wwPDB X-ray Structure Validation Summary Report (i)

May 17, 2020 – 11:07 am BST

PDB ID	:	4CHA
Title	:	STRUCTURE OF ALPHA-*CHYMOTRYPSIN REFINED AT 1.68 ANGSTROMS RESOLUTION
Authors	:	Tsukada, H.; Blow, D.M.
Deposited on	:	1984-11-26
Resolution	:	1.68 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 following versions of software and data (see references (1)) were used in the production of this report:

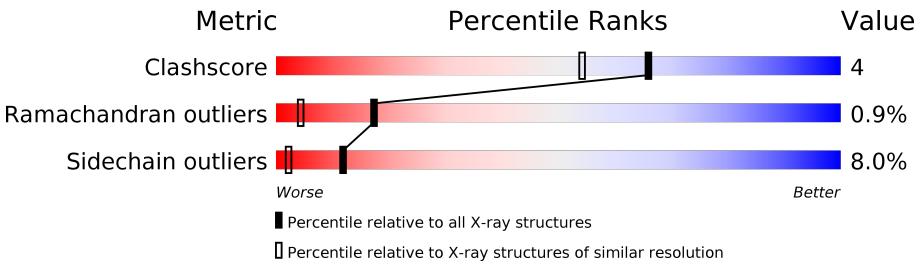
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.68 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	7310 (1.70-1.66)
Ramachandran outliers	138981	7173 (1.70-1.66)
Sidechain outliers	138945	7172 (1.70-1.66)

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%

Mol	Chain	Length	Quality of chain
1	A	13	62% 23% 15%
1	E	13	54% 23% 23%
2	B	131	79% 18% •
2	F	131	74% 21% 5% •
3	C	97	78% 18% • •
3	G	97	81% 13% • •

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 3591 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
			Total	C	N	O	S			
1	A	11	74	48	12	13	1	0	0	0
1	E	10	68	45	11	11	1	0	0	0

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

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

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

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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	1	1	1	0	0
4	B	23	23	23	0	0
4	C	18	18	18	0	0
4	F	21	21	21	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	G	22	22	22	0	0

3 Residue-property plots (i)

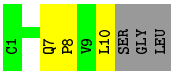
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according 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.

- Molecule 1: ALPHA-CHYMOTRYPSIN A

Chain A:

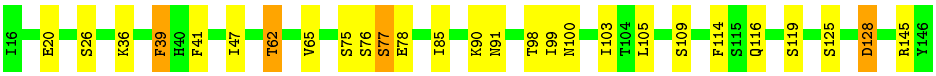
- Molecule 1: ALPHA-CHYMOTRYPSIN A

Chain E:



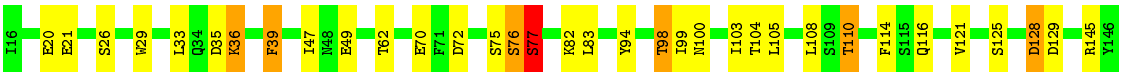
- Molecule 2: ALPHA-CHYMOTRYPSIN A

Chain B:



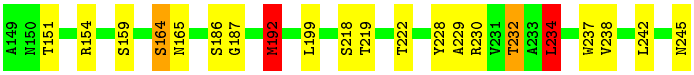
- Molecule 2: ALPHA-CHYMOTRYPSIN A

Chain F:



- Molecule 3: ALPHA-CHYMOTRYPSIN A

Chain C:



- Molecule 3: ALPHA-CHYMOTRYPSIN A

Chain G:

A149
N150
T151
R154
Q157
S164
N165
T166
N167
K177
S190
G191
M192
K203
T208
T219
Y228
A229
R230
V231
T232
A233
L234
V238
L242
N245

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	49.23Å 67.39Å 65.99Å 90.00° 101.80° 90.00°	Depositor
Resolution (Å)	(Not available) – 1.68 29.87 – 1.69	Depositor EDS
% Data completeness (in resolution range)	(Not available) ((Not available)-1.68) 94.4 (29.87-1.69)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.93 (at 1.70Å)	Xtriage
Refinement program	unknown	Depositor
R, R_{free}	(Not available) , (Not available) 0.235 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	15.2	Xtriage
Anisotropy	0.273	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 46.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	3591	wwPDB-VP
Average B, all atoms (Å^2)	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.18% 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 (i)

5.1 Standard geometry (i)

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	0.71	0/75	1.79	2/103 (1.9%)
1	E	0.78	0/69	1.54	0/95
2	B	0.91	0/1000	1.79	14/1361 (1.0%)
2	F	0.92	0/1000	1.91	24/1361 (1.8%)
3	C	0.94	1/715 (0.1%)	1.69	12/973 (1.2%)
3	G	0.93	0/715	1.72	14/973 (1.4%)
All	All	0.92	1/3574 (0.0%)	1.79	66/4866 (1.4%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	245	ASN	C-OXT	5.16	1.33	1.23

The worst 5 of 66 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	128	ASP	CB-CG-OD1	23.73	139.66	118.30
2	B	128	ASP	CB-CG-OD1	15.12	131.91	118.30
2	B	114	PHE	CB-CG-CD1	11.31	128.71	120.80
2	F	145	ARG	NE-CZ-NH2	10.08	125.34	120.30
2	B	125	SER	N-CA-CB	9.08	124.12	110.50

There are no chirality outliers.

There are no planarity outliers.

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	74	0	81	1	0
1	E	68	0	76	3	0
2	B	980	0	950	7	0
2	F	980	0	951	9	1
3	C	702	0	698	10	0
3	G	702	0	698	11	1
4	A	1	0	0	0	0
4	B	23	0	0	0	0
4	C	18	0	0	0	0
4	F	21	0	0	0	0
4	G	22	0	0	1	0
All	All	3591	0	3454	28	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 4.

The worst 5 of 28 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)
3:C:192:MET:HG2	3:G:192:MET:HG2	1.66	0.78
2:F:98:THR:HG22	2:F:100:ASN:HB2	1.73	0.70
2:B:39:PHE:HE2	3:G:151:THR:HG21	1.64	0.61
3:G:165:ASN:HD21	3:G:230:ARG:HH11	1.48	0.60
2:B:98:THR:HG22	2:B:100:ASN:HB2	1.85	0.58

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:F:76:SER:CB	3:G:167:ASN:ND2[1_455]	2.16	0.04

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	9/13 (69%)	9 (100%)	0	0	100	100
1	E	8/13 (62%)	8 (100%)	0	0	100	100
2	B	129/131 (98%)	125 (97%)	2 (2%)	2 (2%)	9	1
2	F	129/131 (98%)	125 (97%)	2 (2%)	2 (2%)	9	1
3	C	95/97 (98%)	91 (96%)	4 (4%)	0	100	100
3	G	95/97 (98%)	90 (95%)	5 (5%)	0	100	100
All	All	465/482 (96%)	448 (96%)	13 (3%)	4 (1%)	17	4

All (4) Ramachandran outliers are listed below:

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

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	9/10 (90%)	9 (100%)	0	100	100
1	E	8/10 (80%)	8 (100%)	0	100	100
2	B	109/109 (100%)	98 (90%)	11 (10%)	7	1
2	F	109/109 (100%)	98 (90%)	11 (10%)	7	1
3	C	77/77 (100%)	72 (94%)	5 (6%)	17	4
3	G	77/77 (100%)	73 (95%)	4 (5%)	23	7
All	All	389/392 (99%)	358 (92%)	31 (8%)	12	2

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	232	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	F	39	PHE
3	G	203	LYS
3	C	234	LEU
2	F	49	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	18	ASN
2	B	34	GLN
2	B	48	ASN
3	C	165	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)

There are no ligands in this entry.

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)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.