONIC ACID (three-letter code: PBA) (formula: $C_8H_{11}BO_2$).

PBA
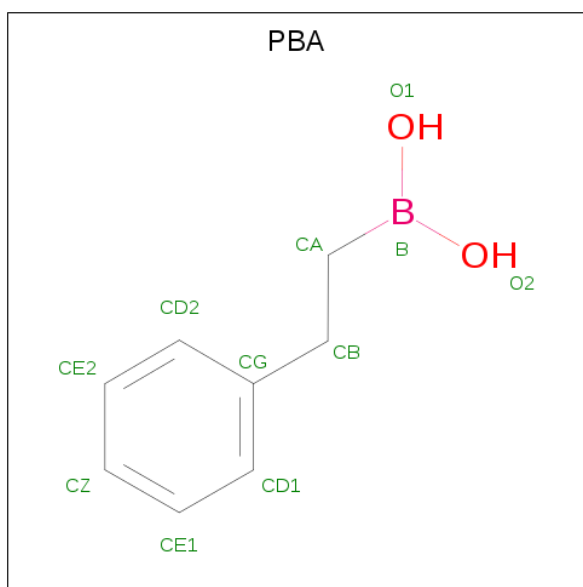
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	C	1	Total	B	C	O	0	0
			11	1	8	2		
4	G	1	Total	B	C	O	0	0
			11	1	8	2		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	6	Total	O	0	0
			6	6		
5	B	43	Total	O	0	0
			43	43		
5	C	39	Total	O	0	0
			39	39		
5	E	2	Total	O	0	0
			2	2		
5	F	55	Total	O	0	0
			55	55		
5	G	40	Total	O	0	0
			40	40		

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.

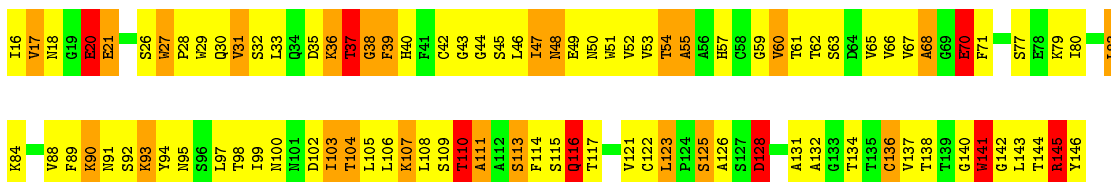- Molecule 1: ALPHA-CHYMOTRYPSIN A

Chain A: 23% 46% 31%

C1 G2 V3 P4 A5 P8 V9 LEU SER GLY LEU

- Molecule 1: ALPHA-CHYMOTRYPSIN A

Chain E: 31% 31% 8% 31%

C1 G2 V3 P4 Q7 P8 V9 LEU SER GLY LEU

- Molecule 2: ALPHA-CHYMOTRYPSIN A

Chain B: 27% 49% 18% 6%

I16 V17 N18 G19 E20 E21 S26 W27 P28 W29 Q30 V31 S32 L33 Q34 D35 K36 T37 G38 F39 H40 F41 C42 G43 G44 S45 L46 I47 N48 E49 N50 W51 V52 V53 T54 A55 A56 H57 C58 G59 V60 T61 T62 S63 D64 V65 V66 V67 A68 G69 E70 F71 S77 E78 K79 I80 L83

K84 V88 F89 K90 N91 S92 K93 Y94 N95 S96 L97 T98 I99 N100 N101 D102 I103 T104 L105 L106 K107 L108 S109 T110 A111 A112 S113 F114 S115 Q116 T117 V121 C122 L123 P124 S125 A126 S127 D128 A131 A132 G133 T134 T135 C136 V137 T138 T139 G140 W141 G142 L143 T144 R145 Y146

- Molecule 2: ALPHA-CHYMOTRYPSIN A

Chain F: 34% 45% 17% •

I16 E20 E21 S26 W27 P28 W29 Q30 V31 S32 L33 Q34 D35 K36 T37 G38 F39 H40 F41 C42 L46 I47 N48 E49 V52 V53 H57 V60 T61 T62 S63 D64 V65 V66 E70 F71 D72 Q73 G74 S75 S76 S77 E78 K79 I80 Q81 K82 L83 K84 I85 A86 K87 V88

F89 K90 N91 S92 K93 Y94 N95 S96 L97 T98 I99 N100 N101 D102 I103 T104 L105 L106 K107 L108 S109 T110 S115 Q116 T117 V118 S119 A120 V121 C122 S125 A126 S127 D128 D129 F130 A131 T135 C136 V137 T138 T139 G140 W141 T144 R145 Y146

- Molecule 3: ALPHA-CHYMOTRYPSIN A

Chain C: 33% 39% 23% 5%

A149 N150 T151 P152 D153 R154 L155 Q156 Q157 A158 S159 L160 P161 L162 L163 S164 N165 T166 N167 C168 K169 K170 Y171 W172 G173 T174 K175 I176 K177 D178 A179 M180 I181 C182 A183 G184 A185 S186 G187 V188 S189 S190 C191 M192 S195 G196 L199 V200 C201 K202 K203 T208 I212 W215 G216

S217 S218 T219 C220 S221 T222 S223 T224 Y228 T232 A233 L234 V235 N236 W237 V238 Q239 Q240 T241 L242 N245

• Molecule 3: ALPHA-CHYMOTRYPSIN A

Chain G:

33%   44%   21%

A149 N150 T151 P152 D153 R154 L155 Q156 Q157 A158 S159 L160 P161 L162 L163 S164 N165 T166 N167 C168 K169 K170 Y171 W172 G173 T174 K175 I176 K177 D178 A179 M180 I181 S186 S189 S190 C191 M192 G193 D194 S195 G196 G197 V200 C201 K202 K203 N204 G205 A206 W207 T208 L209 V210 G211 I212 V213

S214 W215 G216 S217 S218 T219 S223 T224 P225 Y228 A229 R230 V231 T232 A233 L234 V235 N236 W237 V238 Q239 Q240 T241 L242 A243 A244 N245

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	49.27Å 67.16Å 65.91Å 90.00° 101.68° 90.00°	Depositor
Resolution (Å)	5.00 – 1.80	Depositor
% Data completeness (in resolution range)	(Not available) (5.00-1.80)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	PROLSQ	Depositor
R, R_{free}	0.200 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3679	wwPDB-VP
Average B, all atoms (Å²)	17.0	wwPDB-VP

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: PBA

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	1.40	0/55	1.78	1/76 (1.3%)
1	E	1.50	0/55	1.89	1/76 (1.3%)
2	B	1.73	13/1000 (1.3%)	2.59	75/1361 (5.5%)
2	F	1.71	13/1000 (1.3%)	2.32	49/1361 (3.6%)
3	C	1.84	15/715 (2.1%)	2.45	48/973 (4.9%)
3	G	2.08	18/715 (2.5%)	2.63	52/973 (5.3%)
All	All	1.81	59/3540 (1.7%)	2.48	226/4820 (4.7%)

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
2	F	0	1

The worst 5 of 59 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	G	217	SER	CB-OG	13.45	1.59	1.42
2	F	93	LYS	CD-CE	10.11	1.76	1.51
2	F	96	SER	CB-OG	10.03	1.55	1.42
3	G	215	TRP	C-O	8.39	1.39	1.23
3	C	170	LYS	CE-NZ	7.88	1.68	1.49

The worst 5 of 226 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	20	GLU	OE1-CD-OE2	18.66	145.70	123.30
2	F	145	ARG	NE-CZ-NH1	-16.57	112.02	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	G	154	ARG	NE-CZ-NH2	-16.27	112.17	120.30
3	G	153	ASP	CB-CG-OD2	-14.58	105.18	118.30
2	B	20	GLU	CG-CD-OE2	-13.62	91.06	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	F	145	ARG	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	A	54	0	56	8	0
1	E	54	0	56	9	4
2	B	980	0	950	101	7
2	F	980	0	951	91	0
3	C	702	0	697	71	2
3	G	702	0	698	53	0
4	C	11	0	11	3	0
4	G	11	0	11	3	0
5	A	6	0	0	4	0
5	B	43	0	0	30	0
5	C	39	0	0	12	0
5	E	2	0	0	1	0
5	F	55	0	0	24	1
5	G	40	0	0	24	0
All	All	3679	0	3430	297	7

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 44.

The worst 5 of 297 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:F:93:LYS:CE	2:F:93:LYS:CD	1.76	1.61
3:C:170:LYS:CE	3:C:170:LYS:NZ	1.68	1.56
1:E:7:GLN:NE2	1:E:7:GLN:H	1.34	1.24
2:F:91:ASN:ND2	2:F:93:LYS:H	1.36	1.21
2:F:104:THR:HG21	5:F:698:HOH:O	1.40	1.18

The worst 5 of 7 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:116:GLN:CG	1:E:7:GLN:OE1[2_746]	1.51	0.69
2:B:116:GLN:NE2	1:E:7:GLN:OE1[2_746]	1.71	0.49
2:B:116:GLN:CD	1:E:7:GLN:OE1[2_746]	1.87	0.33
2:B:79:LYS:CE	3:C:170:LYS:CD[1_655]	2.07	0.13
2:B:79:LYS:CD	3:C:170:LYS:NZ[1_655]	2.18	0.02

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	7/13 (54%)	7 (100%)	0	0	100	100
1	E	7/13 (54%)	7 (100%)	0	0	100	100
2	B	129/131 (98%)	119 (92%)	7 (5%)	3 (2%)	6	1
2	F	129/131 (98%)	119 (92%)	7 (5%)	3 (2%)	6	1
3	C	95/97 (98%)	87 (92%)	7 (7%)	1 (1%)	14	4
3	G	95/97 (98%)	91 (96%)	4 (4%)	0	100	100
All	All	462/482 (96%)	430 (93%)	25 (5%)	7 (2%)	10	2

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	75	SER
2	F	99	ILE
2	F	77	SER
2	B	99	ILE
2	B	77	SER

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	6/10 (60%)	6 (100%)	0	100	100
1	E	6/10 (60%)	5 (83%)	1 (17%)	2	0
2	B	109/109 (100%)	89 (82%)	20 (18%)	1	0
2	F	109/109 (100%)	94 (86%)	15 (14%)	3	1
3	C	77/77 (100%)	66 (86%)	11 (14%)	3	0
3	G	77/77 (100%)	70 (91%)	7 (9%)	9	2
All	All	384/392 (98%)	330 (86%)	54 (14%)	3	0

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	178	ASP
3	C	234	LEU
3	G	177	LYS
3	C	181	ILE
3	C	192	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	7	GLN
3	G	165	ASN
2	F	50	ASN
3	C	165	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	F	48	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)

2 ligands are 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	PBA	C	1	3	9,11,11	3.58	3 (33%)	11,13,13	1.79	3 (27%)
4	PBA	G	1	-	9,11,11	1.29	1 (11%)	11,13,13	1.66	1 (9%)

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	PBA	C	1	3	-	2/3/5/5	0/1/1/1
4	PBA	G	1	-	-	3/3/5/5	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	1	PBA	B-CA	10.01	1.67	1.56
4	G	1	PBA	B-CA	-2.62	1.53	1.56
4	C	1	PBA	CD2-CG	2.27	1.43	1.38
4	C	1	PBA	CB-CG	2.09	1.57	1.51

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	1	PBA	O1-B-CA	-4.46	112.43	121.20
4	C	1	PBA	O2-B-CA	-3.43	114.45	121.20
4	C	1	PBA	O1-B-CA	-2.87	115.55	121.20
4	C	1	PBA	CD2-CG-CD1	2.13	121.51	118.17

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	G	1	PBA	CA-CB-CG-CD1
4	G	1	PBA	CA-CB-CG-CD2
4	G	1	PBA	B-CA-CB-CG
4	C	1	PBA	CA-CB-CG-CD1
4	C	1	PBA	CA-CB-CG-CD2

There are no ring outliers.

2 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	1	PBA	3	0
4	G	1	PBA	3	0

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.