wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 16, 2021 – 11:05 PM EDT

PDB ID	:	1QXP
Title	:	Crystal Structure of a mu-like calpain
Authors	:	Pal, G.P.; Veyra, T.D.; Elce, J.S.; Jia, Z.
Deposited on	:	2003-09-08
Resolution	:	2.80 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

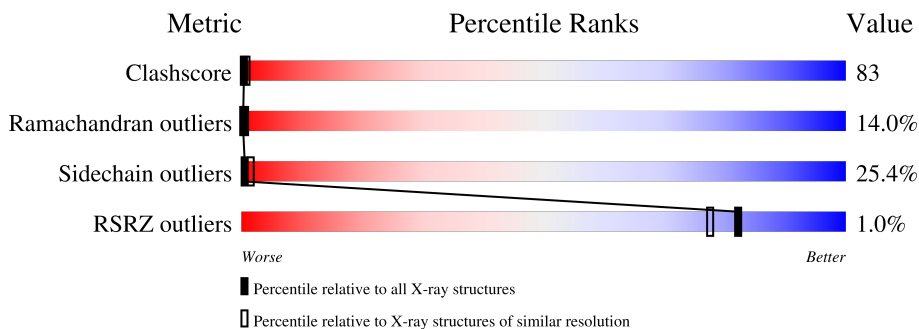|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	Percentile Ranks	Value
Clashscore		83
Ramachandran outliers		14.0%
Sidechain outliers		25.4%
RSRZ outliers		1.0%

Worse — *Better*

█ Percentile relative to all X-ray structures
❑ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	900	5% · 26% · 38% · 19% · 13%
1	B	900	% · 6% · 27% · 36% · 19% · 12%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 12368 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 mu-like calpain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	783	Total 6053	C 3846	N 1037	O 1143	S 27	0	0	0
1	B	788	Total 6003	C 3830	N 1015	O 1129	S 29	0	0	0

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	105	SER	CYS	engineered mutation	UNP P97571
B	105	SER	CYS	engineered mutation	UNP P97571
A	702A	GLY	-	cloning artifact	UNP Q07009
A	702B	LYS	-	cloning artifact	UNP Q07009
A	702C	LEU	-	cloning artifact	UNP Q07009
A	702D	ALA	-	cloning artifact	UNP Q07009
A	702E	ALA	-	cloning artifact	UNP Q07009
A	702F	ALA	-	cloning artifact	UNP Q07009
A	702G	ILE	-	cloning artifact	UNP Q07009
A	702H	GLU	-	cloning artifact	UNP Q07009
A	702I	HIS	-	expression tag	UNP Q07009
A	702J	HIS	-	expression tag	UNP Q07009
A	702K	HIS	-	expression tag	UNP Q07009
A	702L	HIS	-	expression tag	UNP Q07009
A	702M	HIS	-	expression tag	UNP Q07009
A	702N	HIS	-	expression tag	UNP Q07009
B	702A	GLY	-	cloning artifact	UNP Q07009
B	702B	LYS	-	cloning artifact	UNP Q07009
B	702C	LEU	-	cloning artifact	UNP Q07009
B	702D	ALA	-	cloning artifact	UNP Q07009
B	702E	ALA	-	cloning artifact	UNP Q07009
B	702F	ALA	-	cloning artifact	UNP Q07009
B	702G	ILE	-	cloning artifact	UNP Q07009
B	702H	GLU	-	cloning artifact	UNP Q07009
B	702I	HIS	-	expression tag	UNP Q07009

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	702J	HIS	-	expression tag	UNP Q07009
B	702K	HIS	-	expression tag	UNP Q07009
B	702L	HIS	-	expression tag	UNP Q07009
B	702M	HIS	-	expression tag	UNP Q07009
B	702N	HIS	-	expression tag	UNP Q07009

• Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	172	Total 172	O 172	0	0
2	B	140	Total 140	O 140	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: mu-like calpain

Chain A:

5%　　　26%　　　38%　　　19%　　　13%

[Color-coded residue grid: GLU ALA ASN SER GLU GLU GLU ARG GLN PHE ARG LYS LEU F730 V731 Q732 L733 A734 G735 D736 D737 M738 E739 V740 S741 A742 T743 E744 L745 M746 N747 I748 L749 K750 K751 VAL VAL THR ARG H756 P757 D758 L759 K760 THR ASP G763 F764 G765 I766 D767 T768 C769 R770 S771 M772 V773 A774 V775]

[M776 D777 S778 D779 T780 T781 G782 K783 L784 G785 F786 E787 E788 F789 K790 Y791 L792 W793 N794 N795 I796 K797 K798 W799 Q800 G801 I802 Y803 LYS ARG PHE GLU T808 D809 R810 S811 G812 THR ILE GLY SER ASN GLU P820 G821 A822 F823 E824 A825 A826 G827 F828 H829 L830 N831 N832 H833 I834 Y835]

[S836 MET ILE ILE ARG ARG TYR SER ASP E845 T846 G847 N848 D850 F851 D852 M853 F854 I855 T856 C857 L858 V859 R860 L861 D862 A863 M864 F865 R866 A867 F868 R869 S870 L871 D872 K873 T874 G875 T876 G877 Q878 I879 Q880 V881 N882 I883 Q884 E885 W886 L887 Q888 L889 T890 M891 Y892 S893]

- Molecule 1: mu-like calpain

%
Chain B: 6%   27%   36%   19%   12%

[Color-coded residue grids for Chain B continuing through multiple rows, beginning MET A2 G3 I4 A5 M6 K7 L8 A9 K10 D11 R12 E13 A14 A15 E16 G17 L18 G19 S20 H21 E22 R23 A24 I25 K26 Y27 L28 N29 Q30 D31 Y32 E33 T34 L35 R36 N37 E38 C39 L40 E41 A42 G43 A44 L45 F46 Q47 D48 P49 A50 F51 P52 P53 V54 S55 H56 H57 L58 G59 F60]

[K61 E62 L63 G64 P65 N66 S67 S68 K69 T70 Y71 G72 I73 K74 W75 K76 R77 P78 T79 E80 L81 L82 S83 N84 P85 Q86 F87 I88 V89 D90 G91 A92 T93 R94 T95 D96 I97 C98 Q99 G100 A101 L102 G103 D104 S105 W106 L107 L108 A109 A110 I111 A112 S113 L114 T115 L116 N117 E118 T119 I120]

[L121 H122 R123 V124 V125 P126 Y127 G128 Q129 S130 F131 Q132 E133 G134 Y35 A136 G137 I138 F139 H140 F141 Q142 L143 W144 Q145 F146 G147 E148 W149 V150 D151 V152 V153 V154 D155 D156 L157 L158 P159 T160 K161 D162 G163 K164 L165 V166 F167 V168 H169 S170 A171 Q72 G173 N74 E175 F176 W177 S178 A179 L180]

[L181 E182 K183 A184 Y185 A186 K187 V188 N189 G190 S191 Y192 E193 A194 I195 S196 G197 G198 C199 T200 S201 E202 A203 D206 E205 D206 P207 T208 G209 G210 V211 T212 E213 W214 Y215 D216 L217 Q218 K219 A220 P221 S222 D223 L224 Y225 Q226 I227 I228 L229 K230 A231 L232 E233 R234 G235 S236 L237 L238 G239 C240]

[S241 I242 N243 ILE SER ASP ILE ARG ASP LEU GLU ALA ILE THR PHE K256 N257 L258 V259 ARG G261 H262 A263 Y264 S265 V266 T267 D268 A269 K270 Q271 V272 T273 Y274 Q275 Q276 Q277 R278 V279 N280 L281 I282 R283 M284 R285 N286 P287 W288 G289 E290 V291 E292 W293 K294 G295 P296 W297 S298 ASP ASN]

[SER TYR E303 W304 N305 LYS VAL D308 P309 Y310 E311 R312 E313 Q314 I315 R316 V317 K318 M319 E320 D321 G322 E323 F324 W325 M326 S327 F328 R329 D330 F331 I332 T333 E334 F335 T336 K337 L338 E339 I340 C341 N342 L343 T344 P345 D346 A347 L348 K349 S350 R351 THR LEU N355 W356 N357 T358 T359 F360]

[Y361 E362 G363 T364 W365 R366 R367 G368 S369 T370 A371 G372 G373 C374 H375 R376 Y377 P378 A379 T380 F381 W382 V383 N384 P385 Q386 F387 K388 I389 R390 L391 E392 E393 V394 D395 D396 A397 D398 D399 Y400 D401 S402 R403 E404 S405 G406 C407 S408 F409 L410 L411 A412 L413 M414 Q415 K416 H417 R418 ARG ARG]

[GLU ARG ARG F424 G425 R426 D427 M428 E429 T430 I431 G432 F433 A434 V435 Y436 Q437 V438 P439 R440 E441 L442 A443 G444 GLN P446 V447 H448 L449 K450 R451 D452 F453 F454 L455 A456 N457 A458 S459 R460 A461 Q462 S463 E464 H465 F466 I467 N468 L469 R470 E471 V472 S473 N474 R475 I476 R477 L478 P479 P480]

[G481 E482 Y483 I484 V485 V486 P487 S488 T489 F490 E491 P492 N493 K494 E495 G496 D497 F498 L499 L500 R501 F502 F503 S504 E505 K506 K507 A508 G509 T510 Q511 E512 L513 D514 D515 Q516 I517 Q518 A519 N520 L521 P522 D523 E524 K525 V526 LEU SER GLU GLU GLU ILE ASP ASP ASN PHE K537 T538 L539 F540]

[SER LYS LEU ALA GLY ASP ASP MET GLU I550 S551 V552 K553 E554 L555 Q556 T557 I558 L559 N560 ARG ILE ILE SER LYS LEU R570 T571 N572 G573 F574 S575 L576 E577 S578 C579 R580 S581 M582 V583 N584 L585 M586 V587 V588 G589 N600]

[I601 L602 W603 N604 R605 I606 R607 N608 Y609 L610 T611 I612 F613 R614 K615 F616 D617 L618 D619 K620 S621 G622 S623 M624 S625 A626 Y627 E628 M629 R630 M631 A632 I633 E634 A635 A636 G637 F638 K639 L640 P641 C642 Q643 L644 H645 Q646 V647 I648 V649 A650 R651 F652 A653 D654 D655 E656 L657 I658 L659 D660]

[F661 D662 N663 F664 V665 R666 C667 L668 V669 R670 L671 E672 I673 L674 F675 K676 I677 F678 K679 Q680 L681 D682 P683 E684 N685 T686 G687 T688 I689 Q690 L691 D692 L693 I694 S695 W696 L697 S698 F699 S700 V701 L702 GLY LYS LEU ALA ALA ILE GLU HIS HIS HIS HIS HIS M710 H711 Y712 S713]

[N714 ILE GLU ALA N718 E719 S720 E0 GLU E723 R724 Q725 F726 R727 K728 L729 F730 V731 Q732 L733 ALA GLY ASP ASP M738 E739 V740 S741 A742 T743 E744 L745 M746 ASN ILE LEU LYS LYS VAL VAL THR ARG HIS PRO ASP LEU LYS THR ASP GLY G765 I766 D767 T768 C769 R770 S771 M772 V773]

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

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	72.74Å 184.60Å 86.37Å 90.00° 100.74° 90.00°	Depositor
Resolution (Å)	91.29 – 2.80 49.81 – 2.69	Depositor EDS
% Data completeness (in resolution range)	91.6 (91.29-2.80) 87.3 (49.81-2.69)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.50 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.229 , 0.311 0.232 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	54.9	Xtriage
Anisotropy	0.068	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 98.8	EDS
L-test for twinning [2]	$<|L|> = 0.46, <L^2> = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	12368	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.84% 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	3.70	964/6177 (15.6%)	2.67	470/8354 (5.6%)
1	B	3.71	932/6128 (15.2%)	2.67	489/8288 (5.9%)
All	All	3.70	1896/12305 (15.4%)	2.67	959/16642 (5.8%)

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	45
1	B	0	39
All	All	0	84

The worst 5 of 1896 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	67	SER	CA-CB	28.73	1.96	1.52
1	A	339	GLU	CD-OE1	23.66	1.51	1.25
1	A	320	GLU	CD-OE2	20.06	1.47	1.25
1	B	429	GLU	CD-OE1	19.54	1.47	1.25
1	A	811	SER	CA-CB	-18.37	1.25	1.52

The worst 5 of 959 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	460	ARG	NE-CZ-NH1	-29.77	105.41	120.30
1	A	329	ARG	NE-CZ-NH1	26.14	133.37	120.30
1	A	329	ARG	NE-CZ-NH2	-24.17	108.22	120.30
1	B	514	ASP	CB-CG-OD2	-18.56	101.59	118.30
1	A	285	ARG	NE-CZ-NH2	-15.96	112.32	120.30

There are no chirality outliers.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

5 of 84 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	120	ILE	Mainchain
1	A	152	VAL	Mainchain
1	A	24	ALA	Mainchain
1	A	29	ASN	Mainchain
1	A	99	GLN	Mainchain,Peptide

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	6053	0	5615	959	3
1	B	6003	0	5477	969	1
2	A	172	0	0	53	1
2	B	140	0	0	45	1
All	All	12368	0	11092	1928	3

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 83.

The worst 5 of 1928 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:583:VAL:CB	1:A:583:VAL:CA	1.74	1.65
1:B:0:GLU:CA	1:B:0:GLU:CB	1.75	1.64
1:A:550:ILE:CB	1:A:550:ILE:CA	1.75	1.64
1:B:786:PHE:CA	1:B:786:PHE:CB	1.74	1.64
1:B:8:LEU:CD2	1:B:8:LEU:CG	1.76	1.63

All (3) 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:A:457:ASN:OD1	2:B:1033:HOH:O[1_454]	1.56	0.64
1:A:442:LEU:CD2	2:A:922:HOH:O[1_455]	1.95	0.25

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:300:ASN:OD1	1:B:303:GLU:N[2_656]	2.08	0.12

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	747/900 (83%)	500 (67%)	143 (19%)	104 (14%)	0	0
1	B	748/900 (83%)	497 (66%)	146 (20%)	105 (14%)	0	0
All	All	1495/1800 (83%)	997 (67%)	289 (19%)	209 (14%)	0	0

5 of 209 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	54	VAL
1	A	60	PHE
1	A	63	LEU
1	A	67	SER
1	A	83	SER

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	606/782 (78%)	459 (76%)	147 (24%)	0	2
1	B	584/782 (75%)	429 (74%)	155 (26%)	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1190/1564 (76%)	888 (75%)	302 (25%)	0 1

5 of 302 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	427	ASP
1	B	779	ASP
1	B	457	ASN
1	B	616	PHE
1	B	866	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 53 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	37	ASN
1	B	243	ASN
1	B	663	ASN
1	B	47	GLN
1	B	140	HIS

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 ⓘ

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(Å²)	Q<0.9
1	A	783/900 (87%)	-0.36	4 (0%)	91	88	16, 49, 81, 98	0
1	B	788/900 (87%)	-0.32	11 (1%)	75	70	17, 51, 84, 102	0
All	All	1571/1800 (87%)	-0.34	15 (0%)	82	77	16, 50, 83, 102	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	596	LEU	3.7
1	B	587	ASP	3.7
1	A	596	LEU	3.3
1	B	712	TYR	3.2
1	B	402	SER	3.2

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 ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.