wwPDB X-ray Structure Validation Summary Report (i)

Mar 24, 2026 – 06:22 PM UTC

PDB ID	:	3IEC / pdb_00003iec
Title	:	Helicobacter pylori CagA Inhibits PAR1/MARK Family Kinases by Mimicking Host Substrates
Authors	:	Stebbins, C.E.; Nesic, D.; Miller, M.
Deposited on	:	2009-07-22
Resolution	:	2.20 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 (1)) were used in the production of this report:

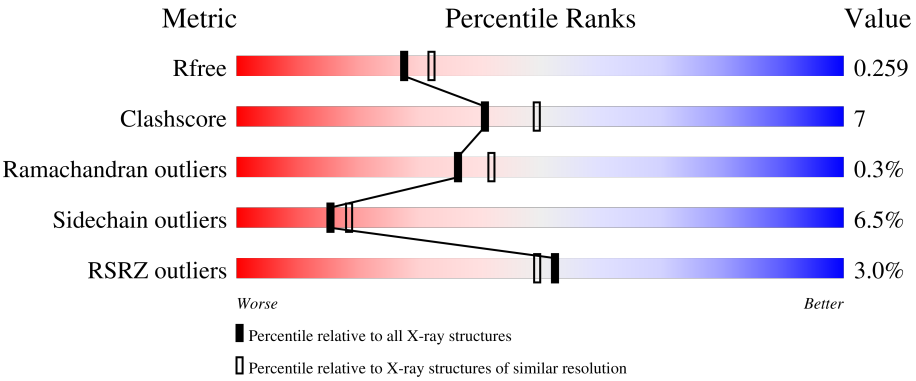
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

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 2.20 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))
R_{free}	180053	6164 (2.20-2.20)
Clashscore	190562	6851 (2.20-2.20)
Ramachandran outliers	187476	6768 (2.20-2.20)
Sidechain outliers	187428	6769 (2.20-2.20)
RSRZ outliers	180081	6166 (2.20-2.20)

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% 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.

Mol	Chain	Length	Quality of chain
1	A	319	2% 79% 17%
1	B	319	3% 82% 13%
1	C	319	2% 81% 13%
1	D	319	3% 81% 13%
2	E	125	10% 89%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	125	2% 8% 89%
2	G	125	% 10% 89%
2	H	125	2% 9% 89%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 11312 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 Serine/threonine-protein kinase MARK2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	311	Total	C	N	O	S	0	0	0
			2534	1633	429	459	13			
1	B	313	Total	C	N	O	S	0	0	0
			2547	1641	432	461	13			
1	C	311	Total	C	N	O	S	0	0	0
			2534	1633	429	459	13			
1	D	311	Total	C	N	O	S	0	0	0
			2534	1633	429	459	13			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	45	MET	-	expression tag	UNP Q7KZI7
A	46	ALA	-	expression tag	UNP Q7KZI7
A	47	ASP	-	expression tag	UNP Q7KZI7
A	48	LEU	-	expression tag	UNP Q7KZI7
A	146	TRP	ARG	engineered mutation	UNP Q7KZI7
B	45	MET	-	expression tag	UNP Q7KZI7
B	46	ALA	-	expression tag	UNP Q7KZI7
B	47	ASP	-	expression tag	UNP Q7KZI7
B	48	LEU	-	expression tag	UNP Q7KZI7
B	146	TRP	ARG	engineered mutation	UNP Q7KZI7
C	45	MET	-	expression tag	UNP Q7KZI7
C	46	ALA	-	expression tag	UNP Q7KZI7
C	47	ASP	-	expression tag	UNP Q7KZI7
C	48	LEU	-	expression tag	UNP Q7KZI7
C	146	TRP	ARG	engineered mutation	UNP Q7KZI7
D	45	MET	-	expression tag	UNP Q7KZI7
D	46	ALA	-	expression tag	UNP Q7KZI7
D	47	ASP	-	expression tag	UNP Q7KZI7
D	48	LEU	-	expression tag	UNP Q7KZI7
D	146	TRP	ARG	engineered mutation	UNP Q7KZI7

- Molecule 2 is a protein called Cytotoxicity-associated immunodominant antigen.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	E	14	Total	C	N	O	0	0	0
			119	76	22	21			
2	F	14	Total	C	N	O	0	0	0
			119	76	22	21			
2	G	14	Total	C	N	O	0	0	0
			119	76	22	21			
2	H	14	Total	C	N	O	0	0	0
			119	76	22	21			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	881	GLY	-	expression tag	UNP P55980
E	882	PRO	-	expression tag	UNP P55980
E	883	VAL	-	expression tag	UNP P55980
E	884	ASP	-	expression tag	UNP P55980
F	881	GLY	-	expression tag	UNP P55980
F	882	PRO	-	expression tag	UNP P55980
F	883	VAL	-	expression tag	UNP P55980
F	884	ASP	-	expression tag	UNP P55980
G	881	GLY	-	expression tag	UNP P55980
G	882	PRO	-	expression tag	UNP P55980
G	883	VAL	-	expression tag	UNP P55980
G	884	ASP	-	expression tag	UNP P55980
H	881	GLY	-	expression tag	UNP P55980
H	882	PRO	-	expression tag	UNP P55980
H	883	VAL	-	expression tag	UNP P55980
H	884	ASP	-	expression tag	UNP P55980

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	185	Total	O	0	0
			185	185		
3	B	166	Total	O	0	0
			166	166		
3	C	161	Total	O	0	0
			161	161		
3	D	147	Total	O	0	0
			147	147		
3	E	6	Total	O	0	0
			6	6		

Continued on next page...

Continued from previous page...

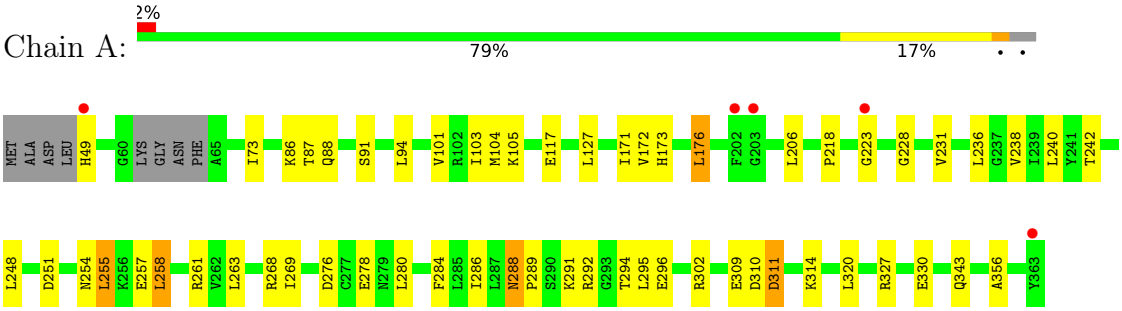
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	F	8	Total O 8 8	0	0
3	G	7	Total O 7 7	0	0
3	H	7	Total O 7 7	0	0

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 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.

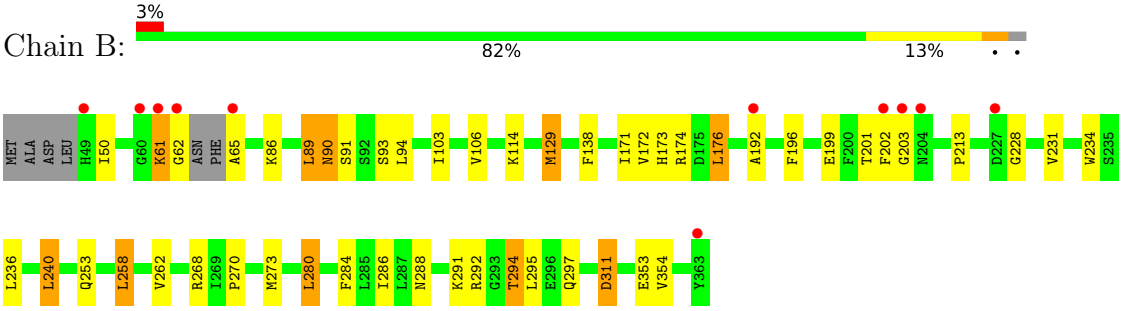
• Molecule 1: Serine/threonine-protein kinase MARK2

Chain A:



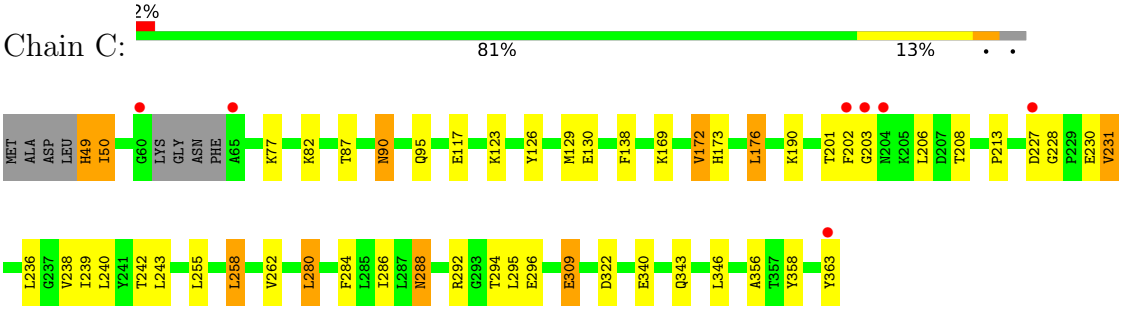
• Molecule 1: Serine/threonine-protein kinase MARK2

Chain B:



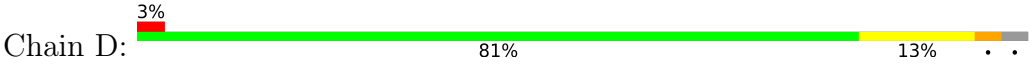
• Molecule 1: Serine/threonine-protein kinase MARK2

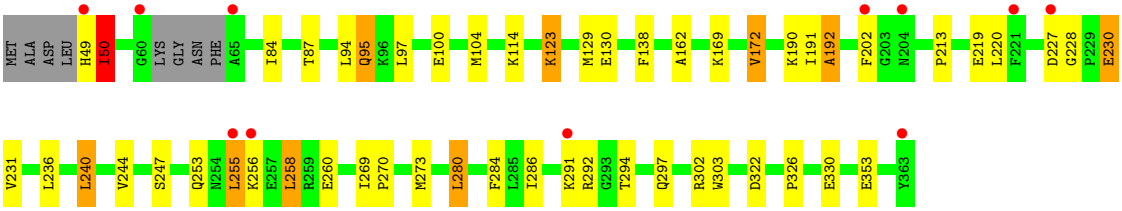
Chain C:



• Molecule 1: Serine/threonine-protein kinase MARK2

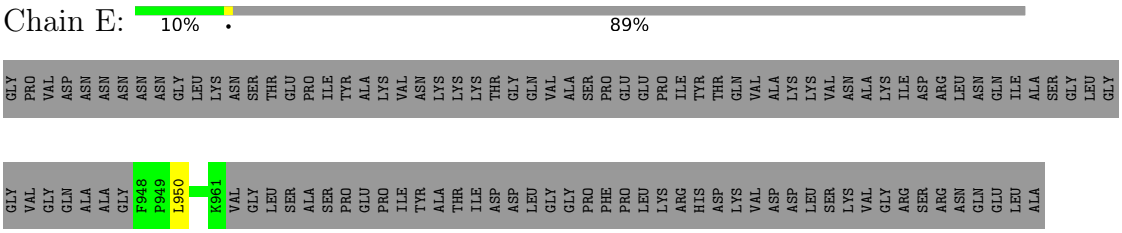
Chain D:





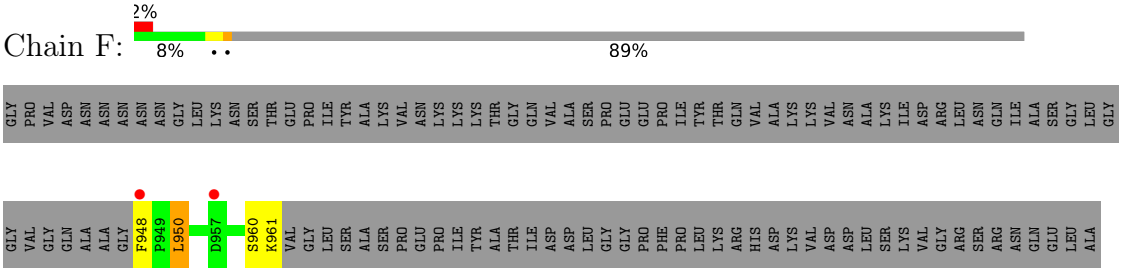
• Molecule 2: Cytotoxicity-associated immunodominant antigen

Chain E:



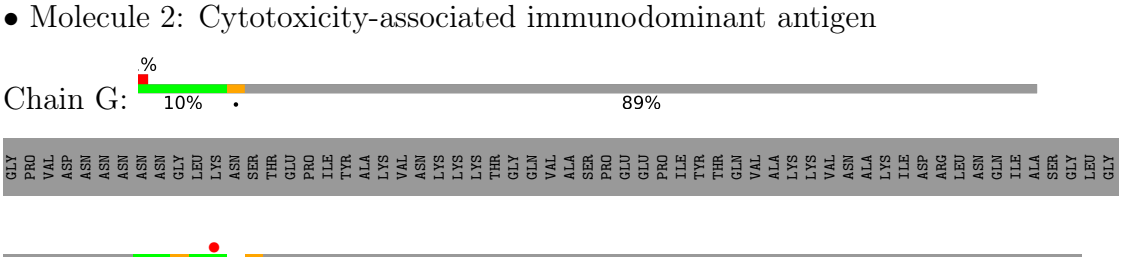
• Molecule 2: Cytotoxicity-associated immunodominant antigen

Chain F:



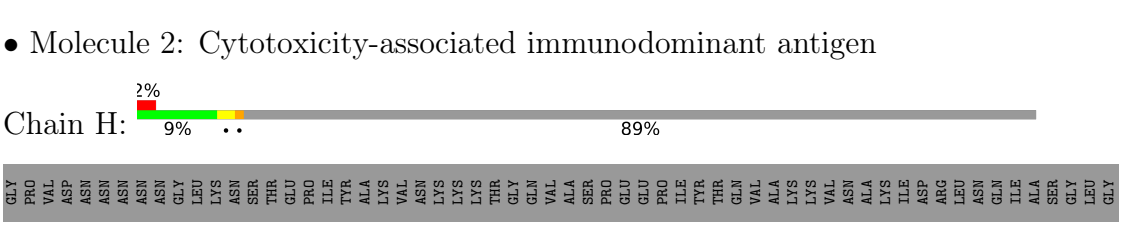
• Molecule 2: Cytotoxicity-associated immunodominant antigen

Chain G:



• Molecule 2: Cytotoxicity-associated immunodominant antigen

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	93.47Å 93.26Å 113.95Å 90.00° 100.94° 90.00°	Depositor
Resolution (Å)	40.96 – 2.20 40.96 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.8 (40.96-2.20) 99.8 (40.96-2.20)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.58 (at 2.20Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.204 , 0.248 0.219 , 0.259	Depositor DCC
R_{free} test set	4906 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	38.5	Xtriage
Anisotropy	0.102	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 43.2	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	0.94	EDS
Total number of atoms	11312	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.26% 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

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	A	0.91	0/2587	0.92	0/3485
1	B	0.88	1/2600 (0.0%)	0.93	4/3501 (0.1%)
1	C	0.89	0/2587	0.89	3/3485 (0.1%)
1	D	0.86	0/2587	0.91	4/3485 (0.1%)
2	E	0.74	0/121	0.88	0/160
2	F	0.99	0/121	1.23	2/160 (1.2%)
2	G	0.67	0/121	0.76	0/160
2	H	0.72	0/121	0.87	0/160
All	All	0.88	1/10845 (0.0%)	0.92	13/14596 (0.1%)

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	A	0	1
1	D	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	192	ALA	C-O	6.95	1.33	1.23

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	192	ALA	N-CA-C	8.22	123.86	112.93
1	C	213	PRO	N-CA-C	6.22	118.29	110.70
2	F	948	PHE	CA-C-N	-6.05	113.73	120.14
2	F	948	PHE	C-N-CA	-6.05	113.73	120.14
1	B	213	PRO	N-CA-C	5.92	117.92	110.70

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	223	GLY	Peptide
1	D	192	ALA	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	A	2534	0	2570	43	0
1	B	2547	0	2586	27	0
1	C	2534	0	2570	32	0
1	D	2534	0	2570	37	0
2	E	119	0	122	0	0
2	F	119	0	122	1	0
2	G	119	0	122	2	0
2	H	119	0	122	3	0
3	A	185	0	0	9	0
3	B	166	0	0	4	1
3	C	161	0	0	4	0
3	D	147	0	0	2	0
3	E	6	0	0	0	0
3	F	8	0	0	0	0
3	G	7	0	0	1	1
3	H	7	0	0	1	0
All	All	11312	0	10784	140	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 7.

The worst 5 of 140 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)
1:C:286:ILE:O	1:C:292:ARG:NH1	1.78	1.16
1:B:286:ILE:O	1:B:292:ARG:NH1	1.89	1.05
1:B:294:THR:HG23	1:B:297:GLN:HE21	1.32	0.91

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:103:ILE:HG23	1:B:171:ILE:HD13	1.57	0.85
1:A:302:ARG:NE	3:A:463:HOH:O	2.17	0.77

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)
3:B:382:HOH:O	3:G:153:HOH:O[1_655]	2.08	0.12

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	307/319 (96%)	296 (96%)	11 (4%)	0	100	100
1	B	309/319 (97%)	301 (97%)	7 (2%)	1 (0%)	36	42
1	C	307/319 (96%)	298 (97%)	8 (3%)	1 (0%)	36	42
1	D	307/319 (96%)	297 (97%)	9 (3%)	1 (0%)	36	42
2	E	12/125 (10%)	11 (92%)	1 (8%)	0	100	100
2	F	12/125 (10%)	11 (92%)	0	1 (8%)	0	0
2	G	12/125 (10%)	11 (92%)	1 (8%)	0	100	100
2	H	12/125 (10%)	11 (92%)	1 (8%)	0	100	100
All	All	1278/1776 (72%)	1236 (97%)	38 (3%)	4 (0%)	36	42

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	202	PHE
1	C	202	PHE
1	D	202	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	F	960	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	277/283 (98%)	260 (94%)	17 (6%)	17	20
1	B	278/283 (98%)	261 (94%)	17 (6%)	17	20
1	C	277/283 (98%)	258 (93%)	19 (7%)	14	16
1	D	277/283 (98%)	260 (94%)	17 (6%)	17	20
2	E	14/103 (14%)	13 (93%)	1 (7%)	13	16
2	F	14/103 (14%)	12 (86%)	2 (14%)	3	3
2	G	14/103 (14%)	12 (86%)	2 (14%)	3	3
2	H	14/103 (14%)	13 (93%)	1 (7%)	13	16
All	All	1165/1544 (76%)	1089 (94%)	76 (6%)	15	18

5 of 76 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	114	LYS
2	F	950	LEU
1	D	169	LYS
1	D	255	LEU
2	H	950	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	180	ASN
1	C	297	GLN
1	C	288	ASN
1	D	95	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	180	ASN

5.3.3 RNA

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 oligosaccharides 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

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	311/319 (97%)	-0.13	5 (1%) 70 67	16, 24, 37, 47	0
1	B	313/319 (98%)	0.08	11 (3%) 47 44	15, 25, 40, 61	0
1	C	311/319 (97%)	0.02	7 (2%) 61 58	13, 23, 36, 47	0
1	D	311/319 (97%)	0.31	11 (3%) 47 44	17, 25, 39, 48	0
2	E	14/125 (11%)	0.04	0 100 100	16, 19, 22, 23	0
2	F	14/125 (11%)	0.49	2 (14%) 6 4	13, 16, 17, 18	0
2	G	14/125 (11%)	0.69	1 (7%) 22 19	9, 16, 18, 19	0
2	H	14/125 (11%)	0.96	2 (14%) 6 4	30, 41, 49, 50	0
All	All	1302/1776 (73%)	0.09	39 (2%) 52 49	9, 24, 39, 61	0

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	202	PHE	4.6
1	D	202	PHE	4.4
1	C	363	TYR	4.1
1	B	203	GLY	4.1
1	B	202	PHE	3.9

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

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

6.3 Carbohydrates

There are no oligosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.