wwPDB X-ray Structure Validation Summary Report (i)

Nov 5, 2023 – 02:14 AM EST

PDB ID	:	1CZI
Title	:	CHYMOSIN COMPLEX WITH THE INHIBITOR CP-113972
Authors	:	Groves, M.R.; Dhanaraj, V.; Pitts, J.E.; Badasso, M.; Hoover, D.; Nugent, P.; Blundell, T.L.
Deposited on	:	1997-01-15
Resolution	:	2.30 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:

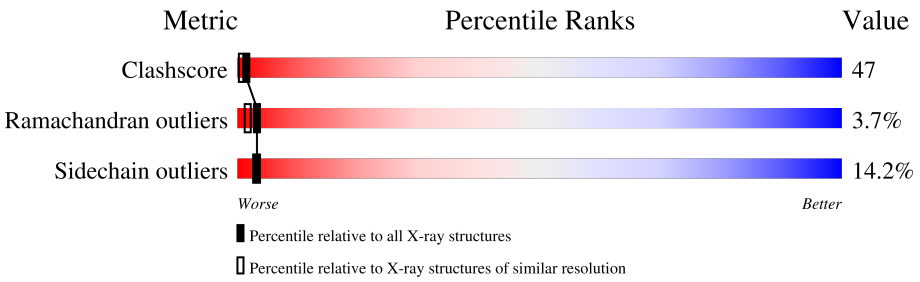
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.30 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	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)

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%

Mol	Chain	Length	Quality of chain
1	E	323	37% 49% 11% •
2	P	4	25% 75%

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
2	PHI	P	3	-	-	X	-

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2745 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 CHYMOSIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	E	323	2511	1598	401	498	14	0	0	0

- Molecule 2 is a protein called CP-113972 (NORSTATINE-S-METHYL CYSTEINE-IODO-PHENYLALANINE-PROLINE).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	I	N	O	S			
2	P	4	43	31	1	4	6	1	0	0	0

- Molecule 3 is water.

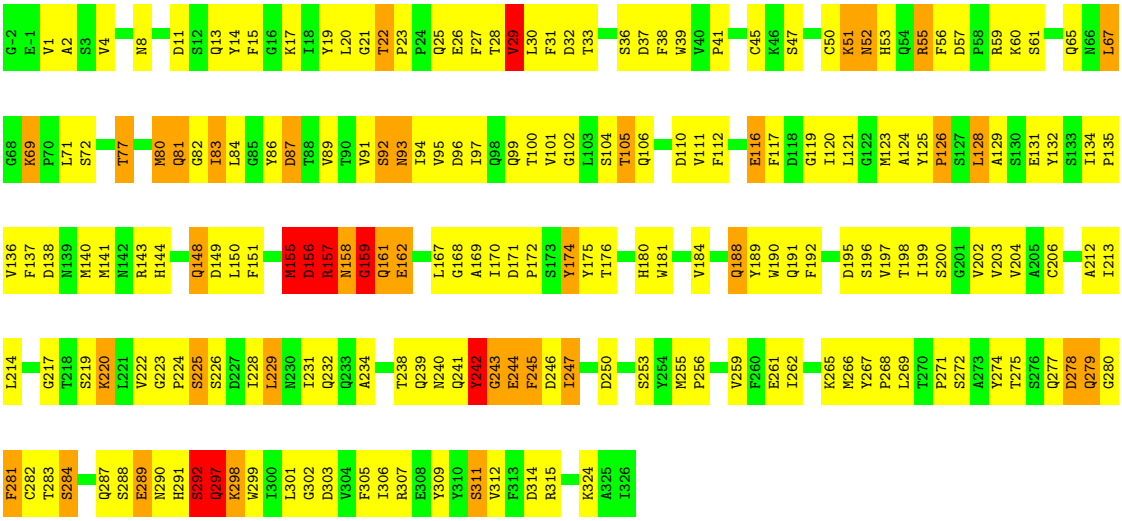
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	E	191	191	191	0	0

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.

• Molecule 1: CHYMOSIN

Chain E: 37% 49% 11%



• Molecule 2: CP-113972 (NORSTATINE-S-METHYL CYSTEINE-IODO-PHENYLALANINE-PROLINE)

Chain P: 25% 75%

4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α, β, γ	132.78Å 132.78Å 81.95Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	9.97 – 2.30 21.01 – 2.30	Depositor EDS
% Data completeness (in resolution range)	98.0 (9.97-2.30) 97.6 (21.01-2.30)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.50 (at 2.30Å)	Xtriage
Refinement program	RESTRAIN	Depositor
R, R_{free}	0.195 , (Not available) 0.265 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	14.5	Xtriage
Anisotropy	0.227	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 20.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.023 for -1/3*h+1/3*k+4/3*l,-k,2/3*h+1/3*k+1/3*l 0.035 for -2/3*h-1/3*k-4/3*l,-1/3*h-2/3*k+4/3*l,-1/3*h+1/3*k+1/3*l 0.017 for -h,1/3*h-1/3*k-4/3*l,-1/3*h-2/3*k+1/3*l	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	2745	wwPDB-VP
Average B, all atoms (Å^2)	0.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.41% 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)

Bond lengths and bond angles in the following residue types are not validated in this section: NOR, PHI, SMC

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	0.46	0/2574	1.40	18/3506 (0.5%)
2	P	0.24	0/7	0.89	0/8
All	All	0.46	0/2581	1.40	18/3514 (0.5%)

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	4

There are no bond length outliers.

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	159	GLY	C-N-CA	16.02	161.74	121.70
1	E	156	ASP	C-N-CA	10.95	149.09	121.70
1	E	278	ASP	CB-CG-OD1	-7.57	111.49	118.30
1	E	309	TYR	CB-CG-CD1	-7.48	116.51	121.00
1	E	157	ARG	C-N-CA	7.44	140.30	121.70

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	143	ARG	Sidechain
1	E	155	MET	Peptide
1	E	157	ARG	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	E	159	GLY	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	E	2511	0	2382	227	6
2	P	43	0	36	19	0
3	E	191	0	0	25	3
All	All	2745	0	2418	232	9

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 47.

The worst 5 of 232 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:P:3:PHI:I	2:P:3:PHI:CZ	1.36	1.74
2:P:3:PHI:I	2:P:3:PHI:CE2	2.42	1.36
2:P:3:PHI:I	2:P:3:PHI:CE1	2.44	1.35
1:E:77:THR:HG22	1:E:77:THR:O	1.54	1.03
1:E:243:GLY:O	1:E:244:GLU:HB2	1.58	1.02

The worst 5 of 9 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)
3:E:568:HOH:O	3:E:568:HOH:O[6_766]	0.58	1.62
1:E:242:TYR:CD2	1:E:244:GLU:CG[18_655]	1.47	0.73
1:E:241:GLN:O	1:E:246:ASP:OD2[18_655]	1.69	0.51
1:E:240:ASN:OD1	1:E:240:ASN:OD1[18_655]	1.87	0.33
3:E:450:HOH:O	3:E:450:HOH:O[10_546]	1.92	0.28

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	E	321/323 (99%)	284 (88%)	25 (8%)	12 (4%)	3 2

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	157	ARG
1	E	158	ASN
1	E	242	TYR
1	E	244	GLU
1	E	297	GLN

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	E	280/280 (100%)	240 (86%)	40 (14%)	3 3
2	P	1/1 (100%)	1 (100%)	0	100 100
All	All	281/281 (100%)	241 (86%)	40 (14%)	3 3

5 of 40 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	242	TYR
1	E	284	SER
1	E	245	PHE
1	E	253	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	297	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	144	HIS
1	E	148	GLN
1	E	318	ASN
1	E	251	ASN
1	E	277	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)

3 non-standard protein/DNA/RNA residues 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	$\#\|Z\| > 2$
2	SMC	P	2	2	5,6,7	0.48	0	2,6,8	6.04	2 (100%)
2	PHI	P	3	2	11,12,13	8.96	1 (9%)	12,15,17	2.63	6 (50%)
2	NOR	P	1	2	17,17,17	2.80	1 (5%)	18,22,22	2.61	5 (27%)

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
2	SMC	P	2	2	-	2/3/5/7	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	PHI	P	3	2	-	1/5/6/8	0/1/1/1
2	NOR	P	1	2	-	7/16/24/24	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	P	3	PHI	CZ-I	-29.63	1.36	2.10
2	P	1	NOR	CH-CA	-11.05	1.41	1.54

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	P	1	NOR	CE2-CD2-CG	-7.05	98.82	112.15
2	P	2	SMC	CA-CB-SG	-7.05	102.64	114.04
2	P	3	PHI	CE2-CZ-CE1	-5.51	112.95	120.63
2	P	1	NOR	CM-OXT-C	-4.99	111.84	117.83
2	P	2	SMC	CS-SG-CB	4.83	110.18	101.30

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	P	3	PHI	O-C-CA-CB
2	P	2	SMC	N-CA-CB-SG
2	P	1	NOR	CA-CB-CG-CD1
2	P	1	NOR	CA-CB-CG-CD2
2	P	1	NOR	CM2-CM-OXT-C

There are no ring outliers.

3 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	P	2	SMC	1	0
2	P	3	PHI	8	0
2	P	1	NOR	5	0

5.5 Carbohydrates (i)

There are no monosaccharides 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.