



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 24, 2024 – 12:59 AM EDT

PDB ID : 1S4I
Title : Crystal structure of a SOD-like protein from *Bacillus subtilis*
Authors : Banci, L.; Bertini, I.; Calderone, V.; Cramaro, F.; Del Conte, R.; Fantoni, A.; Mangani, S.; Quattrone, A.; Viezzoli, M.S.
Deposited on : 2004-01-16
Resolution : 1.80 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

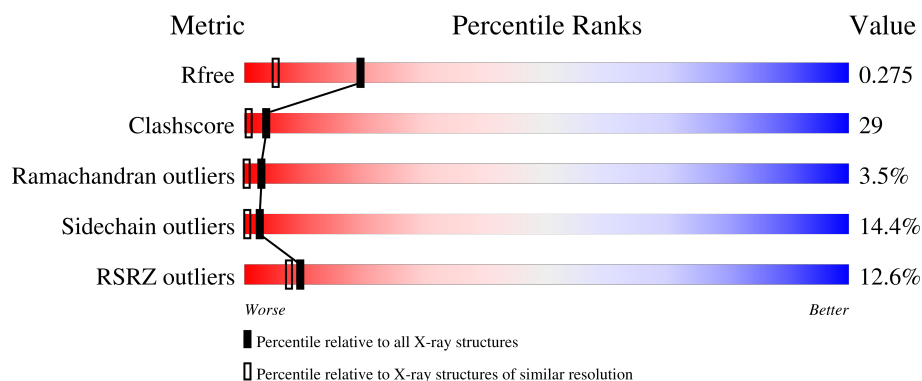
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	7108 (1.80-1.80)
Clashscore	180529	8162 (1.80-1.80)
Ramachandran outliers	177936	8077 (1.80-1.80)
Sidechain outliers	177891	8076 (1.80-1.80)
RSRZ outliers	164620	7108 (1.80-1.80)

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 $\leq 5\%$. 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	175	
1	B	175	
1	C	175	
1	D	175	

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 4920 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 superoxide dismutase-like protein yojM.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	152	Total	C	N	O	S	0	0	0
			1133	704	201	224	4			
1	A	151	Total	C	N	O	S	0	0	0
			1129	702	200	223	4			
1	C	151	Total	C	N	O	S	0	0	0
			1129	702	200	223	4			
1	D	151	Total	C	N	O	S	0	0	0
			1129	702	200	223	4			

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	2	Total	Zn	0	0
			2	2		
2	A	2	Total	Zn	0	0
			2	2		
2	C	1	Total	Zn	0	0
			1	1		
2	D	1	Total	Zn	0	0
			1	1		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Cl	0	0
			1	1		
3	A	1	Total	Cl	0	0
			1	1		
3	C	1	Total	Cl	0	0
			1	1		
3	D	1	Total	Cl	0	0
			1	1		

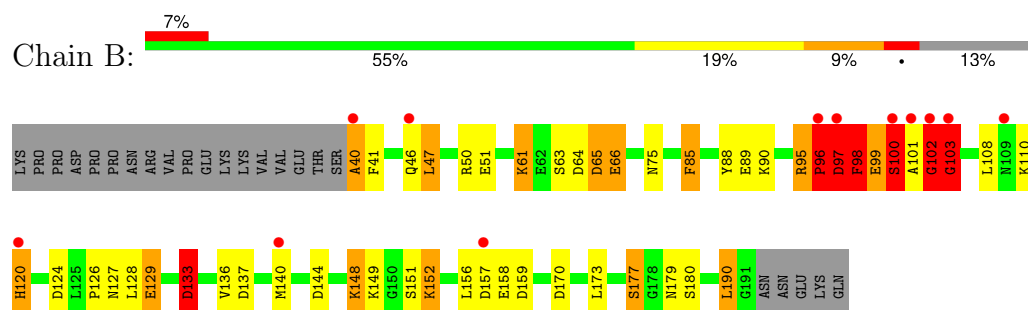
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	117	Total 117	O 117	0	0
4	A	93	Total 93	O 93	0	0
4	C	83	Total 83	O 83	0	0
4	D	97	Total 97	O 97	0	0

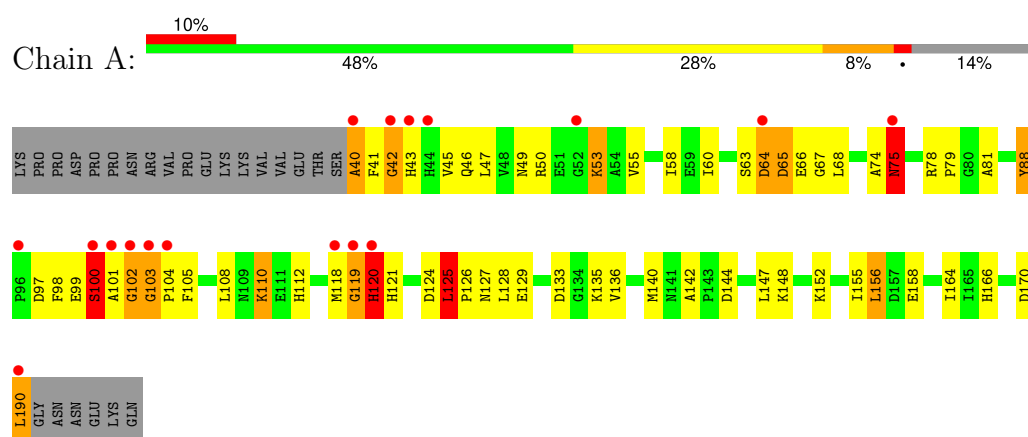
3 Residue-property plots [i](#)

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.

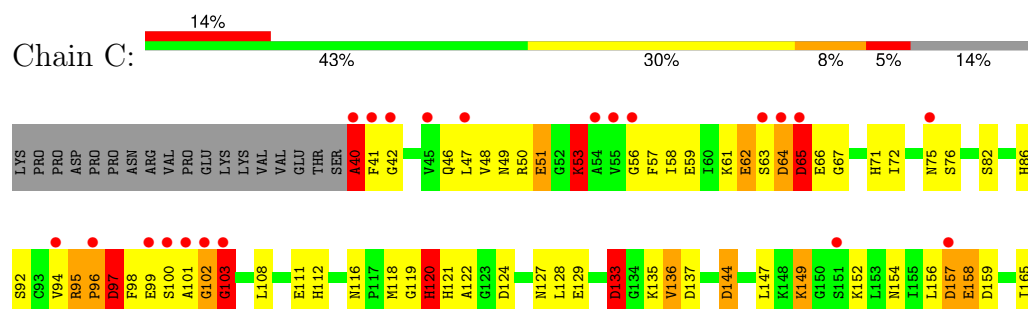
- Molecule 1: superoxide dismutase-like protein yojM

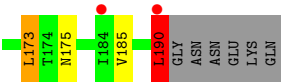


- Molecule 1: superoxide dismutase-like protein yojM

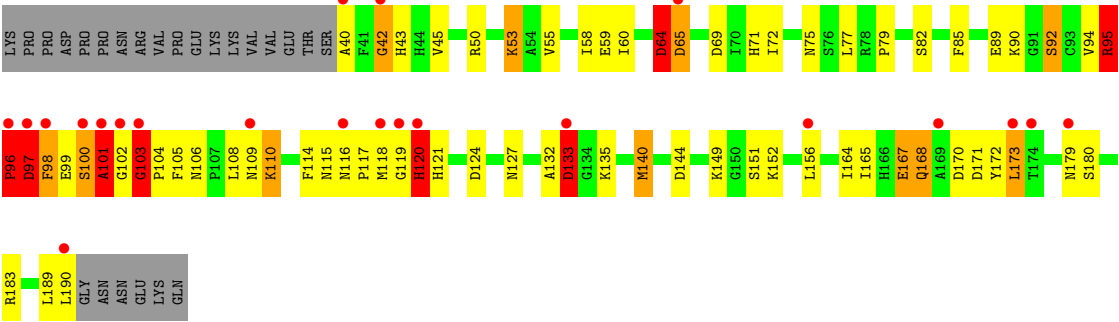
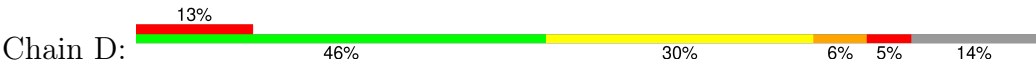


- Molecule 1: superoxide dismutase-like protein yojM





● Molecule 1: superoxide dismutase-like protein yojM



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	38.23Å 61.11Å 64.92Å 84.35° 76.02° 90.42°	Depositor
Resolution (Å)	37.01 – 1.80 37.01 – 1.80	Depositor EDS
% Data completeness (in resolution range)	100.0 (37.01-1.80) 94.6 (37.01-1.80)	Depositor EDS
R_{merge}	0.03	Depositor
R_{sym}	0.03	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.54 (at 1.81Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.221 , 0.258 0.250 , 0.275	Depositor DCC
R_{free} test set	4150 reflections (8.35%)	wwPDB-VP
Wilson B-factor (Å ²)	27.4	Xtriage
Anisotropy	0.372	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 51.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	4920	wwPDB-VP
Average B, all atoms (Å ²)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.09% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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](#)

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

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	1.19	4/1153 (0.3%)	1.35	12/1558 (0.8%)
1	B	1.43	7/1157 (0.6%)	1.66	24/1563 (1.5%)
1	C	1.17	2/1153 (0.2%)	1.49	22/1558 (1.4%)
1	D	1.39	12/1153 (1.0%)	1.62	29/1558 (1.9%)
All	All	1.30	25/4616 (0.5%)	1.54	87/6237 (1.4%)

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	1	7
1	B	0	7
1	C	0	4
1	D	1	10
All	All	2	28

The worst 5 of 25 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	120	HIS	C-N	10.92	1.59	1.34
1	D	96	PRO	N-CD	10.68	1.62	1.47
1	B	97	ASP	C-N	-9.45	1.12	1.34
1	B	96	PRO	C-N	-9.45	1.12	1.34
1	D	65	ASP	C-O	8.58	1.39	1.23

The worst 5 of 87 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	96	PRO	C-N-CA	19.17	169.61	121.70
1	B	97	ASP	O-C-N	-18.30	93.42	122.70
1	C	137	ASP	CB-CG-OD1	12.74	129.77	118.30
1	D	104	PRO	CA-N-CD	-10.21	97.21	111.50
1	D	65	ASP	CA-C-O	-10.05	98.99	120.10

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	120	HIS	CA
1	D	101	ALA	CA

5 of 28 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	102	GLY	Peptide
1	B	103	GLY	Peptide
1	B	40	ALA	Peptide
1	B	96	PRO	Peptide
1	B	97	ASP	Peptide,Mainchain

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	1129	0	1075	81	0
1	B	1133	0	1076	58	0
1	C	1129	0	1075	63	0
1	D	1129	0	1072	55	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	93	0	0	24	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	117	0	0	18	0
4	C	83	0	0	17	0
4	D	97	0	0	4	0
All	All	4920	0	4298	253	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 29.

The worst 5 of 253 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:156:LEU:HD23	1:A:190:LEU:CD2	1.64	1.28
1:D:110:LYS:HE2	1:D:120:HIS:CB	1.66	1.25
1:D:156:LEU:HD21	1:D:189:LEU:HG	1.19	1.17
1:A:156:LEU:CD2	1:A:190:LEU:CD2	2.25	1.14
1:A:140:MET:HB3	4:A:894:HOH:O	1.46	1.12

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	149/175 (85%)	132 (89%)	15 (10%)	2 (1%)	10	3
1	B	150/175 (86%)	134 (89%)	10 (7%)	6 (4%)	2	0
1	C	149/175 (85%)	125 (84%)	15 (10%)	9 (6%)	1	0
1	D	149/175 (85%)	136 (91%)	9 (6%)	4 (3%)	4	0
All	All	597/700 (85%)	527 (88%)	49 (8%)	21 (4%)	3	0

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	97	ASP
1	B	98	PHE
1	B	120	HIS
1	B	133	ASP
1	A	120	HIS

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	120/143 (84%)	106 (88%)	14 (12%)	4	1
1	B	120/143 (84%)	100 (83%)	20 (17%)	2	0
1	C	120/143 (84%)	105 (88%)	15 (12%)	3	1
1	D	120/143 (84%)	100 (83%)	20 (17%)	2	0
All	All	480/572 (84%)	411 (86%)	69 (14%)	2	0

5 of 69 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	108	LEU
1	D	120	HIS
1	D	167	GLU
1	A	92	SER
1	A	75	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	175	ASN
1	D	75	ASN
1	D	127	ASN
1	D	115	ASN
1	A	127	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 10 ligands modelled in this entry, 10 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	D	2
1	B	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	102:GLY	C	103:GLY	N	1.17

Continued on next page...

Continued from previous page...

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	65:ASP	C	66:GLU	N	1.14
1	B	96:PRO	C	97:ASP	N	1.12
1	B	97:ASP	C	98:PHE	N	1.12

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	151/175 (86%)	1.05	18 (11%) 10 8	17, 39, 51, 61	0
1	B	152/175 (86%)	0.63	12 (7%) 20 18	16, 33, 50, 63	0
1	C	151/175 (86%)	1.26	24 (15%) 6 4	17, 40, 58, 70	0
1	D	151/175 (86%)	1.10	22 (14%) 7 6	17, 38, 51, 57	0
All	All	605/700 (86%)	1.01	76 (12%) 9 7	16, 38, 54, 70	0

The worst 5 of 76 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	97	ASP	5.8
1	D	120	HIS	5.7
1	A	103	GLY	5.7
1	D	101	ALA	5.6
1	A	120	HIS	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(\AA^2)	Q<0.9
2	ZN	A	802	1/1	0.93	0.10	32,32,32,32	0
2	ZN	A	501	1/1	0.95	0.08	36,36,36,36	0
3	CL	A	502	1/1	0.96	0.09	34,34,34,34	0
3	CL	B	402	1/1	0.97	0.05	28,28,28,28	0
2	ZN	B	801	1/1	0.97	0.10	27,27,27,27	0
2	ZN	B	401	1/1	0.98	0.06	33,33,33,33	0
2	ZN	C	601	1/1	0.98	0.08	33,33,33,33	0
3	CL	D	702	1/1	0.98	0.06	38,38,38,38	0
3	CL	C	602	1/1	0.99	0.05	30,30,30,30	0
2	ZN	D	701	1/1	0.99	0.05	38,38,38,38	0

6.5 Other polymers [i](#)

There are no such residues in this entry.