wwPDB X-ray Structure Validation Summary Report (i)

Aug 23, 2023 – 03:29 PM EDT

PDB ID	:	3C0Y
Title	:	Crystal structure of catalytic domain of human histone deacetylase HDAC7
Authors	:	Min, J.R.; Schuetz, A.; Allali-Hassani, A.; Loppnau, P.; Kwiatkowski, N.P.; Mazitschek, R.; Weigelt, J.; Sundstrom, M.; Arrowsmith, C.H.; Edwards, A.M.; Vedadi, M.; Bochkarev, A.; Plotnikov, A.N.; Structural Genomics Consortium (SGC)
Deposited on	:	2008-01-21
Resolution	:	2.10 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:

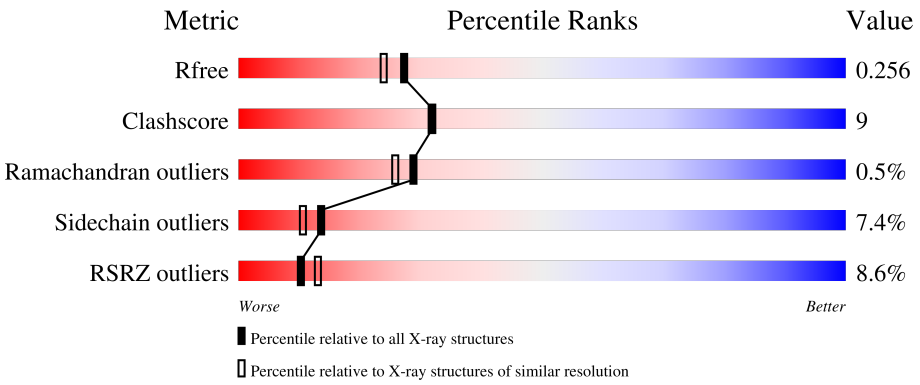
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	423	8% 73% 15% • 9%
1	B	423	6% 66% 18% • • 13%
1	C	423	9% 67% 17% • 15%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 8791 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 Histone deacetylase 7a.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	386	2925	1835	527	543	20	0	0	0
1	B	367	2786	1745	505	517	19	0	0	0
1	C	358	2718	1698	493	509	18	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	481	GLY	-	expression tag	UNP Q8WUI4
B	481	GLY	-	expression tag	UNP Q8WUI4
C	481	GLY	-	expression tag	UNP Q8WUI4

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
2	A	2	2	2	0	0
2	B	2	2	2	0	0
2	C	2	2	2	0	0

- Molecule 3 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	K		
3	A	2	2	2	0	0
3	B	2	2	2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	K		
3	C	2	2	2	0	0

- Molecule 4 is water.

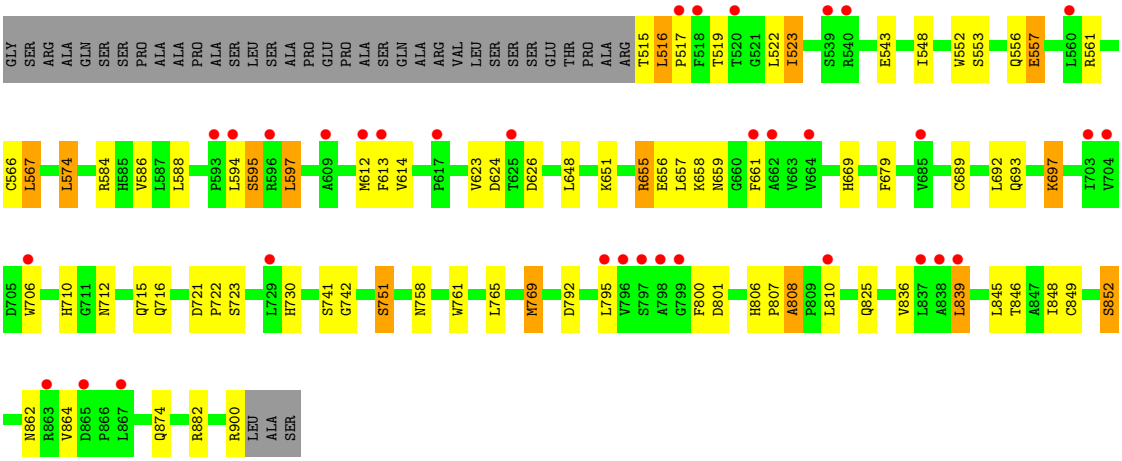
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	122	122	122	0	0
			Total	O		
4	B	129	129	129	0	0
			Total	O		
4	C	99	99	99	0	0

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.

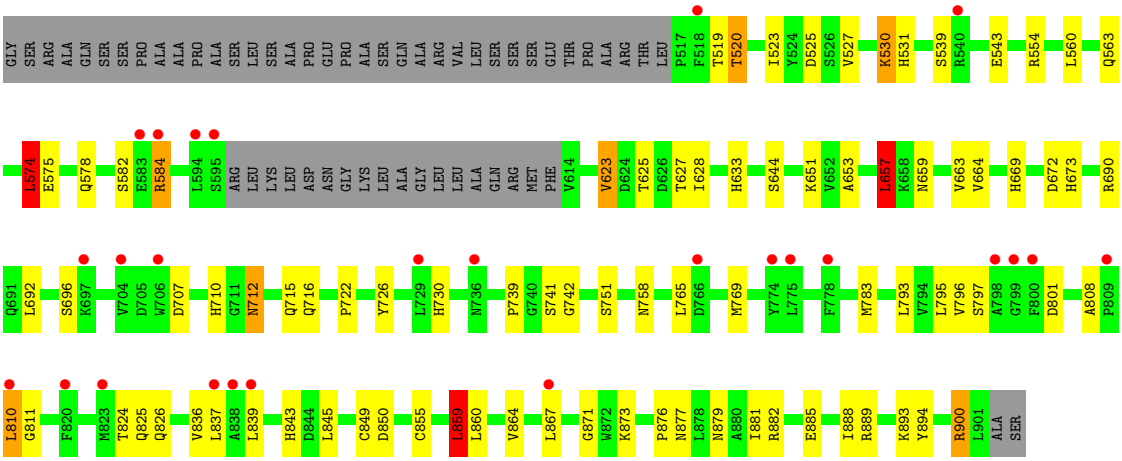
• Molecule 1: Histone deacetylase 7a

Chain A:



• Molecule 1: Histone deacetylase 7a

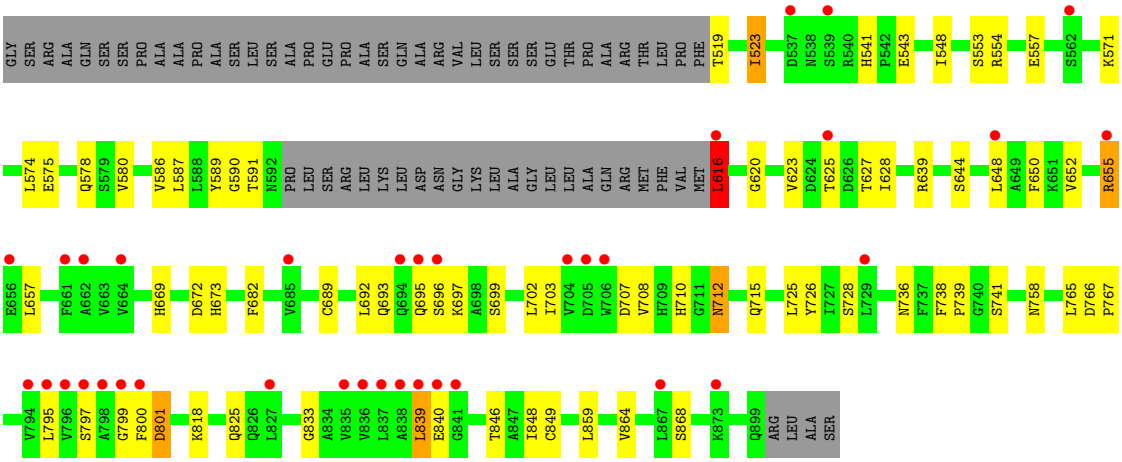
Chain B:



• Molecule 1: Histone deacetylase 7a

Chain C:

9%
67%
17%
15%


GLY SER ARG ALA GLN SER SER PRO ALA ALA PRO ALA SER LEU SER ALA PRO GLU PRO ALA SER GLN ALA ARG VAL LEU SER SER SER GLU THR PRO ALA ARG THR LEU PRO PHE T519 I523 D537 N538 S539 R540 H541 P542 E543 I548 S553 R554 E557 S562 K571
L574 E575 Q578 S579 V580 V586 L587 L588 Y589 G590 T591 N592 PRO LEU SER ARG LEU LYS LEU ASP ASN GLY LYS LEU ALA GLY LEU LEU ALA GLN ARG MET PHE VAL MET L616 G620 V623 D624 T625 D626 T627 I628 R639 S644 L648 A649 F650 K651 V652 R655
E656 L657 F661 A662 V663 V664 H669 D672 H673 F682 V685 C689 L692 Q693 Q694 Q695 S696 K697 A698 S699 L702 I703 V704 D705 W706 D707 V708 H709 H710 G711 N712 Q715 L725 Y726 I727 S728 L729 N736 F737 F738 P739 G740 S741 N758 L765 D766 P767
V794 L795 V796 S797 A798 G799 F800 D801 K818 Q825 Q826 L827 G833 A834 V835 V836 L837 A838 L839 E840 G841 T846 A847 I848 C849 L859 V864 L867 S868 K873 Q899 ARG LEU ALA SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α, β, γ	81.75Å 81.75Å 148.90Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	34.44 – 2.10 34.44 – 2.10	Depositor EDS
% Data completeness (in resolution range)	98.1 (34.44-2.10) 98.1 (34.44-2.10)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.69 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.202 , 0.264 0.198 , 0.256	Depositor DCC
R_{free} test set	3222 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	41.7	Xtriage
Anisotropy	0.026	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 45.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.008 for -h,-k,l 0.035 for h,-h-k,-l 0.019 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	8791	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.24% 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: K, ZN

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	0.90	1/2995 (0.0%)	0.91	6/4063 (0.1%)
1	B	0.90	3/2854 (0.1%)	0.92	6/3871 (0.2%)
1	C	0.76	0/2783	0.78	1/3775 (0.0%)
All	All	0.86	4/8632 (0.0%)	0.87	13/11709 (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

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	855	CYS	CB-SG	-6.00	1.72	1.82
1	A	716	GLN	CB-CG	5.53	1.67	1.52
1	B	726	TYR	CD2-CE2	5.18	1.47	1.39
1	B	663	VAL	CA-CB	5.14	1.65	1.54

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	690	ARG	NE-CZ-NH2	-7.32	116.64	120.30
1	A	769	MET	CG-SD-CE	-6.60	89.64	100.20
1	A	516	LEU	CA-CB-CG	6.36	129.93	115.30
1	B	657	LEU	CA-CB-CG	6.29	129.77	115.30
1	B	707	ASP	CB-CG-OD2	6.18	123.86	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	612	MET	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	2925	0	2823	55	0
1	B	2786	0	2678	51	0
1	C	2718	0	2601	43	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
3	C	2	0	0	0	0
4	A	122	0	0	7	0
4	B	129	0	0	9	0
4	C	99	0	0	6	0
All	All	8791	0	8102	147	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 9.

The worst 5 of 147 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:B:769:MET:HE3	4:B:329:HOH:O	1.63	0.98
1:B:575:GLU:HG3	4:B:65:HOH:O	1.70	0.91
1:B:769:MET:CE	4:B:329:HOH:O	2.17	0.91
1:A:595:SER:HB3	1:A:597:LEU:HB2	1.55	0.89
1:A:825:GLN:HE22	1:A:864:VAL:H	1.22	0.86

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	384/423 (91%)	363 (94%)	18 (5%)	3 (1%)	19	15
1	B	363/423 (86%)	344 (95%)	17 (5%)	2 (1%)	25	21
1	C	354/423 (84%)	337 (95%)	17 (5%)	0	100	100
All	All	1101/1269 (87%)	1044 (95%)	52 (5%)	5 (0%)	29	26

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	613	PHE
1	A	808	ALA
1	B	531	HIS
1	A	557	GLU
1	B	811	GLY

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	306/336 (91%)	280 (92%)	26 (8%)	10	7
1	B	292/336 (87%)	273 (94%)	19 (6%)	17	14
1	C	284/336 (84%)	264 (93%)	20 (7%)	15	12
All	All	882/1008 (88%)	817 (93%)	65 (7%)	13	10

5 of 65 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	699	SER
1	C	712	ASN
1	A	882	ARG
1	A	874	GLN
1	C	765	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	673	HIS
1	C	825	GLN
1	C	693	GLN
1	C	715	GLN
1	B	592	ASN

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 12 ligands modelled in this entry, 12 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	386/423 (91%)	0.38	34 (8%)	10	12	24, 40, 62, 71	0
1	B	367/423 (86%)	0.22	26 (7%)	16	20	25, 38, 59, 71	0
1	C	358/423 (84%)	0.58	36 (10%)	7	9	31, 48, 67, 74	0
All	All	1111/1269 (87%)	0.39	96 (8%)	10	13	24, 42, 63, 74	0

The worst 5 of 96 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	837	LEU	8.6
1	C	796	VAL	6.9
1	C	798	ALA	6.0
1	C	797	SER	5.3
1	C	706	TRP	5.3

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
3	K	C	905	1/1	0.95	0.12	61,61,61,61	0
3	K	C	904	1/1	0.97	0.09	46,46,46,46	0
3	K	B	905	1/1	0.97	0.12	46,46,46,46	0
2	ZN	C	405	1/1	0.99	0.09	43,43,43,43	0
2	ZN	C	406	1/1	0.99	0.04	54,54,54,54	0
3	K	A	904	1/1	0.99	0.08	33,33,33,33	0
3	K	A	905	1/1	0.99	0.06	43,43,43,43	0
3	K	B	904	1/1	0.99	0.09	38,38,38,38	0
2	ZN	A	401	1/1	0.99	0.12	34,34,34,34	0
2	ZN	B	403	1/1	0.99	0.11	38,38,38,38	0
2	ZN	B	404	1/1	0.99	0.05	49,49,49,49	0
2	ZN	A	402	1/1	1.00	0.04	46,46,46,46	0

6.5 Other polymers (i)

There are no such residues in this entry.