wwPDB X-ray Structure Validation Summary Report (i)

May 31, 2020 – 01:00 am BST

PDB ID	:	1FSE
Title	:	CRYSTAL STRUCTURE OF THE BACILLUS SUBTILIS REGULATORY PROTEIN GERE
Authors	:	Ducros, V.M.-A.; Lewis, R.J.; Verma, C.S.; Dodson, E.J.; Leonard, G.; Turkenburg, J.P.; Murshudov, G.N.; Wilkinson, A.J.; Brannigan, J.A.
Deposited on	:	2000-09-08
Resolution	:	2.05 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 following versions of software and data (see references (i)) were used in the production of this report:

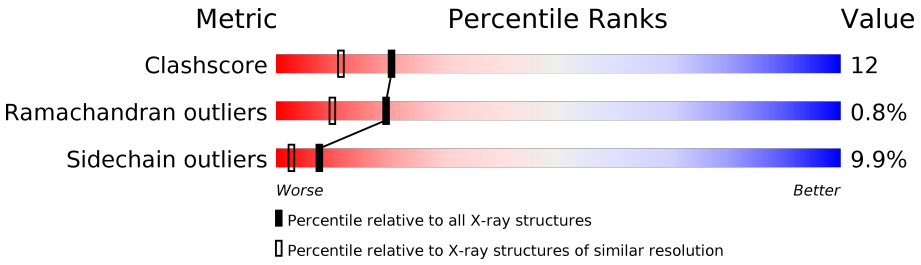
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.05 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))
Clashscore	141614	1773 (2.04-2.04)
Ramachandran outliers	138981	1752 (2.04-2.04)
Sidechain outliers	138945	1752 (2.04-2.04)

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 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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	74	70% 19% • 9%
1	B	74	68% 20% • • 5%
1	C	74	74% 15% • 9%
1	D	74	62% 22% 5% 11%
1	E	74	58% 20% 8% 14%
1	F	74	42% 23% • 32%

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	GOL	D	601	-	X	-	-

2 Entry composition (i)

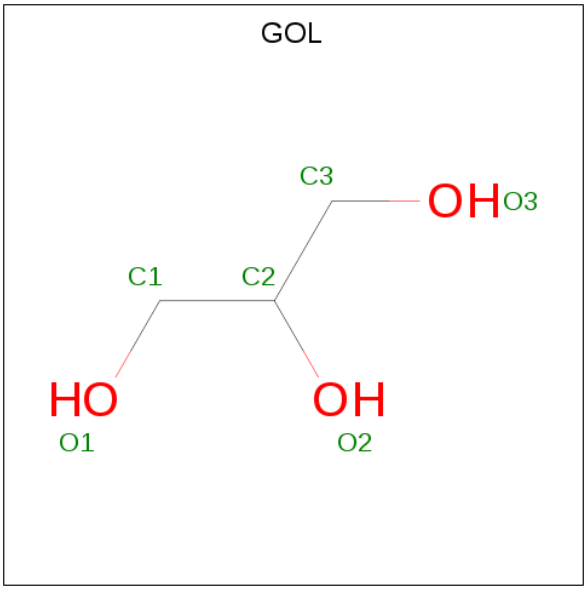
There are 4 unique types of molecules in this entry. The entry contains 3484 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 GERE.

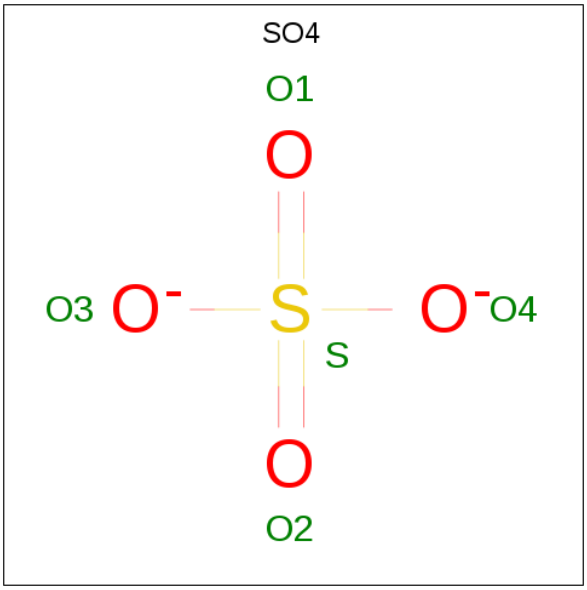
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	67	Total	C	N	O	S	0	3	0
			543	341	97	103	2			
1	B	70	Total	C	N	O	S	0	2	0
			571	358	103	108	2			
1	C	67	Total	C	N	O	S	0	2	0
			540	339	96	103	2			
1	D	66	Total	C	N	O	S	0	4	0
			541	340	97	102	2			
1	E	64	Total	C	N	O	S	0	4	0
			530	331	98	99	2			
1	F	50	Total	C	N	O	S	0	0	0
			399	250	75	72	2			

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total C O 6 3 3	0	0
2	D	1	Total C O 6 3 3	0	0
2	F	1	Total C O 6 3 3	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	C	1	Total O S 5 4 1	0	0
3	C	1	Total O S 5 4 1	0	0
3	D	1	Total O S 5 4 1	0	0
3	D	1	Total O S 5 4 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	43	Total O 43 43	0	0
4	B	66	Total O 66 66	0	0
4	C	69	Total O 69 69	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	D	74	Total O 74 74	0	0
4	E	34	Total O 34 34	0	0
4	F	36	Total O 36 36	0	0

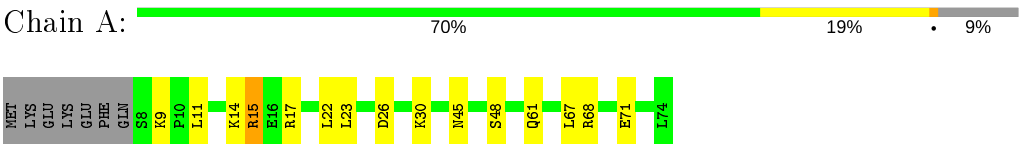
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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. 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. 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.

Note EDS was not executed.

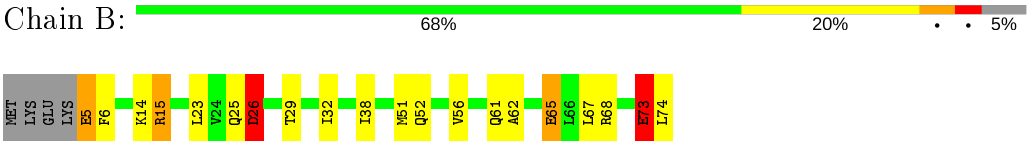
- Molecule 1: GERE

Chain A:



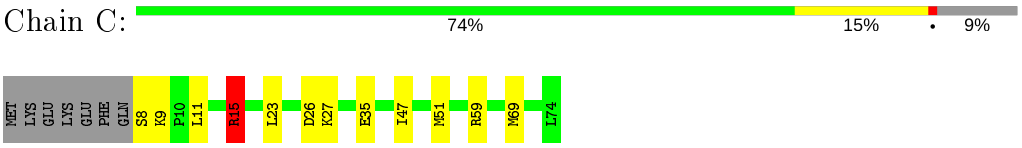
- Molecule 1: GERE

Chain B:



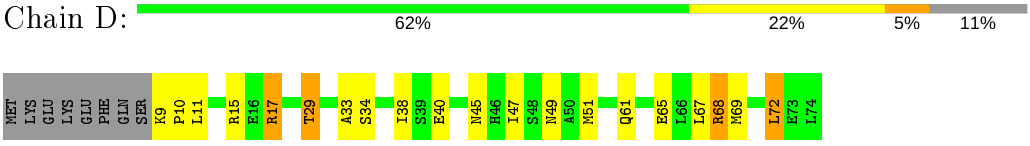
- Molecule 1: GERE

Chain C:



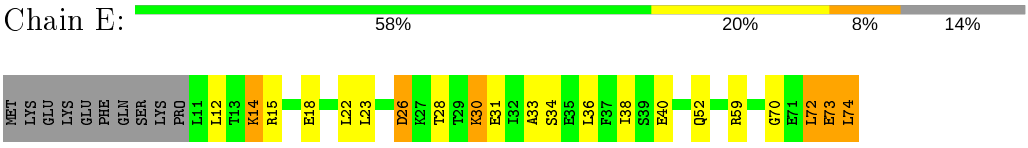
- Molecule 1: GERE

Chain D:



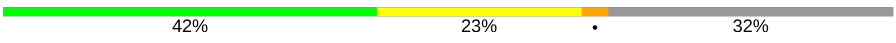
- Molecule 1: GERE

Chain E:



- Molecule 1: GERE

Chain F:




MET
LYS
GLU
LYS
GLU
PHE
GLN
SER
LYS
P10
L11
L12
R17
E18
E21
L22
L23
V24
Q25
D26
K27
THR
THR
LYS
GLU
ILE
ALA
SER
GLU
LEU
PHE
ILE
SER
GLU
LYS
THR
V43
R44
N45
H46
I47
S48
N49
Q52
K53
L54
L66
L67
R68
M69
L74

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	109.02Å 61.75Å 71.74Å 90.00° 97.08° 90.00°	Depositor
Resolution (Å)	20.00 – 2.05	Depositor
% Data completeness (in resolution range)	99.0 (20.00-2.05)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC	Depositor
R, R_{free}	0.214 , 0.272	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3484	wwPDB-VP
Average B, all atoms (Å^2)	18.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, 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	0.38	0/561	1.11	4/748 (0.5%)
1	B	0.42	0/585	1.28	6/781 (0.8%)
1	C	0.50	0/553	1.21	2/739 (0.3%)
1	D	0.38	0/564	1.15	3/754 (0.4%)
1	E	0.33	0/552	1.00	1/735 (0.1%)
1	F	0.34	0/400	1.10	2/532 (0.4%)
All	All	0.40	0/3215	1.15	18/4289 (0.4%)

There are no bond length outliers.

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	15	ARG	NE-CZ-NH1	11.06	125.83	120.30
1	B	15	ARG	NE-CZ-NH2	-10.28	115.16	120.30
1	C	59	ARG	NE-CZ-NH2	-9.20	115.70	120.30
1	B	68	ARG	NE-CZ-NH1	-8.14	116.23	120.30
1	B	29	THR	OG1-CB-CG2	-6.58	94.87	110.00

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	543	0	580	15	0
1	B	571	0	597	18	1
1	C	540	0	573	8	0
1	D	541	0	578	15	0
1	E	530	0	566	13	0
1	F	399	0	422	17	0
2	B	6	0	8	2	0
2	D	6	0	7	3	0
2	F	6	0	8	2	0
3	C	10	0	0	0	0
3	D	10	0	0	0	0
4	A	43	0	0	2	0
4	B	66	0	0	3	1
4	C	69	0	0	2	0
4	D	74	0	0	3	0
4	E	34	0	0	0	0
4	F	36	0	0	3	0
All	All	3484	0	3339	81	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 12.

The worst 5 of 81 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:61[A]:GLN:HG2	4:B:651:HOH:O	1.63	0.98
1:B:38:ILE:HA	2:D:601:GOL:H2	1.48	0.96
1:A:14[B]:LYS:HG2	4:A:91:HOH:O	1.69	0.91
1:B:5:GLU:HG2	1:B:6:PHE:H	1.32	0.91
1:F:26:ASP:OD2	4:F:602:HOH:O	1.94	0.86

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:B:5:GLU:OE2	4:B:628:HOH:O[2_656]	2.03	0.17

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	68/74 (92%)	65 (96%)	3 (4%)	0	100 100
1	B	70/74 (95%)	69 (99%)	1 (1%)	0	100 100
1	C	67/74 (90%)	67 (100%)	0	0	100 100
1	D	68/74 (92%)	68 (100%)	0	0	100 100
1	E	66/74 (89%)	63 (96%)	1 (2%)	2 (3%)	4 0
1	F	46/74 (62%)	45 (98%)	0	1 (2%)	6 1
All	All	385/444 (87%)	377 (98%)	5 (1%)	3 (1%)	19 10

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	12	LEU
1	E	73	GLU
1	F	25	GLN

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	63/68 (93%)	60 (95%)	3 (5%)	25 18
1	B	65/68 (96%)	60 (92%)	5 (8%)	13 5
1	C	62/68 (91%)	57 (92%)	5 (8%)	11 5
1	D	63/68 (93%)	53 (84%)	10 (16%)	2 0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	E	61/68 (90%)	51 (84%)	10 (16%)	2 0
1	F	44/68 (65%)	39 (89%)	5 (11%)	5 1
All	All	358/408 (88%)	320 (89%)	38 (11%)	8 2

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	29[A]	THR
1	D	72[A]	LEU
1	F	48	SER
1	D	34	SER
1	D	72[B]	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	25	GLN
1	E	25	GLN
1	F	46	HIS
1	F	49	ASN
1	F	52	GLN

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)

7 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	GOL	F	600	-	5,5,5	0.28	0	5,5,5	1.00	0
3	SO4	C	502	-	4,4,4	0.34	0	6,6,6	0.65	0
2	GOL	D	601	-	5,5,5	1.04	1 (20%)	5,5,5	2.89	4 (80%)
3	SO4	D	503	-	4,4,4	0.33	0	6,6,6	0.35	0
3	SO4	D	504	-	4,4,4	0.15	0	6,6,6	0.98	0
2	GOL	B	602	-	5,5,5	0.46	0	5,5,5	1.28	1 (20%)
3	SO4	C	501	-	4,4,4	0.19	0	6,6,6	0.80	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	GOL	F	600	-	-	4/4/4/4	-
2	GOL	D	601	-	-	3/4/4/4	-
2	GOL	B	602	-	-	0/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	601	GOL	O2-C2	-2.23	1.36	1.43

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	601	GOL	O2-C2-C3	-3.96	91.67	109.12
2	D	601	GOL	O2-C2-C1	-3.73	92.69	109.12
2	B	602	GOL	C3-C2-C1	2.43	121.16	111.70
2	D	601	GOL	O3-C3-C2	2.20	120.75	110.20
2	D	601	GOL	C3-C2-C1	2.03	119.62	111.70

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	F	600	GOL	O1-C1-C2-C3
2	D	601	GOL	O1-C1-C2-C3
2	F	600	GOL	O1-C1-C2-O2
2	F	600	GOL	C1-C2-C3-O3
2	F	600	GOL	O2-C2-C3-O3

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	600	GOL	2	0
2	D	601	GOL	3	0
2	B	602	GOL	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

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.