W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 01:57 PM GMT

|		
PDB ID	:	3VKU
Title	:	Penta mutant of Lactobacillus casei lactate dehydrogenase
Authors	:	Arai, K.; Miyanaga, A.; Uchikoba, H.; Fushinobu, S.; Taguchi, H.
Deposited on	:	2011-11-22
Resolution	:	1.96 Å(reported)

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 1.96 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.
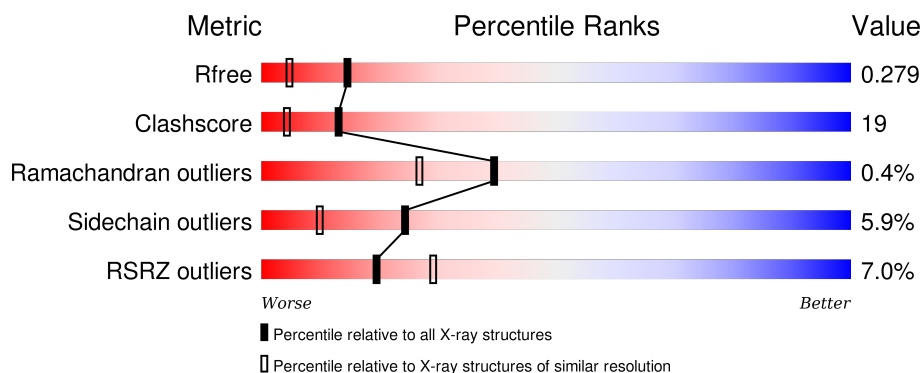
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1833 (1.96-1.96)
Clashscore	102246	1953 (1.96-1.96)
Ramachandran outliers	100387	1936 (1.96-1.96)
Sidechain outliers	100360	1936 (1.96-1.96)
RSRZ outliers	91569	1835 (1.96-1.96)

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. 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	326	9% / 55% / 30% · 11%
1	B	326	5% / 70% / 21% · · 5%
1	C	326	3% / 62% / 25% · 8%
1	D	326	10% / 67% / 20% · · 8%
1	E	326	5% / 74% / 15% 5% · ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	326	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	SO4	B	401	-	-	X	-
2	SO4	F	401	-	-	X	-

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 14731 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 L-lactate dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	290	Total	C	N	O	S	0	0	0
			2225	1424	369	426	6			
1	B	311	Total	C	N	O	S	0	0	0
			2393	1530	399	458	6			
1	C	299	Total	C	N	O	S	0	0	0
			2296	1468	381	441	6			
1	D	300	Total	C	N	O	S	0	0	0
			2304	1472	382	444	6			
1	E	312	Total	C	N	O	S	0	0	0
			2401	1534	400	461	6			
1	F	315	Total	C	N	O	S	0	0	0
			2421	1547	403	465	6			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	55	GLU	SER	ENGINEERED MUTATION	UNP P00343
A	56	ASP	ASN	ENGINEERED MUTATION	UNP P00343
A	164	LYS	GLU	ENGINEERED MUTATION	UNP P00343
A	223	ASN	ASP	ENGINEERED MUTATION	UNP P00343
A	224	LYS	ALA	ENGINEERED MUTATION	UNP P00343
B	55	GLU	SER	ENGINEERED MUTATION	UNP P00343
B	56	ASP	ASN	ENGINEERED MUTATION	UNP P00343
B	164	LYS	GLU	ENGINEERED MUTATION	UNP P00343
B	223	ASN	ASP	ENGINEERED MUTATION	UNP P00343
B	224	LYS	ALA	ENGINEERED MUTATION	UNP P00343
C	55	GLU	SER	ENGINEERED MUTATION	UNP P00343
C	56	ASP	ASN	ENGINEERED MUTATION	UNP P00343
C	164	LYS	GLU	ENGINEERED MUTATION	UNP P00343
C	223	ASN	ASP	ENGINEERED MUTATION	UNP P00343
C	224	LYS	ALA	ENGINEERED MUTATION	UNP P00343
D	55	GLU	SER	ENGINEERED MUTATION	UNP P00343
D	56	ASP	ASN	ENGINEERED MUTATION	UNP P00343

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	164	LYS	GLU	ENGINEERED MUTATION	UNP P00343
D	223	ASN	ASP	ENGINEERED MUTATION	UNP P00343
D	224	LYS	ALA	ENGINEERED MUTATION	UNP P00343
E	55	GLU	SER	ENGINEERED MUTATION	UNP P00343
E	56	ASP	ASN	ENGINEERED MUTATION	UNP P00343
E	164	LYS	GLU	ENGINEERED MUTATION	UNP P00343
E	223	ASN	ASP	ENGINEERED MUTATION	UNP P00343
E	224	LYS	ALA	ENGINEERED MUTATION	UNP P00343
F	55	GLU	SER	ENGINEERED MUTATION	UNP P00343
F	56	ASP	ASN	ENGINEERED MUTATION	UNP P00343
F	164	LYS	GLU	ENGINEERED MUTATION	UNP P00343
F	223	ASN	ASP	ENGINEERED MUTATION	UNP P00343
F	224	LYS	ALA	ENGINEERED MUTATION	UNP P00343

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).
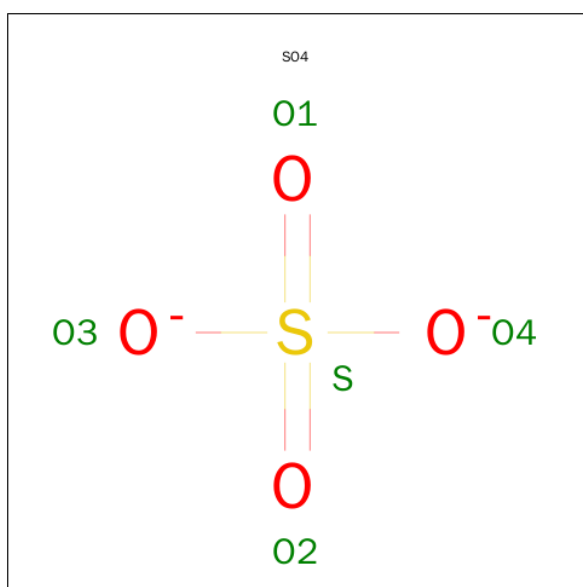
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	C	1	Total 5	O 4	S 1	0	0
2	D	1	Total 5	O 4	S 1	0	0
2	E	1	Total 5	O 4	S 1	0	0
2	E	1	Total 5	O 4	S 1	0	0
2	E	1	Total 5	O 4	S 1	0	0
2	F	1	Total 5	O 4	S 1	0	0
2	F	1	Total 5	O 4	S 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	77	Total 77	O 77	0	0
3	B	133	Total 133	O 133	0	0
3	C	93	Total 93	O 93	0	0
3	D	101	Total 101	O 101	0	0
3	E	142	Total 142	O 142	0	0
3	F	85	Total 85	O 85	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.
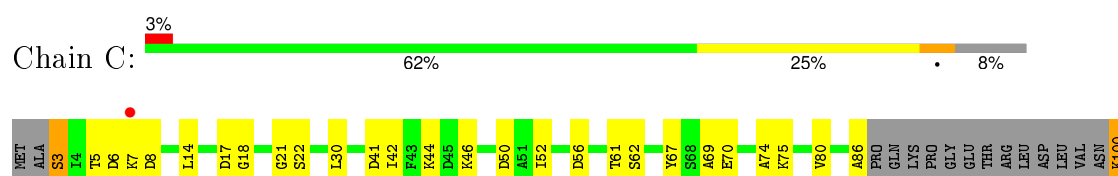
- Molecule 1: L-lactate dehydrogenase
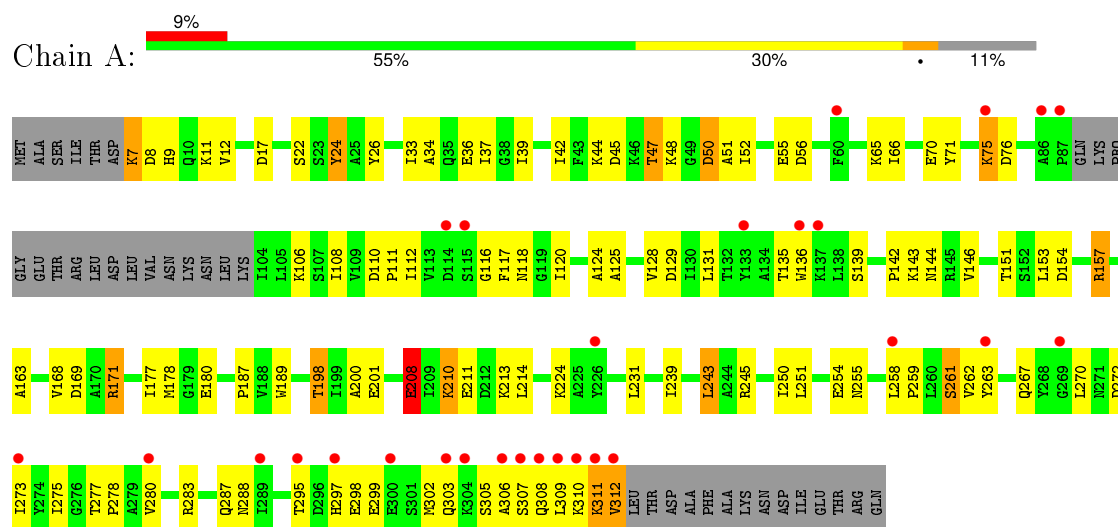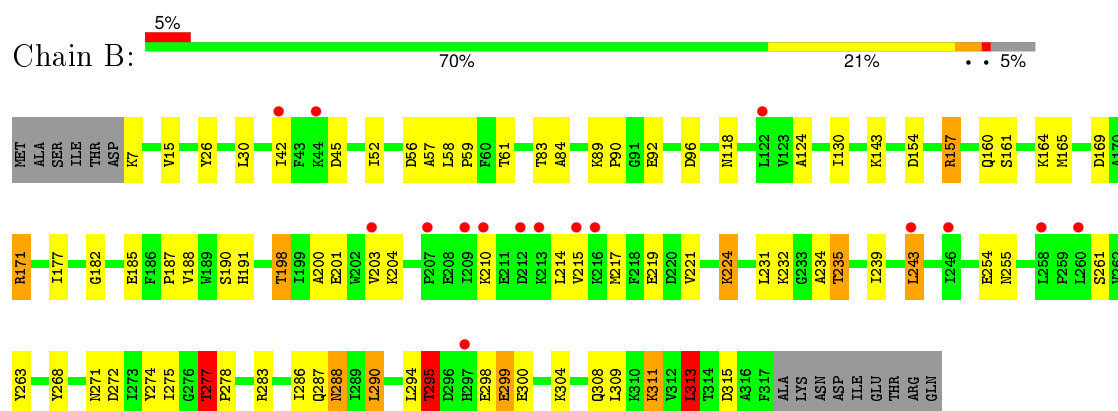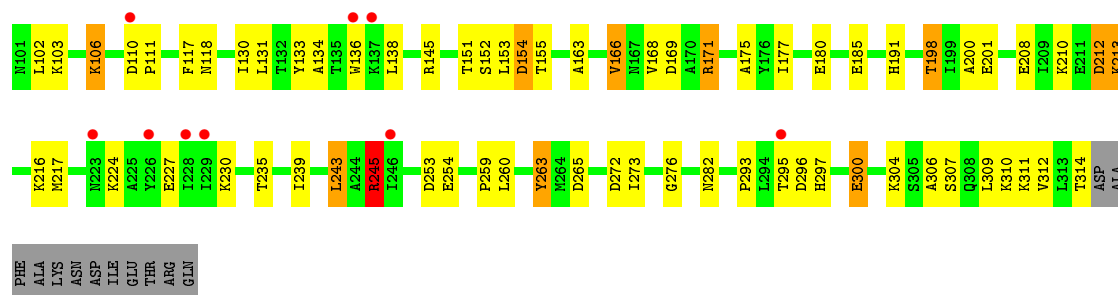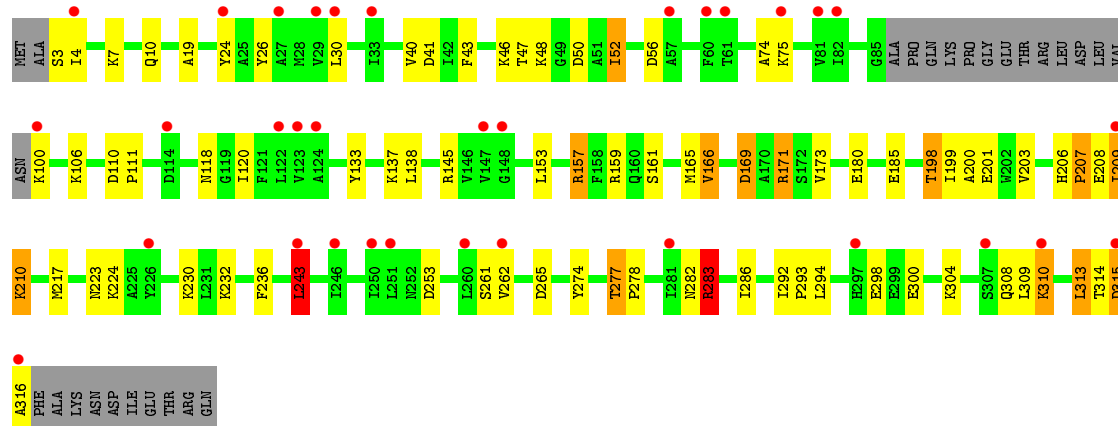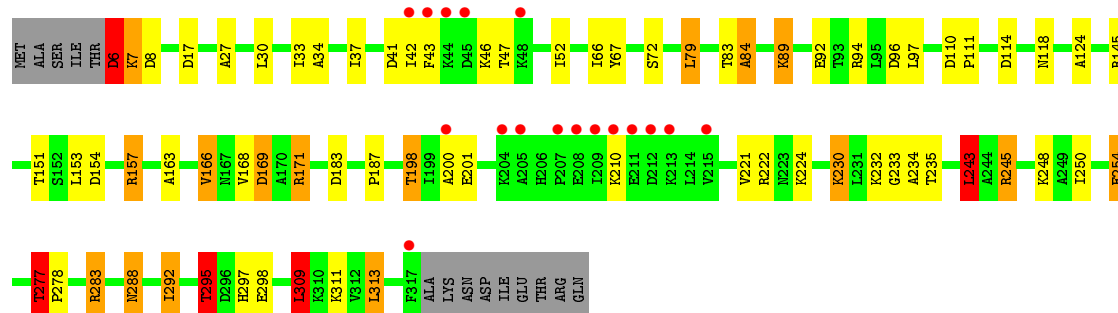

- Molecule 1: L-lactate dehydrogenase
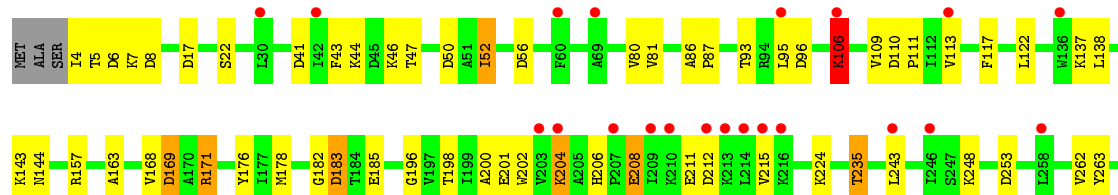

- Molecule 1: L-lactate dehydrogenase

- Molecule 1: L-lactate dehydrogenase

Chain D:

- Molecule 1: L-lactate dehydrogenase

Chain E:

- Molecule 1: L-lactate dehydrogenase

Chain F:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	163.50Å 82.69Å 180.85Å 90.00° 91.49° 90.00°	Depositor
Resolution (Å)	50.00 – 1.96 38.74 – 1.96	Depositor EDS
% Data completeness (in resolution range)	84.1 (50.00-1.96) 84.2 (38.74-1.96)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.99 (at 1.97Å)	Xtriage
Refinement program	REFMAC 5.5.0066	Depositor
R, R_{free}	0.206 , 0.260 0.234 , 0.279	Depositor DCC
R_{free} test set	7275 reflections (5.29%)	DCC
Wilson B-factor (Å^2)	34.0	Xtriage
Anisotropy	0.267	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 55.4	EDS
Estimated twinning fraction	0.032 for -h,-k,l	Xtriage
L-test for twinning [2]	$<|L|>=0.47$, $<L^2>=0.30$	Xtriage
Outliers	0 of 144848 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	14731	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.69% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: SO4

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	0.99	2/2266 (0.1%)	1.06	5/3071 (0.2%)
1	B	1.10	3/2437 (0.1%)	1.17	14/3303 (0.4%)
1	C	1.04	2/2336 (0.1%)	1.01	8/3164 (0.3%)
1	D	1.09	3/2344 (0.1%)	1.14	14/3175 (0.4%)
1	E	1.17	4/2445 (0.2%)	1.32	21/3314 (0.6%)
1	F	0.97	1/2465 (0.0%)	0.97	7/3342 (0.2%)
All	All	1.06	15/14293 (0.1%)	1.12	69/19369 (0.4%)

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	263	TYR	CD1-CE1	7.47	1.50	1.39
1	F	248	LYS	CE-NZ	6.37	1.65	1.49
1	B	221	VAL	CB-CG1	5.99	1.65	1.52
1	D	173	VAL	CB-CG2	5.91	1.65	1.52
1	B	15	VAL	CB-CG1	5.85	1.65	1.52

The worst 5 of 69 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	157	ARG	NE-CZ-NH2	17.28	128.94	120.30
1	E	245	ARG	NE-CZ-NH1	-17.21	111.70	120.30
1	E	171	ARG	NE-CZ-NH1	-16.41	112.09	120.30
1	E	157	ARG	NE-CZ-NH1	-16.24	112.18	120.30
1	E	157	ARG	NE-CZ-NH2	16.01	128.30	120.30

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	2225	0	2246	138	0
1	B	2393	0	2424	87	0
1	C	2296	0	2327	120	0
1	D	2304	0	2331	98	0
1	E	2401	0	2428	85	0
1	F	2421	0	2451	86	0
2	A	10	0	0	1	0
2	B	10	0	0	2	0
2	C	10	0	0	0	0
2	D	5	0	0	0	0
2	E	15	0	0	0	0
2	F	10	0	0	2	0
3	A	77	0	0	4	0
3	B	133	0	0	0	0
3	C	93	0	0	3	0
3	D	101	0	0	1	0
3	E	142	0	0	7	0
3	F	85	0	0	5	0
All	All	14731	0	14207	538	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 19.

The worst 5 of 538 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:C:198:THR:CG2	1:C:201:GLU:H	1.43	1.31
1:D:209:ILE:O	1:D:210:LYS:HG3	1.33	1.27
1:C:224:LYS:HE3	1:E:52:ILE:CG2	1.63	1.27
1:A:44:LYS:HE3	1:A:70:GLU:OE2	1.42	1.20
1:D:224:LYS:HE2	1:F:52:ILE:CG2	1.72	1.18

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	286/326 (88%)	264 (92%)	21 (7%)	1 (0%)	46	35
1	B	309/326 (95%)	294 (95%)	14 (4%)	1 (0%)	46	35
1	C	295/326 (90%)	280 (95%)	15 (5%)	0	100	100
1	D	296/326 (91%)	279 (94%)	13 (4%)	4 (1%)	14	4
1	E	310/326 (95%)	297 (96%)	11 (4%)	2 (1%)	30	16
1	F	313/326 (96%)	298 (95%)	15 (5%)	0	100	100
All	All	1809/1956 (92%)	1712 (95%)	89 (5%)	8 (0%)	39	27

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	8	ASP
1	D	209	ILE
1	D	210	LYS
1	D	315	ASP
1	E	8	ASP

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	238/270 (88%)	223 (94%)	15 (6%)	22	8
1	B	257/270 (95%)	245 (95%)	12 (5%)	32	16
1	C	247/270 (92%)	232 (94%)	15 (6%)	23	9

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	248/270 (92%)	231 (93%)	17 (7%)	19	7
1	E	258/270 (96%)	242 (94%)	16 (6%)	23	8
1	F	260/270 (96%)	246 (95%)	14 (5%)	27	12
All	All	1508/1620 (93%)	1419 (94%)	89 (6%)	24	10

5 of 89 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	307	SER
1	D	232	LYS
1	F	235	THR
1	C	311	LYS
1	D	100	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	287	GLN
1	D	118	ASN
1	F	206	HIS
1	C	118	ASN
1	C	126	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 carbohydrates in this entry.

5.6 Ligand geometry (i)

12 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
2	SO4	A	401	-	4,4,4	0.28	0	6,6,6	0.96	1 (16%)
2	SO4	A	402	-	4,4,4	0.35	0	6,6,6	0.66	0
2	SO4	B	401	-	4,4,4	0.21	0	6,6,6	0.60	0
2	SO4	B	402	-	4,4,4	0.18	0	6,6,6	0.49	0
2	SO4	C	401	-	4,4,4	0.37	0	6,6,6	1.65	1 (16%)
2	SO4	C	402	-	4,4,4	0.34	0	6,6,6	0.45	0
2	SO4	D	401	-	4,4,4	0.66	0	6,6,6	0.97	0
2	SO4	E	327	-	4,4,4	0.86	0	6,6,6	0.81	0
2	SO4	E	401	-	4,4,4	0.23	0	6,6,6	0.69	0
2	SO4	E	402	-	4,4,4	1.11	0	6,6,6	0.61	0
2	SO4	F	401	-	4,4,4	0.22	0	6,6,6	0.41	0
2	SO4	F	402	-	4,4,4	0.50	0	6,6,6	0.28	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SO4	A	401	-	-	0/0/0/0	0/0/0/0
2	SO4	A	402	-	-	0/0/0/0	0/0/0/0
2	SO4	B	401	-	-	0/0/0/0	0/0/0/0
2	SO4	B	402	-	-	0/0/0/0	0/0/0/0
2	SO4	C	401	-	-	0/0/0/0	0/0/0/0
2	SO4	C	402	-	-	0/0/0/0	0/0/0/0
2	SO4	D	401	-	-	0/0/0/0	0/0/0/0
2	SO4	E	327	-	-	0/0/0/0	0/0/0/0
2	SO4	E	401	-	-	0/0/0/0	0/0/0/0
2	SO4	E	402	-	-	0/0/0/0	0/0/0/0
2	SO4	F	401	-	-	0/0/0/0	0/0/0/0
2	SO4	F	402	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	401	SO4	O2-S-O1	-3.61	98.06	109.50
2	A	401	SO4	O2-S-O1	-2.31	102.17	109.50

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	401	SO4	1	0
2	B	401	SO4	2	0
2	F	401	SO4	2	0

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	290/326 (88%)	0.68	28 (9%)	10	16	32, 52, 78, 94	0
1	B	311/326 (95%)	0.50	16 (5%)	32	43	29, 44, 59, 75	0
1	C	299/326 (91%)	0.49	10 (3%)	50	61	28, 50, 69, 85	0
1	D	300/326 (92%)	0.65	33 (11%)	7	12	31, 42, 67, 78	0
1	E	312/326 (95%)	0.25	17 (5%)	29	41	31, 42, 58, 69	0
1	F	315/326 (96%)	0.60	23 (7%)	18	28	33, 51, 68, 82	0
All	All	1827/1956 (93%)	0.52	127 (6%)	19	29	28, 46, 68, 94	0

The worst 5 of 127 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	312	VAL	5.3
1	B	212	ASP	5.0
1	D	316	ALA	4.8
1	F	318	ALA	4.6
1	D	246	ILE	4.6

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)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
2	SO4	F	402	5/5	0.97	0.17	1.65	54,54,59,60	0
2	SO4	F	401	5/5	0.94	0.14	1.14	62,65,65,66	0
2	SO4	A	401	5/5	0.93	0.14	0.43	68,69,72,75	0
2	SO4	B	402	5/5	0.98	0.15	0.13	48,53,55,56	0
2	SO4	B	401	5/5	0.97	0.12	-0.26	59,60,67,67	0
2	SO4	E	402	5/5	0.99	0.09	-0.44	40,45,46,49	0
2	SO4	D	401	5/5	0.96	0.10	-0.57	48,50,57,57	0
2	SO4	A	402	5/5	0.97	0.13	-0.98	47,48,55,56	0
2	SO4	E	401	5/5	0.95	0.09	-1.05	59,61,67,69	0
2	SO4	E	327	5/5	0.98	0.08	-1.17	47,49,52,52	0
2	SO4	C	401	5/5	0.94	0.10	-1.65	62,63,67,70	0
2	SO4	C	402	5/5	0.98	0.09	-2.16	52,55,57,60	0

6.5 Other polymers (i)

There are no such residues in this entry.