wwPDB X-ray Structure Validation Summary Report (i)

Oct 11, 2021 – 03:40 AM EDT

PDB ID	:	3C6L
Title	:	Crystal structure of mouse MHC class II I-Ab/3K peptide complexed with mouse TCR 2W20
Authors	:	Dai, S.
Deposited on	:	2008-02-04
Resolution	:	3.40 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:

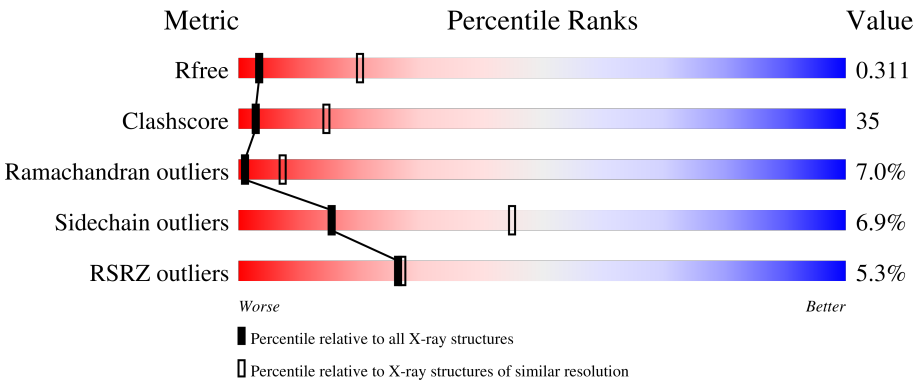
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 3.40 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}	130704	1026 (3.48-3.32)
Clashscore	141614	1055 (3.48-3.32)
Ramachandran outliers	138981	1038 (3.48-3.32)
Sidechain outliers	138945	1038 (3.48-3.32)
RSRZ outliers	127900	2173 (3.50-3.30)

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% 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	185	11% 49% 41% 8% •
1	E	185	13% 47% 43% 9% ••
2	B	236	5% 50% 39% 7% • •
2	F	236	3% 52% 39% 8%
3	C	182	2% 46% 48% 5% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	182	3% 51% 46%
4	D	217	33% 45% 8% 12%
4	H	217	4% 35% 46% 7% 12%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 12716 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 TCR 2W20 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	185	1476	935	243	291	7	0	0	0
1	E	184	1464	926	242	289	7	0	0	0

- Molecule 2 is a protein called TCR 2W20 beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	227	1797	1129	319	343	6	0	0	0
2	F	236	1870	1175	330	359	6	0	0	0

- Molecule 3 is a protein called H-2 class II histocompatibility antigen, A-B alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	182	1459	944	230	282	3	0	0	0
3	G	182	1459	944	230	282	3	0	0	0

- Molecule 4 is a protein called 3K peptide, Linker,and H-2 class II histocompatibility antigen (A beta chain).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	192	1594	1002	288	297	7	0	0	0
4	H	192	1594	1002	288	297	7	0	0	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	14	GLY	-	linker	UNP P14483
D	15	GLY	-	linker	UNP P14483
D	16	GLY	-	linker	UNP P14483
D	17	GLY	-	linker	UNP P14483
D	18	SER	-	linker	UNP P14483
D	19	LEU	-	linker	UNP P14483
D	20	VAL	-	linker	UNP P14483
D	21	PRO	-	linker	UNP P14483
D	22	ARG	-	linker	UNP P14483
D	23	GLY	-	linker	UNP P14483
D	24	SER	-	linker	UNP P14483
D	25	GLY	-	linker	UNP P14483
D	26	GLY	-	linker	UNP P14483
D	27	GLY	-	linker	UNP P14483
D	28	GLY	-	linker	UNP P14483
D	216	LYS	ARG	engineered mutation	UNP P14483
H	14	GLY	-	linker	UNP P14483
H	15	GLY	-	linker	UNP P14483
H	16	GLY	-	linker	UNP P14483
H	17	GLY	-	linker	UNP P14483
H	18	SER	-	linker	UNP P14483
H	19	LEU	-	linker	UNP P14483
H	20	VAL	-	linker	UNP P14483
H	21	PRO	-	linker	UNP P14483
H	22	ARG	-	linker	UNP P14483
H	23	GLY	-	linker	UNP P14483
H	24	SER	-	linker	UNP P14483
H	25	GLY	-	linker	UNP P14483
H	26	GLY	-	linker	UNP P14483
H	27	GLY	-	linker	UNP P14483
H	28	GLY	-	linker	UNP P14483
H	216	LYS	ARG	engineered mutation	UNP P14483

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

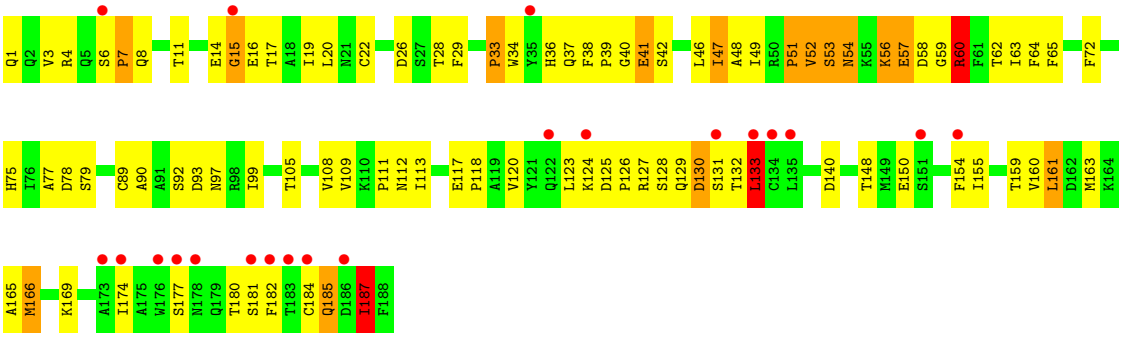
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
5	A	1	1	1	0	0
			Total	Ca		
5	E	2	2	2	0	0

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.

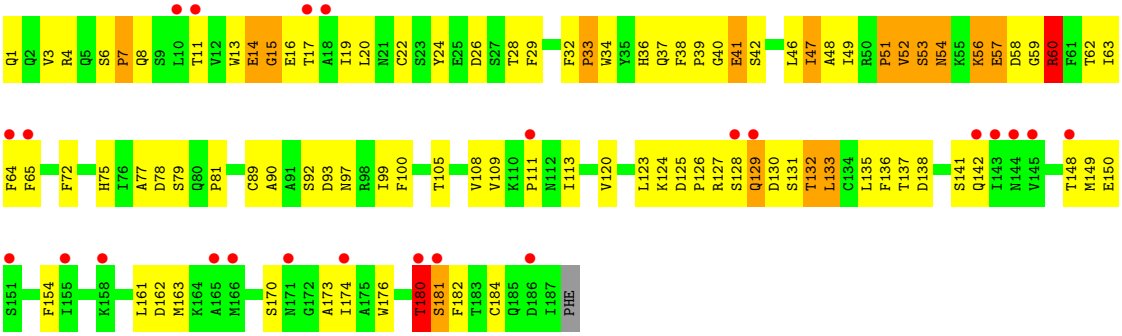
• Molecule 1: TCR 2W20 alpha chain

Chain A:



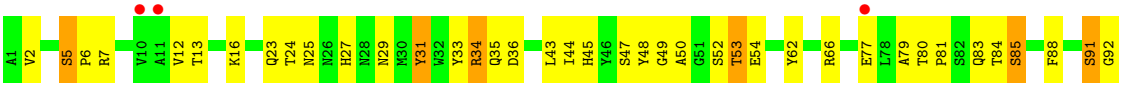
• Molecule 1: TCR 2W20 alpha chain

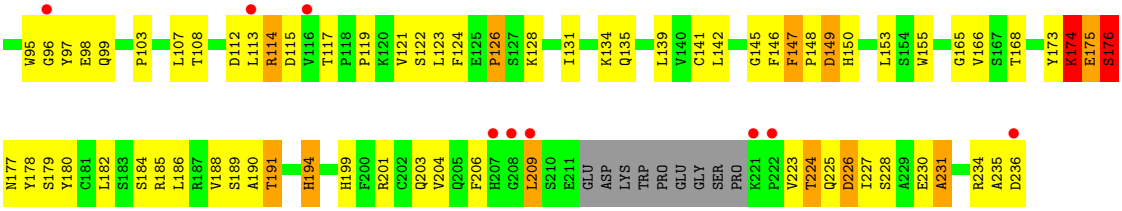
Chain E:



• Molecule 2: TCR 2W20 beta chain

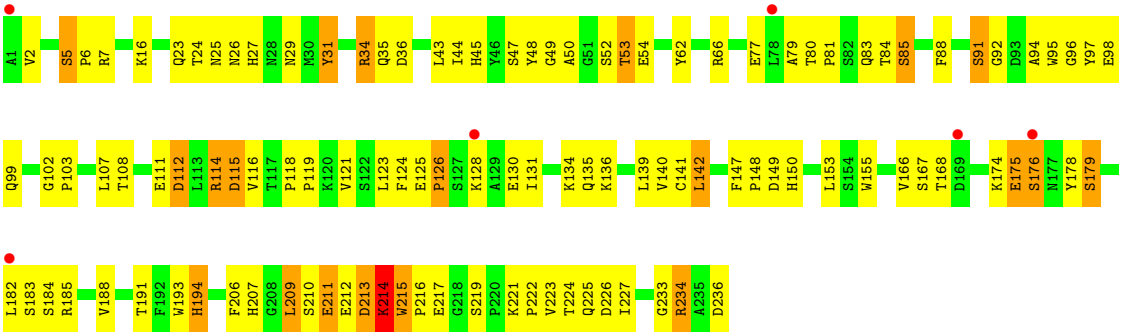
Chain B:





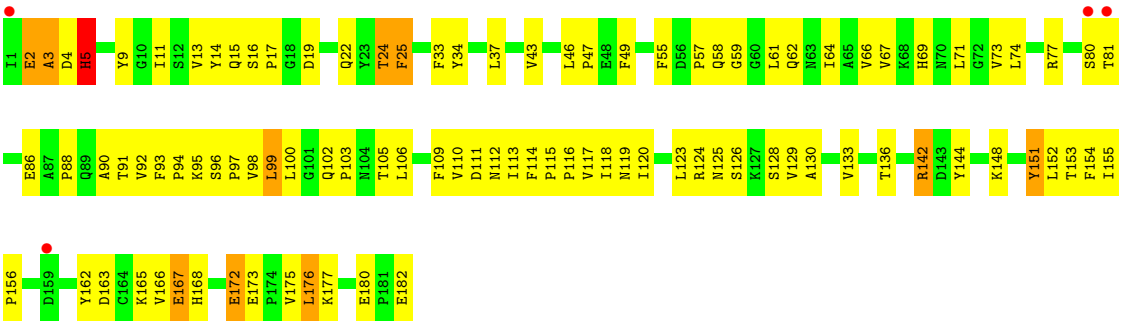
• Molecule 2: TCR 2W20 beta chain

Chain F:



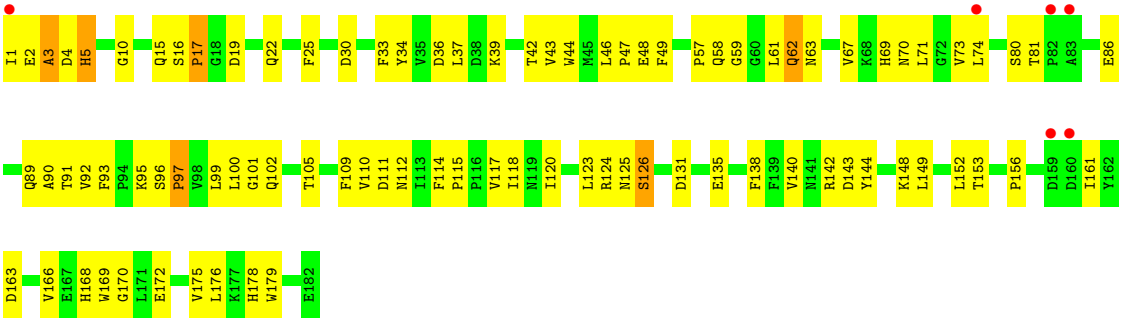
• Molecule 3: H-2 class II histocompatibility antigen, A-B alpha chain

Chain C:



• Molecule 3: H-2 class II histocompatibility antigen, A-B alpha chain

Chain G:



• Molecule 4: 3K peptide, Linker,and H-2 class II histocompatibility antigen (A beta chain)

Chain D: 33% 45% 8% • 12%

F1 E2 A3 Q4 K5 A8 N9 K10 A11 V12 ASP GLY GLY GLY GLY SER LEU VAL PRO ARG GLY SER GLY GLY GLY GLY SER GLU R31 H32 F33 V34 Y35 Q36 F37 M38 G39 E40 F43 T44 N45 Q48 R51 Y52 R55 Y56 I57 Y58 N59 R60 E61 E62 Y63 V64 R65

Y66 D67 S68 E72 H73 V76 T77 E78 L79 G80 R81 P82 N88 E92 R96 T97 E100 T103 V104 H107 N108 Y109 E110 G111 P112 E113 T114 H115 T116 S117 L118 R119 R120 L121 E122 Q123 P124 N125 V126 V127 I128 S129 L130 S131 ARG THR GLU ALA LEU ASN HIS H139

C144 T147 D148 F149 Y150 P151 I154 K155 V156 R157 W158 F159 R160 N161 G162 Q163 E164 E165 T166 V167 G168 V169 S170 S171 T172 Q173 L174 I175 R176 N177 G178 T181 F182 Q183 V184 L185 V186 M187 L188 E189 M190 T191 P192 R193 R194 V197 Y198 T199 C200 H201 V202 E203 H204 P205 S206

L207 P210 I211 T212 V213 E214 W215 K216 A217

• Molecule 4: 3K peptide, Linker,and H-2 class II histocompatibility antigen (A beta chain)

Chain H: 4% 35% 46% 7% • 12%

F1 E2 A3 Q4 K5 K10 A11 V12 ASP GLY GLY GLY GLY SER LEU VAL PRO ARG GLY SER GLY GLY GLY GLY SER GLU R31 H32 F33 V34 Y35 Q36 E40 T44 N45 G46 T47 Q48 Y52 V53 T54 R55 Y56 I57 Y58 N59 R60 E61 E62 Y63 V64 R65 Y66 D67 S68 D69

V70 H73 V76 T77 E78 R81 P82 N88 L94 E95 R96 T97 E100 L101 D102 T103 V104 C105 R106 H107 N108 P112 E113 T114 H115 T116 S117 R120 L121 E122 Q123 P124 N125 V126 V127 I128 S129 L130 S131 ARG THR GLU ALA LEU ASN HIS H139 N140 T141 L142 V143

D148 F149 Y150 P151 A152 K153 I154 K155 V156 R160 N161 G162 Q163 E164 E165 T166 V167 G168 V169 S170 S171 T172 Q173 L174 D179 W180 T181 V184 L185 V186 M187 L188 E189 M190 T191 P192 R193 R194 G195 E196 V197 Y198 T199 C200 H201 V202 E203 H204 P205 S206 L207 K208 S209 P210 I211 K216

A217

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	46.54Å 113.93Å 386.26Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.07 – 3.40 49.06 – 3.35	Depositor EDS
% Data completeness (in resolution range)	97.7 (49.07-3.40) 97.2 (49.06-3.35)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.79 (at 3.33Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.268 , 0.325 0.258 , 0.311	Depositor DCC
R_{free} test set	1482 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	98.7	Xtriage
Anisotropy	0.389	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 71.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	12716	wwPDB-VP
Average B, all atoms (Å^2)	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.86% 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: CA

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	0.55	0/1514	0.76	2/2053 (0.1%)
1	E	0.52	0/1501	0.73	0/2037
2	B	0.53	0/1847	0.70	1/2509 (0.0%)
2	F	0.52	0/1925	0.69	0/2618
3	C	0.67	0/1504	0.86	1/2054 (0.0%)
3	G	0.67	1/1504 (0.1%)	0.81	1/2054 (0.0%)
4	D	0.67	0/1632	0.87	2/2210 (0.1%)
4	H	0.68	0/1632	0.82	0/2210
All	All	0.60	1/13059 (0.0%)	0.78	7/17745 (0.0%)

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
3	C	0	2
4	D	0	1
4	H	0	1
All	All	0	4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	G	44	TRP	CB-CG	-5.51	1.40	1.50

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	161	LEU	CA-CB-CG	6.31	129.82	115.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	187	ILE	N-CA-C	6.11	127.50	111.00
3	C	5	HIS	N-CA-C	5.67	126.29	111.00
2	B	174	LYS	N-CA-C	5.51	125.88	111.00
3	G	5	HIS	N-CA-C	5.44	125.68	111.00

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	144	TYR	Sidechain
3	C	9	TYR	Sidechain
4	D	35	TYR	Sidechain
4	H	150	TYR	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	1476	0	1402	86	0
1	E	1464	0	1394	102	0
2	B	1797	0	1703	125	0
2	F	1870	0	1765	133	0
3	C	1459	0	1386	120	0
3	G	1459	0	1386	101	0
4	D	1594	0	1534	169	0
4	H	1594	0	1534	156	0
5	A	1	0	0	0	0
5	E	2	0	0	0	0
All	All	12716	0	12104	876	0

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 35.

The worst 5 of 876 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:3:ALA:HB1	4:H:164:GLU:H	0.97	1.10
3:C:109:PHE:HE2	3:C:111:ASP:HB2	1.16	1.08
3:C:2:GLU:HA	4:D:45:ASN:OD1	1.55	1.05
3:C:71:LEU:HD23	4:D:35:TYR:HB2	1.40	1.03
4:H:116:THR:HG23	4:H:117:SER:H	1.17	1.02

There are no symmetry-related clashes.

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	183/185 (99%)	135 (74%)	29 (16%)	19 (10%)	0 3
1	E	182/185 (98%)	135 (74%)	30 (16%)	17 (9%)	0 4
2	B	223/236 (94%)	177 (79%)	31 (14%)	15 (7%)	1 8
2	F	234/236 (99%)	189 (81%)	27 (12%)	18 (8%)	1 6
3	C	180/182 (99%)	154 (86%)	19 (11%)	7 (4%)	3 19
3	G	180/182 (99%)	149 (83%)	24 (13%)	7 (4%)	3 19
4	D	186/217 (86%)	151 (81%)	20 (11%)	15 (8%)	1 5
4	H	186/217 (86%)	150 (81%)	25 (13%)	11 (6%)	1 11
All	All	1554/1640 (95%)	1240 (80%)	205 (13%)	109 (7%)	1 7

5 of 109 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	52	VAL
1	A	53	SER
1	A	57	GLU
1	A	60	ARG
1	A	128	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	166/166 (100%)	155 (93%)	11 (7%)	16	46
1	E	165/166 (99%)	154 (93%)	11 (7%)	16	46
2	B	194/202 (96%)	179 (92%)	15 (8%)	13	40
2	F	202/202 (100%)	188 (93%)	14 (7%)	15	45
3	C	163/163 (100%)	155 (95%)	8 (5%)	25	55
3	G	163/163 (100%)	154 (94%)	9 (6%)	21	51
4	D	174/189 (92%)	161 (92%)	13 (8%)	13	41
4	H	174/189 (92%)	158 (91%)	16 (9%)	9	31
All	All	1401/1440 (97%)	1304 (93%)	97 (7%)	15	45

5 of 97 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	180	THR
2	F	234	ARG
2	F	31	TYR
2	F	142	LEU
3	G	62	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 55 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	75	HIS
2	F	29	ASN
4	H	177	ASN
4	H	9	ASN
1	E	129	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)

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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)

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, 95th 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	185/185 (100%)	0.62	21 (11%)	5	6	44, 96, 176, 198	0
1	E	184/185 (99%)	0.76	24 (13%)	3	4	61, 124, 174, 194	0
2	B	227/236 (96%)	0.36	12 (5%)	26	27	50, 98, 156, 186	0
2	F	236/236 (100%)	0.22	6 (2%)	57	55	52, 98, 152, 173	0
3	C	182/182 (100%)	0.15	4 (2%)	62	60	38, 73, 136, 182	0
3	G	182/182 (100%)	0.26	6 (3%)	46	45	44, 78, 142, 178	0
4	D	192/217 (88%)	-0.00	1 (0%)	91	90	43, 77, 122, 152	0
4	H	192/217 (88%)	0.13	9 (4%)	31	31	45, 85, 127, 168	0
All	All	1580/1640 (96%)	0.31	83 (5%)	26	27	38, 91, 160, 198	0

The worst 5 of 83 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	177	SER	10.2
1	A	182	PHE	7.4
3	G	159	ASP	6.2
2	B	222	PRO	5.8
1	E	166	MET	5.2

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	CA	E	190	1/1	0.92	0.22	85,85,85,85	0
5	CA	A	189	1/1	0.95	0.29	45,45,45,45	0
5	CA	E	189	1/1	0.97	0.15	73,73,73,73	0

6.5 Other polymers (i)

There are no such residues in this entry.