



Full wwPDB X-ray Structure Validation Report ⓘ

May 21, 2020 – 07:08 pm BST

PDB ID : 4CW1
Title : COMPLEX OF A B14 CHICKEN MHC CLASS I MOLECULE AND A 9MER CHICKEN PEPTIDE
Authors : Chappell, P.E.; Roversi, P.; Harrison, M.C.; Mears, L.E.; Kaufman, J.F.; Lea, S.M.
Deposited on : 2014-03-31
Resolution : 2.58 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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

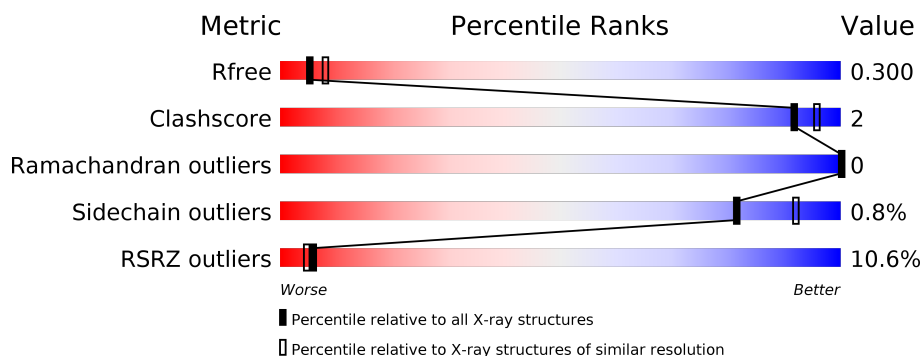
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.58 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	3676 (2.60-2.56)
Clashscore	141614	4049 (2.60-2.56)
Ramachandran outliers	138981	3979 (2.60-2.56)
Sidechain outliers	138945	3979 (2.60-2.56)
RSRZ outliers	127900	3614 (2.60-2.56)

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 $\leq 5\%$. 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	310	<div> <div>9%</div> <div>82%</div> <div>5%</div> <div>13%</div> </div>
1	D	310	<div> <div>10%</div> <div>82%</div> <div>5%</div> <div>13%</div> </div>
2	B	98	<div> <div>10%</div> <div>96%</div> <div>.</div> </div>
2	E	98	<div> <div>9%</div> <div>97%</div> <div>..</div> </div>
3	C	9	<div> <div>22%</div> <div>78%</div> <div>22%</div> </div>
3	F	9	<div> <div>11%</div> <div>100%</div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 6119 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 MAJOR HISTOCOMPATIBILITY COMPLEX CLASS I GLY-COPROTEIN HAPLOTYPE B14.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	271	Total	C	N	O	S	0	0	0
			2168	1359	383	418	8			
1	D	271	Total	C	N	O	S	0	0	0
			2160	1356	380	416	8			

There are 76 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	273	ARG	-	expression tag	UNP A0ZXM3
A	274	SER	-	expression tag	UNP A0ZXM3
A	275	GLY	-	expression tag	UNP A0ZXM3
A	276	GLY	-	expression tag	UNP A0ZXM3
A	277	GLY	-	expression tag	UNP A0ZXM3
A	278	LEU	-	expression tag	UNP A0ZXM3
A	279	ASN	-	expression tag	UNP A0ZXM3
A	280	ASP	-	expression tag	UNP A0ZXM3
A	281	ILE	-	expression tag	UNP A0ZXM3
A	282	PHE	-	expression tag	UNP A0ZXM3
A	283	GLU	-	expression tag	UNP A0ZXM3
A	284	ALA	-	expression tag	UNP A0ZXM3
A	285	GLN	-	expression tag	UNP A0ZXM3
A	286	LYS	-	expression tag	UNP A0ZXM3
A	287	ILE	-	expression tag	UNP A0ZXM3
A	288	GLU	-	expression tag	UNP A0ZXM3
A	289	TRP	-	expression tag	UNP A0ZXM3
A	290	HIS	-	expression tag	UNP A0ZXM3
A	291	GLU	-	expression tag	UNP A0ZXM3
A	292	ASN	-	expression tag	UNP A0ZXM3
A	293	SER	-	expression tag	UNP A0ZXM3
A	294	SER	-	expression tag	UNP A0ZXM3
A	295	SER	-	expression tag	UNP A0ZXM3
A	296	VAL	-	expression tag	UNP A0ZXM3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	297	ASP	-	expression tag	UNP A0ZXM3
A	298	LYS	-	expression tag	UNP A0ZXM3
A	299	LEU	-	expression tag	UNP A0ZXM3
A	300	ALA	-	expression tag	UNP A0ZXM3
A	301	ALA	-	expression tag	UNP A0ZXM3
A	302	ALA	-	expression tag	UNP A0ZXM3
A	303	LEU	-	expression tag	UNP A0ZXM3
A	304	GLU	-	expression tag	UNP A0ZXM3
A	305	HIS	-	expression tag	UNP A0ZXM3
A	306	HIS	-	expression tag	UNP A0ZXM3
A	307	HIS	-	expression tag	UNP A0ZXM3
A	308	HIS	-	expression tag	UNP A0ZXM3
A	309	HIS	-	expression tag	UNP A0ZXM3
A	310	HIS	-	expression tag	UNP A0ZXM3
D	273	ARG	-	expression tag	UNP A0ZXM3
D	274	SER	-	expression tag	UNP A0ZXM3
D	275	GLY	-	expression tag	UNP A0ZXM3
D	276	GLY	-	expression tag	UNP A0ZXM3
D	277	GLY	-	expression tag	UNP A0ZXM3
D	278	LEU	-	expression tag	UNP A0ZXM3
D	279	ASN	-	expression tag	UNP A0ZXM3
D	280	ASP	-	expression tag	UNP A0ZXM3
D	281	ILE	-	expression tag	UNP A0ZXM3
D	282	PHE	-	expression tag	UNP A0ZXM3
D	283	GLU	-	expression tag	UNP A0ZXM3
D	284	ALA	-	expression tag	UNP A0ZXM3
D	285	GLN	-	expression tag	UNP A0ZXM3
D	286	LYS	-	expression tag	UNP A0ZXM3
D	287	ILE	-	expression tag	UNP A0ZXM3
D	288	GLU	-	expression tag	UNP A0ZXM3
D	289	TRP	-	expression tag	UNP A0ZXM3
D	290	HIS	-	expression tag	UNP A0ZXM3
D	291	GLU	-	expression tag	UNP A0ZXM3
D	292	ASN	-	expression tag	UNP A0ZXM3
D	293	SER	-	expression tag	UNP A0ZXM3
D	294	SER	-	expression tag	UNP A0ZXM3
D	295	SER	-	expression tag	UNP A0ZXM3
D	296	VAL	-	expression tag	UNP A0ZXM3
D	297	ASP	-	expression tag	UNP A0ZXM3
D	298	LYS	-	expression tag	UNP A0ZXM3
D	299	LEU	-	expression tag	UNP A0ZXM3
D	300	ALA	-	expression tag	UNP A0ZXM3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	301	ALA	-	expression tag	UNP A0ZXM3
D	302	ALA	-	expression tag	UNP A0ZXM3
D	303	LEU	-	expression tag	UNP A0ZXM3
D	304	GLU	-	expression tag	UNP A0ZXM3
D	305	HIS	-	expression tag	UNP A0ZXM3
D	306	HIS	-	expression tag	UNP A0ZXM3
D	307	HIS	-	expression tag	UNP A0ZXM3
D	308	HIS	-	expression tag	UNP A0ZXM3
D	309	HIS	-	expression tag	UNP A0ZXM3
D	310	HIS	-	expression tag	UNP A0ZXM3

- Molecule 2 is a protein called BETA-2-MICROGLOBULIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	98	Total	C	N	O	S	0	0	0
			779	499	126	149	5			
2	E	97	Total	C	N	O	S	0	0	0
			767	490	125	147	5			

- Molecule 3 is a protein called PEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	9	Total	C	N	O	S	0	0	0
			84	55	17	11	1			
3	F	9	Total	C	N	O	S	0	0	0
			84	55	17	11	1			

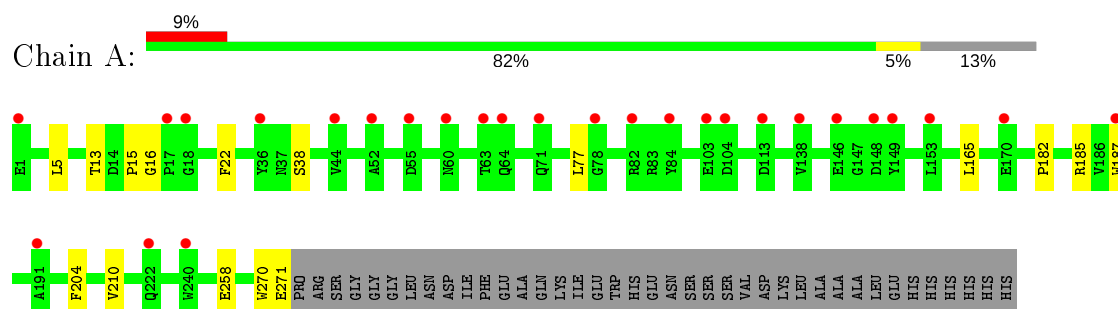
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	34	Total	O	0	0
			34	34		
4	B	9	Total	O	0	0
			9	9		
4	C	2	Total	O	0	0
			2	2		
4	D	23	Total	O	0	0
			23	23		
4	E	8	Total	O	0	0
			8	8		
4	F	1	Total	O	0	0
			1	1		

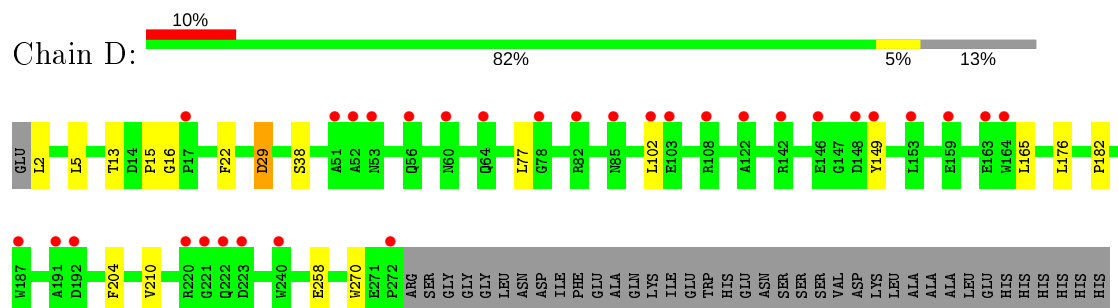
3 Residue-property plots [i](#)

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 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.

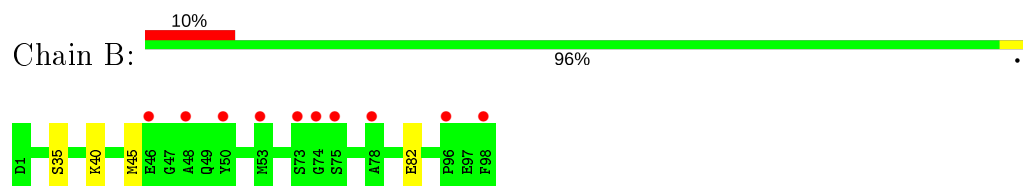
- Molecule 1: MAJOR HISTOCOMPATIBILITY COMPLEX CLASS I GLYCOPROTEIN HAP-LOTYP E B14



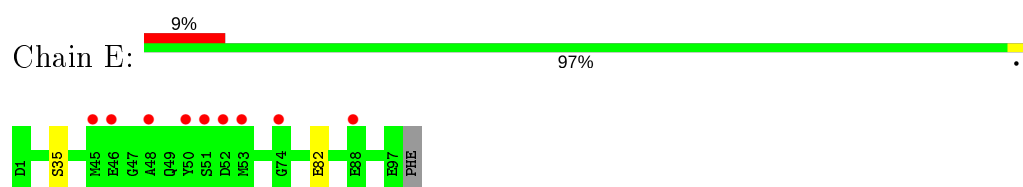
- Molecule 1: MAJOR HISTOCOMPATIBILITY COMPLEX CLASS I GLYCOPROTEIN HAP-LOTYP E B14




- Molecule 2: BETA-2-MICROGLOBULIN

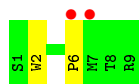


- Molecule 2: BETA-2-MICROGLOBULIN



● Molecule 3: PEPTIDE

Chain C:  22% 78% 22%



● Molecule 3: PEPTIDE

Chain F:  11% 100%



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	62.20 Å 90.58 Å 144.75 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	72.38 – 2.58 72.38 – 2.58	Depositor EDS
% Data completeness (in resolution range)	99.1 (72.38-2.58) 99.1 (72.38-2.58)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.44 (at 2.58 Å)	Xtriage
Refinement program	BUSTER 2.11.4	Depositor
R, R_{free}	0.286 , 0.291 0.291 , 0.300	Depositor DCC
R_{free} test set	1333 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	38.1	Xtriage
Anisotropy	0.569	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 28.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	6119	wwPDB-VP
Average B, all atoms (Å ²)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 45.67 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.2778e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	0.36	0/2226	0.54	0/3029
1	D	0.36	0/2219	0.55	0/3022
2	B	0.36	0/804	0.53	0/1092
2	E	0.36	0/791	0.53	0/1076
3	C	0.33	0/87	0.48	0/115
3	F	0.32	0/87	0.47	0/115
All	All	0.36	0/6214	0.54	0/8449

There are no bond length outliers.

There are no bond angle outliers.

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	2168	0	2030	7	0
1	D	2160	0	2017	9	0
2	B	779	0	737	2	0
2	E	767	0	728	1	0
3	C	84	0	88	1	0
3	F	84	0	88	0	0
4	A	34	0	0	0	0
4	B	9	0	0	0	0
4	C	2	0	0	0	0
4	D	23	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	E	8	0	0	0	0
4	F	1	0	0	0	0
All	All	6119	0	5688	19	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 2.

All (19) 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:A:185:ARG:HD3	1:A:187:TRP:HE1	1.64	0.61
2:B:35:SER:HB3	2:B:82:GLU:HB2	1.86	0.57
2:E:35:SER:HB3	2:E:82:GLU:HB2	1.86	0.56
1:D:210:VAL:HB	1:D:258:GLU:HB2	1.91	0.52
1:A:210:VAL:HB	1:A:258:GLU:HB2	1.91	0.50
1:A:22:PHE:HB3	1:A:38:SER:HB3	1.98	0.46
1:A:5:LEU:HB2	1:A:165:LEU:HD13	1.98	0.46
1:A:182:PRO:HB3	1:A:204:PHE:HB3	1.98	0.46
1:D:5:LEU:HB2	1:D:165:LEU:HD13	1.98	0.45
1:D:22:PHE:HB3	1:D:38:SER:HB3	1.98	0.44
1:D:182:PRO:HB3	1:D:204:PHE:HB3	1.99	0.44
1:A:13:THR:HG23	1:A:77:LEU:HD22	1.99	0.43
1:D:13:THR:HG23	1:D:77:LEU:HD22	1.99	0.43
1:D:2:LEU:HD12	1:D:102:LEU:HA	2.00	0.43
2:B:40:LYS:HB2	2:B:45:MET:HE2	2.01	0.41
1:D:15:PRO:N	1:D:16:GLY:HA2	2.36	0.41
1:D:29:ASP:HB2	1:D:176:LEU:HD22	2.03	0.41
1:A:15:PRO:N	1:A:16:GLY:HA2	2.36	0.40
3:C:6:PRO:HG2	1:D:149:TYR:HB3	2.02	0.40

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	269/310 (87%)	264 (98%)	5 (2%)	0	100	100
1	D	269/310 (87%)	265 (98%)	4 (2%)	0	100	100
2	B	96/98 (98%)	94 (98%)	2 (2%)	0	100	100
2	E	95/98 (97%)	93 (98%)	2 (2%)	0	100	100
3	C	7/9 (78%)	7 (100%)	0	0	100	100
3	F	7/9 (78%)	7 (100%)	0	0	100	100
All	All	743/834 (89%)	730 (98%)	13 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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	223/256 (87%)	221 (99%)	2 (1%)	78	90
1	D	222/256 (87%)	220 (99%)	2 (1%)	78	90
2	B	86/86 (100%)	86 (100%)	0	100	100
2	E	85/86 (99%)	85 (100%)	0	100	100
3	C	9/9 (100%)	8 (89%)	1 (11%)	6	10
3	F	9/9 (100%)	9 (100%)	0	100	100
All	All	634/702 (90%)	629 (99%)	5 (1%)	81	92

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	270	TRP
1	A	271	GLU
3	C	2	TRP
1	D	29	ASP
1	D	270	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 carbohydrates 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, 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	271/310 (87%)	0.89	27 (9%) 7 5	20, 40, 57, 64	0
1	D	271/310 (87%)	0.90	31 (11%) 5 4	22, 40, 57, 66	0
2	B	98/98 (100%)	0.83	10 (10%) 6 5	20, 35, 51, 57	0
2	E	97/98 (98%)	0.69	9 (9%) 8 7	20, 34, 52, 73	0
3	C	9/9 (100%)	1.64	2 (22%) 0 0	37, 44, 55, 75	0
3	F	9/9 (100%)	0.87	1 (11%) 5 4	34, 40, 54, 74	0
All	All	755/834 (90%)	0.87	80 (10%) 6 5	20, 39, 56, 75	0

All (80) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	222	GLN	7.7
2	B	74	GLY	7.3
1	D	272	PRO	4.9
3	C	6	PRO	4.2
1	D	148	ASP	3.8
1	D	220	ARG	3.8
3	C	7	MET	3.7
1	D	149	TYR	3.6
1	A	1	GLU	3.6
1	D	82	ARG	3.5
2	B	98	PHE	3.5
1	D	187	TRP	3.4
1	D	192	ASP	3.4
1	D	52	ALA	3.3
2	E	50	TYR	3.3
1	D	146	GLU	3.3
1	D	191	ALA	3.1
1	A	17	PRO	3.1
1	A	113	ASP	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	60	ASN	3.0
2	B	50	TYR	3.0
2	B	46	GLU	3.0
1	A	240	TRP	3.0
1	D	53	ASN	2.9
1	D	240	TRP	2.9
1	A	82	ARG	2.9
1	A	149	TYR	2.9
1	D	142	ARG	2.8
2	E	46	GLU	2.8
2	E	51	SER	2.8
1	A	104	ASP	2.7
1	A	187	TRP	2.7
2	E	52	ASP	2.7
2	B	78	ALA	2.6
1	D	60	ASN	2.6
2	B	75	SER	2.6
2	E	45	MET	2.6
1	A	44	VAL	2.6
2	E	53	MET	2.5
2	E	74	GLY	2.5
1	A	170	GLU	2.5
1	A	64	GLN	2.5
2	E	88	GLU	2.5
1	A	191	ALA	2.4
1	D	221	GLY	2.4
2	B	53	MET	2.3
1	A	36	TYR	2.3
1	A	103	GLU	2.3
1	D	163	GLU	2.3
1	D	56	GLN	2.2
1	D	102	LEU	2.2
1	A	55	ASP	2.2
1	D	122	ALA	2.2
1	A	71	GLN	2.2
2	B	73	SER	2.2
1	A	78	GLY	2.2
2	B	96	PRO	2.2
1	D	103	GLU	2.2
1	A	153	LEU	2.1
1	A	146	GLU	2.1
3	F	1	SER	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	148	ASP	2.1
1	A	52	ALA	2.1
2	B	48	ALA	2.1
1	D	17	PRO	2.1
1	A	18	GLY	2.1
1	D	64	GLN	2.1
1	A	84	TYR	2.1
1	D	222	GLN	2.1
1	D	159	GLU	2.1
1	D	153	LEU	2.0
1	D	78	GLY	2.0
1	D	223	ASP	2.0
1	A	63	THR	2.0
1	A	138	VAL	2.0
1	D	108	ARG	2.0
1	D	51	ALA	2.0
1	D	164	TRP	2.0
2	E	48	ALA	2.0
1	D	85	ASN	2.0

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 carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.