WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Aug 20, 2023 – 09:40 AM EDT

|		
PDB ID	:	2OJE
Title	:	Mycoplasma arthritidis-derived mitogen complexed with class II MHC molecule HLA-DR1/HA complex in the presence of EDTA
Authors	:	Li, H.; Zhao, Y.; Guo, Y.; Li, Z.; Eisele, L.; Mourad, W.
Deposited on	:	2007-01-12
Resolution	:	3.00 Å(reported)

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 3.00 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.
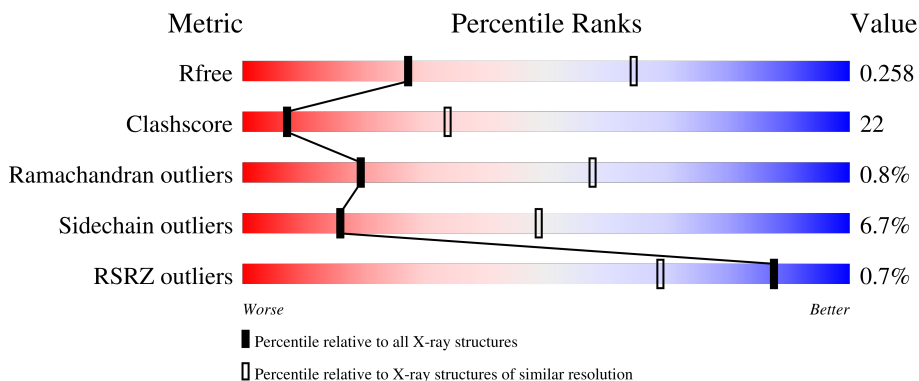
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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	180	61% 34% ··
1	E	180	61% 36% ·
2	B	190	% 58% 35% 6%
2	F	190	3% 55% 42% ·
3	C	13	38% 54% 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	13	54% 46%
4	D	214	60% 35% 5%
4	H	214	59% 39%

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 10010 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 HLA class II histocompatibility antigen, DR alpha chain precursor.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	180	Total	C	N	O	S	Se	0	0	0
			1483	960	241	277	2	3			
1	E	180	Total	C	N	O	S	Se	0	0	0
			1483	960	241	277	2	3			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	23	MSE	MET	modified residue	UNP P01903
A	36	MSE	MET	modified residue	UNP P01903
A	73	MSE	MET	modified residue	UNP P01903
E	23	MSE	MET	modified residue	UNP P01903
E	36	MSE	MET	modified residue	UNP P01903
E	73	MSE	MET	modified residue	UNP P01903

- Molecule 2 is a protein called HLA class II histocompatibility antigen, DRB1-1 beta chain precursor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	190	Total	C	N	O	S	0	0	0
			1557	979	279	293	6			
2	F	190	Total	C	N	O	S	0	0	0
			1557	979	279	293	6			

- Molecule 3 is a protein called haemagglutinin peptide 306-318.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	13	Total	C	N	O	0	0	0
			106	69	18	19			
3	G	13	Total	C	N	O	0	0	0
			102	66	17	19			

- Molecule 4 is a protein called Superantigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	214	Total 1792	C 1157	N 302	O 328	S 5	0	0	0
4	H	214	Total 1792	C 1157	N 302	O 328	S 5	0	0	0

- Molecule 5 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total 5	O 4	P 1	0	0
5	D	1	Total 5	O 4	P 1	0	0
5	E	1	Total 5	O 4	P 1	0	0
5	H	1	Total 5	O 4	P 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	25	Total 25	O 25	0	0
6	B	23	Total 23	O 23	0	0
6	C	1	Total 1	O 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	D	19	Total 19	O 19	0	0
6	E	34	Total 34	O 34	0	0
6	F	4	Total 4	O 4	0	0
6	H	12	Total 12	O 12	0	0

3 Residue-property plots ⓘ

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. 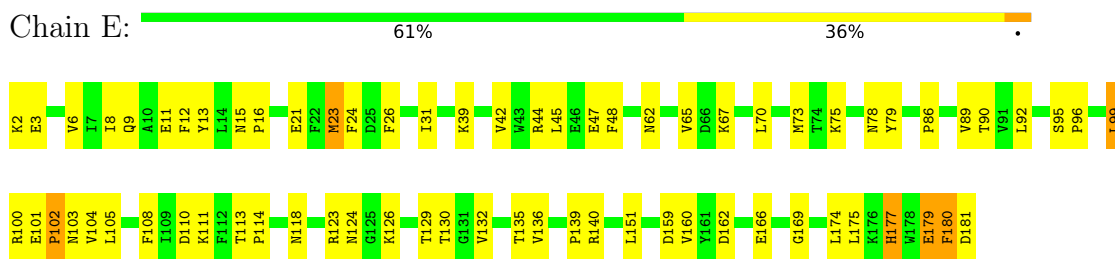A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: HLA class II histocompatibility antigen, DR alpha chain precursor
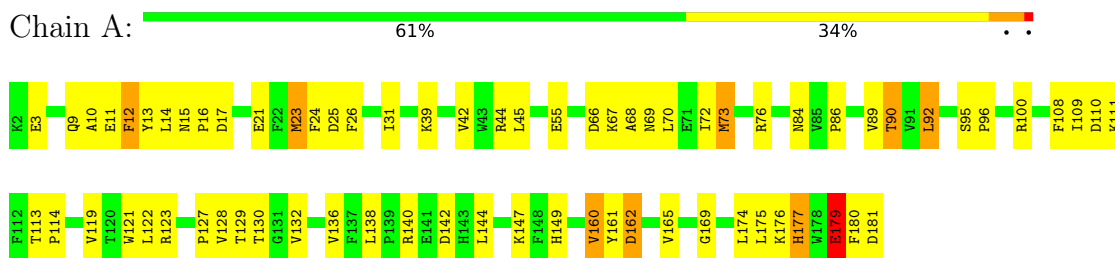
Chain A:

- Molecule 1: HLA class II histocompatibility antigen, DR alpha chain precursor

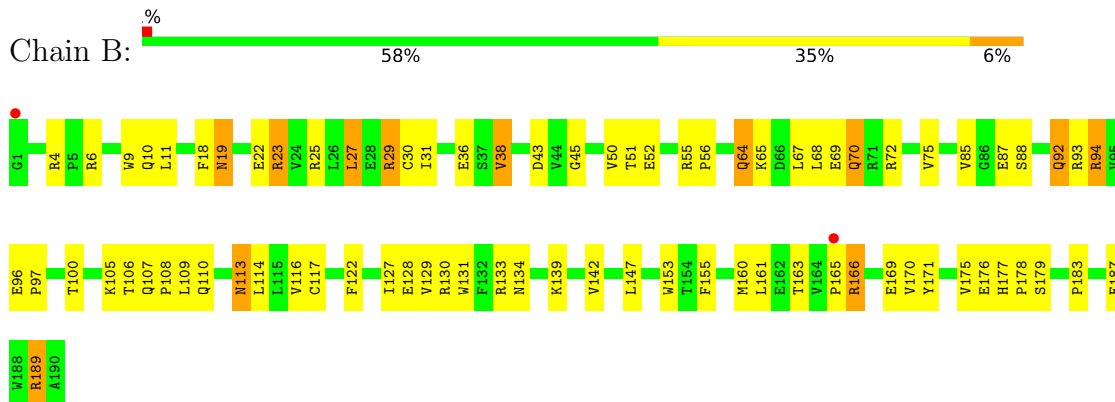Chain E:

- Molecule 2: HLA class II histocompatibility antigen, DRB1-1 beta chain precursor
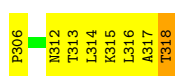
Chain B:

- Molecule 2: HLA class II histocompatibility antigen, DRB1-1 beta chain precursor
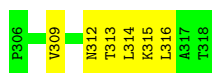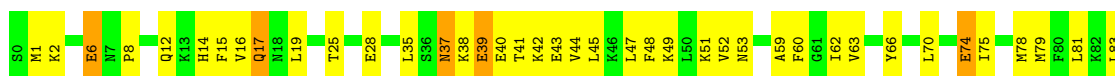
Chain F:

- Molecule 3: haemagglutinin peptide 306-318

Chain C:

- Molecule 3: haemagglutinin peptide 306-318

Chain G:

- Molecule 4: Superantigen

Chain D:

- Molecule 4: Superantigen

Chain H:
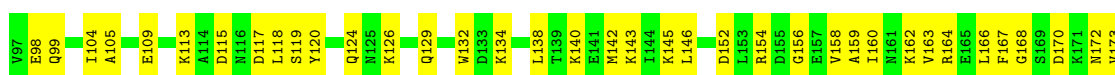
4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	137.74Å 179.56Å 179.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.22 – 3.00 40.22 – 2.99	Depositor EDS
% Data completeness (in resolution range)	89.8 (40.22-3.00) 89.5 (40.22-2.99)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	12.40	Depositor
$< I/\sigma(I) >$ [1]	4.70 (at 3.01Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.211 , 0.261 0.209 , 0.258	Depositor DCC
R_{free} test set	2850 reflections (7.01%)	wwPDB-VP
Wilson B-factor (Å^2)	45.7	Xtriage
Anisotropy	0.454	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 50.7	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.046 for -h,-l,-k	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	10010	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.18% 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 ⓘ

5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: PO4

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	0.44	0/1525	0.67	0/2072
1	E	0.41	0/1525	0.67	0/2072
2	B	0.42	0/1597	0.66	0/2168
2	F	0.39	0/1597	0.65	0/2168
3	C	0.44	0/107	0.80	0/141
3	G	0.46	0/103	0.81	0/137
4	D	0.40	0/1822	0.59	0/2439
4	H	0.41	0/1822	0.61	0/2439
All	All	0.42	0/10098	0.64	0/13636

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 ⓘ

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	1483	0	1420	74	0
1	E	1483	0	1420	59	1
2	B	1557	0	1488	87	0
2	F	1557	0	1488	84	0
3	C	106	0	119	13	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	G	102	0	108	9	0
4	D	1792	0	1836	80	0
4	H	1792	0	1836	74	0
5	C	5	0	0	0	0
5	D	5	0	0	0	0
5	E	5	0	0	0	0
5	H	5	0	0	0	0
6	A	25	0	0	0	0
6	B	23	0	0	0	0
6	C	1	0	0	0	0
6	D	19	0	0	1	0
6	E	34	0	0	0	0
6	F	4	0	0	0	0
6	H	12	0	0	0	0
All	All	10010	0	9715	439	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 22.

The worst 5 of 439 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:B:166:ARG:HB2	2:B:166:ARG:HH21	1.24	1.02
2:F:107:GLN:HE21	2:F:109:LEU:HD21	1.22	1.01
4:H:1:MET:CE	4:H:132:TRP:HB3	1.96	0.95
4:H:1:MET:HE1	4:H:132:TRP:HB3	1.48	0.93
4:H:160:ILE:H	4:H:160:ILE:HD12	1.33	0.91

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)
1:E:2:LYS:NZ	1:E:2:LYS:NZ[3_655]	1.93	0.27

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	178/180 (99%)	163 (92%)	13 (7%)	2 (1%)	14	50
1	E	178/180 (99%)	163 (92%)	13 (7%)	2 (1%)	14	50
2	B	188/190 (99%)	171 (91%)	17 (9%)	0	100	100
2	F	188/190 (99%)	169 (90%)	19 (10%)	0	100	100
3	C	11/13 (85%)	9 (82%)	2 (18%)	0	100	100
3	G	11/13 (85%)	9 (82%)	2 (18%)	0	100	100
4	D	212/214 (99%)	201 (95%)	8 (4%)	3 (1%)	11	43
4	H	212/214 (99%)	201 (95%)	8 (4%)	3 (1%)	11	43
All	All	1178/1194 (99%)	1086 (92%)	82 (7%)	10 (1%)	19	57

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	180	PHE
1	E	180	PHE
4	H	168	GLY
4	H	200	LEU
4	D	94	ALA

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	165/162 (102%)	152 (92%)	13 (8%)	12	41

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	165/162 (102%)	156 (94%)	9 (6%)	21	57
2	B	171/171 (100%)	156 (91%)	15 (9%)	10	36
2	F	171/171 (100%)	161 (94%)	10 (6%)	20	55
3	C	12/12 (100%)	11 (92%)	1 (8%)	11	39
3	G	11/12 (92%)	11 (100%)	0	100	100
4	D	197/197 (100%)	184 (93%)	13 (7%)	16	49
4	H	197/197 (100%)	185 (94%)	12 (6%)	18	53
All	All	1089/1084 (100%)	1016 (93%)	73 (7%)	16	49

5 of 73 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	114	LEU
4	H	187	VAL
2	F	118	SER
4	H	70	LEU
2	B	113	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
2	F	34	GLN
4	H	95	ASN
2	F	107	GLN
2	F	150	ASN
4	H	172	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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

4 ligands 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	#$|Z| > 2$
5	PO4	E	204	-	4,4,4	1.67	0	6,6,6	0.44	0
5	PO4	D	214	-	4,4,4	1.73	1 (25%)	6,6,6	0.46	0
5	PO4	C	202	-	4,4,4	1.67	0	6,6,6	0.42	0
5	PO4	H	214	-	4,4,4	1.73	1 (25%)	6,6,6	0.41	0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	H	214	PO4	P-O4	-2.10	1.48	1.54
5	D	214	PO4	P-O4	-2.02	1.48	1.54

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 ⓘ

There are no such residues in this entry.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

5.8 Polymer linkage issues (i)

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(Å²)	Q<0.9
1	A	177/180 (98%)	-0.54	0	100	100	17, 31, 53, 78	0
1	E	177/180 (98%)	-0.59	0	100	100	19, 32, 52, 81	0
2	B	190/190 (100%)	-0.35	2 (1%)	80	56	20, 34, 71, 95	0
2	F	190/190 (100%)	-0.23	6 (3%)	47	20	22, 36, 74, 94	0
3	C	13/13 (100%)	-0.57	0	100	100	23, 29, 37, 42	0
3	G	13/13 (100%)	-0.38	0	100	100	28, 31, 44, 49	0
4	D	214/214 (100%)	-0.43	0	100	100	20, 36, 57, 67	0
4	H	214/214 (100%)	-0.19	0	100	100	20, 38, 58, 73	0
All	All	1188/1194 (99%)	-0.39	8 (0%)	87	69	17, 34, 62, 95	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	1	GLY	4.5
2	B	1	GLY	2.7
2	B	165	PRO	2.5
2	F	134	ASN	2.3
2	F	111	HIS	2.3

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 monosaccharides in this entry.

6.4 Ligands ⓘ

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	PO4	H	214	5/5	0.97	0.17	51,51,52,53	0
5	PO4	E	204	5/5	0.98	0.14	67,68,69,69	0
5	PO4	C	202	5/5	0.98	0.15	59,59,60,61	0
5	PO4	D	214	5/5	0.99	0.16	41,42,42,43	0

6.5 Other polymers ⓘ

There are no such residues in this entry.