



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 27, 2026 – 05:14 PM EDT

PDB ID : 2HQT / pdb_00002hqt
Title : Crystal structures of the interacting domains from yeast glutamyl-tRNA synthetase and tRNA aminoacylation and nuclear export cofactor Arc1p reveal a novel function for an old fold
Authors : Simader, H.; Hothorn, M.; Suck, D.
Deposited on : 2006-07-19
Resolution : 1.90 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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

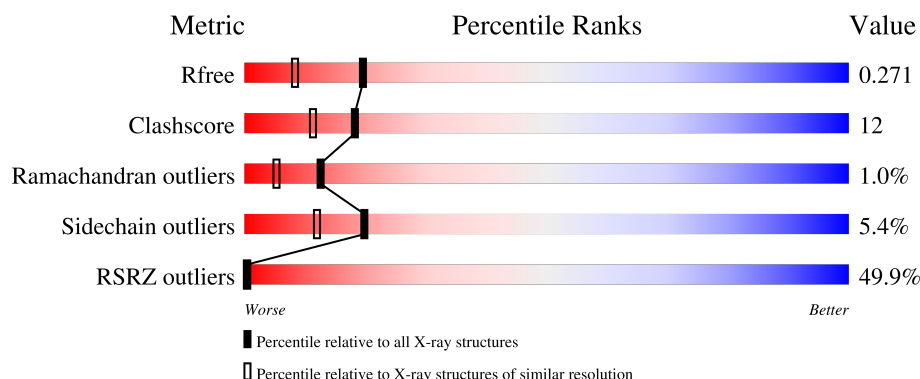
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.90 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}	180053	7789 (1.90-1.90)
Clashscore	190562	8410 (1.90-1.90)
Ramachandran outliers	187476	8333 (1.90-1.90)
Sidechain outliers	187428	8333 (1.90-1.90)
RSRZ outliers	180081	7790 (1.90-1.90)

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	124	55% 66% 20% 6% • 7%
1	B	124	35% 79% 17% • •
1	C	124	46% 67% 22% • 8%
1	D	124	32% 81% 15% • •
1	E	124	62% 73% 18% • 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	124	
1	G	124	
1	H	124	
1	I	124	
1	J	124	
1	K	124	
1	L	124	
1	M	124	
1	N	124	
1	O	124	
1	P	124	
1	Q	124	
1	R	124	
1	S	124	
1	T	124	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 19976 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 GU4 nucleic-binding protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	115	Total	C	N	O	S	0	0	0
			911	584	148	178	1			
1	B	120	Total	C	N	O	S	0	0	0
			954	614	154	185	1			
1	C	114	Total	C	N	O	S	0	0	0
			887	570	145	171	1			
1	D	120	Total	C	N	O	S	0	0	0
			957	615	155	186	1			
1	E	116	Total	C	N	O	S	0	0	0
			902	579	148	174	1			
1	F	120	Total	C	N	O	S	0	0	0
			953	613	153	185	2			
1	G	120	Total	C	N	O	S	0	0	0
			940	604	153	182	1			
1	H	111	Total	C	N	O	S	0	0	0
			864	560	141	162	1			
1	I	116	Total	C	N	O	S	0	0	0
			902	578	148	175	1			
1	J	121	Total	C	N	O	S	0	0	0
			955	615	154	185	1			
1	K	117	Total	C	N	O	S	0	0	0
			919	591	151	176	1			
1	L	119	Total	C	N	O	S	0	0	0
			950	611	154	184	1			
1	M	117	Total	C	N	O	S	0	0	0
			938	604	151	181	2			
1	N	119	Total	C	N	O	S	0	0	0
			921	593	151	176	1			
1	O	119	Total	C	N	O	S	0	0	0
			936	601	154	180	1			
1	P	119	Total	C	N	O	S	0	0	0
			944	608	153	181	2			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Q	117	Total	C	N	O	S	0	0	0
			927	596	151	178	2			
1	R	119	Total	C	N	O	S	0	0	0
			941	606	153	181	1			
1	S	116	Total	C	N	O	S	0	0	0
			895	577	147	170	1			
1	T	121	Total	C	N	O	S	0	0	0
			965	621	156	186	2			

There are 40 discrepancies between the modelled and reference sequences:

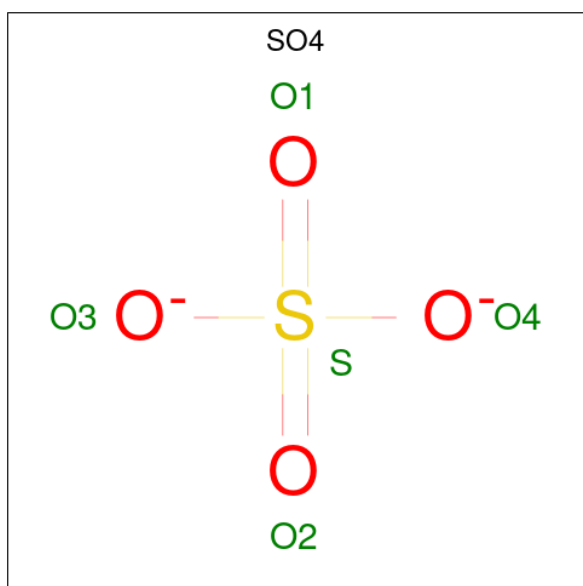
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	cloning artifact	UNP P46672
A	2	HIS	-	cloning artifact	UNP P46672
B	1	GLY	-	cloning artifact	UNP P46672
B	2	HIS	-	cloning artifact	UNP P46672
C	1	GLY	-	cloning artifact	UNP P46672
C	2	HIS	-	cloning artifact	UNP P46672
D	1	GLY	-	cloning artifact	UNP P46672
D	2	HIS	-	cloning artifact	UNP P46672
E	1	GLY	-	cloning artifact	UNP P46672
E	2	HIS	-	cloning artifact	UNP P46672
F	1	GLY	-	cloning artifact	UNP P46672
F	2	HIS	-	cloning artifact	UNP P46672
G	1	GLY	-	cloning artifact	UNP P46672
G	2	HIS	-	cloning artifact	UNP P46672
H	1	GLY	-	cloning artifact	UNP P46672
H	2	HIS	-	cloning artifact	UNP P46672
I	1	GLY	-	cloning artifact	UNP P46672
I	2	HIS	-	cloning artifact	UNP P46672
J	1	GLY	-	cloning artifact	UNP P46672
J	2	HIS	-	cloning artifact	UNP P46672
K	1	GLY	-	cloning artifact	UNP P46672
K	2	HIS	-	cloning artifact	UNP P46672
L	1	GLY	-	cloning artifact	UNP P46672
L	2	HIS	-	cloning artifact	UNP P46672
M	1	GLY	-	cloning artifact	UNP P46672
M	2	HIS	-	cloning artifact	UNP P46672
N	1	GLY	-	cloning artifact	UNP P46672
N	2	HIS	-	cloning artifact	UNP P46672
O	1	GLY	-	cloning artifact	UNP P46672
O	2	HIS	-	cloning artifact	UNP P46672
P	1	GLY	-	cloning artifact	UNP P46672

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
P	2	HIS	-	cloning artifact	UNP P46672
Q	1	GLY	-	cloning artifact	UNP P46672
Q	2	HIS	-	cloning artifact	UNP P46672
R	1	GLY	-	cloning artifact	UNP P46672
R	2	HIS	-	cloning artifact	UNP P46672
S	1	GLY	-	cloning artifact	UNP P46672
S	2	HIS	-	cloning artifact	UNP P46672
T	1	GLY	-	cloning artifact	UNP P46672
T	2	HIS	-	cloning artifact	UNP P46672

- Molecule 2 is SULFATE ION (CCD ID: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	E	1	Total	O	S	0	0
			5	4	1		
2	E	1	Total	O	S	0	0
			5	4	1		
2	I	1	Total	O	S	0	0
			5	4	1		
2	J	1	Total	O	S	0	0
			5	4	1		
2	M	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	N	1	Total	O	S	0	0
			5	4	1		
2	Q	1	Total	O	S	0	0
			5	4	1		
2	S	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	75	Total	O	0	0
			75	75		
3	B	97	Total	O	0	0
			97	97		
3	C	40	Total	O	0	0
			40	40		
3	D	82	Total	O	0	0
			82	82		
3	E	61	Total	O	0	0
			61	61		
3	F	55	Total	O	0	0
			55	55		
3	G	86	Total	O	0	0
			86	86		
3	H	44	Total	O	0	0
			44	44		
3	I	57	Total	O	0	0
			57	57		
3	J	77	Total	O	0	0
			77	77		
3	K	49	Total	O	0	0
			49	49		
3	L	56	Total	O	0	0
			56	56		
3	M	64	Total	O	0	0
			64	64		
3	N	52	Total	O	0	0
			52	52		
3	O	80	Total	O	0	0
			80	80		
3	P	77	Total	O	0	0
			77	77		

Continued on next page...

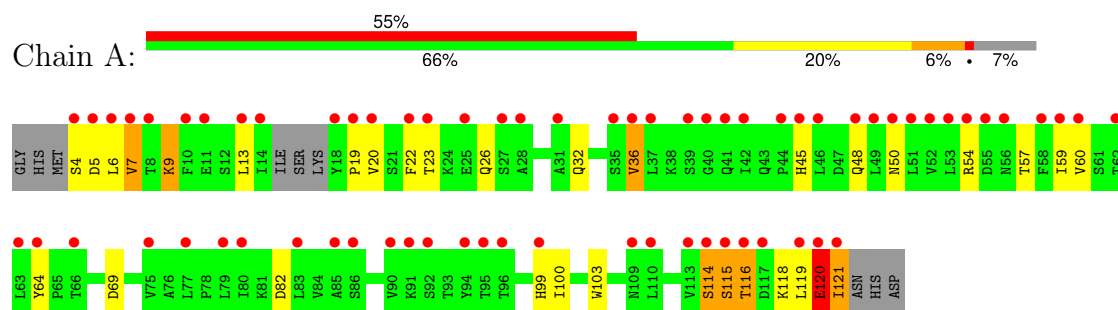
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	Q	84	Total 84	O 84	0	0
3	R	93	Total 93	O 93	0	0
3	S	62	Total 62	O 62	0	0
3	T	74	Total 74	O 74	0	0

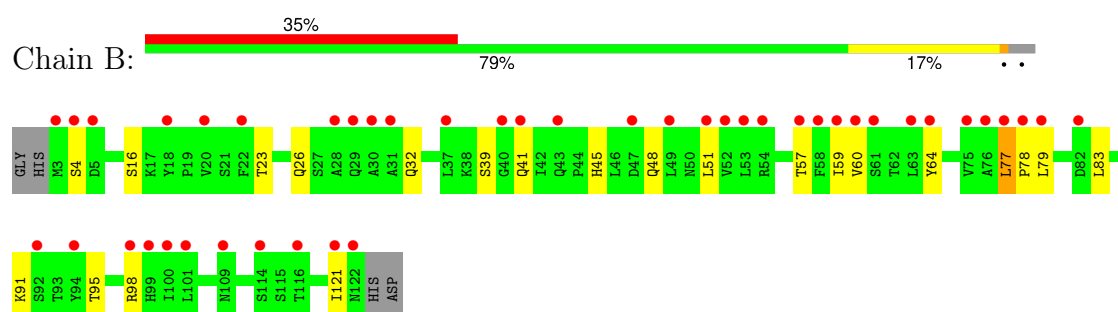
3 Residue-property plots [i](#)

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.

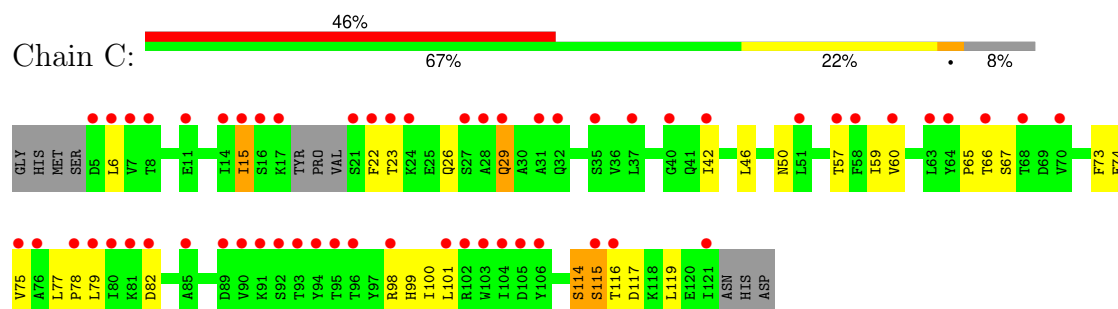
• Molecule 1: GU4 nucleic-binding protein 1



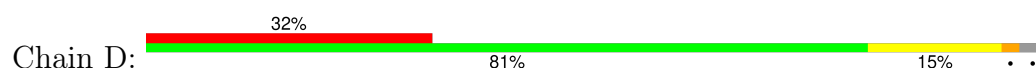
• Molecule 1: GU4 nucleic-binding protein 1

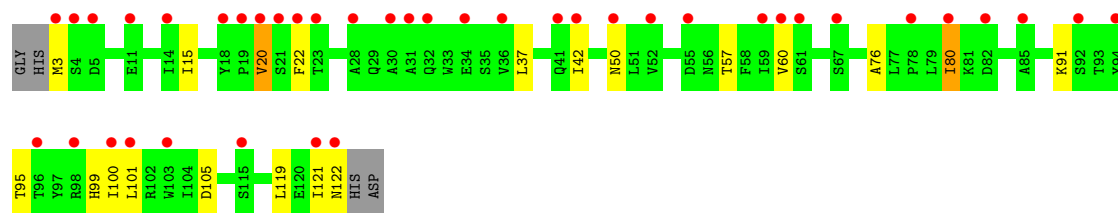


• Molecule 1: GU4 nucleic-binding protein 1

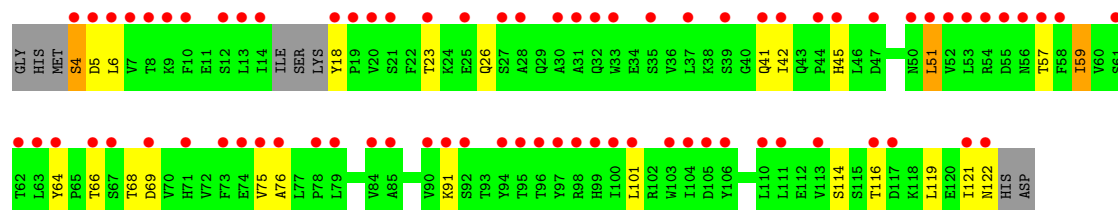
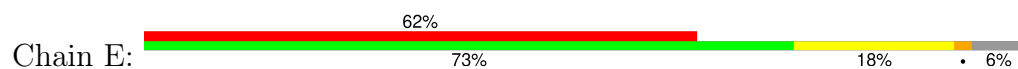


• Molecule 1: GU4 nucleic-binding protein 1

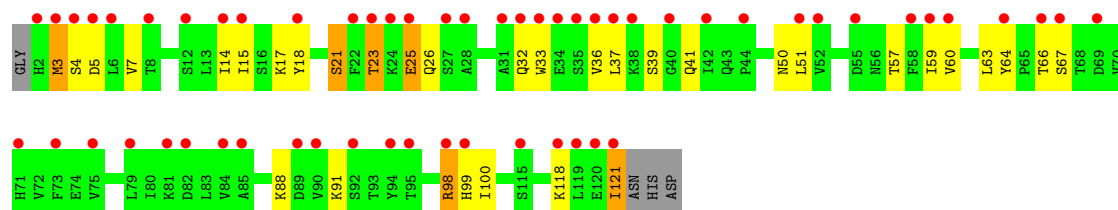




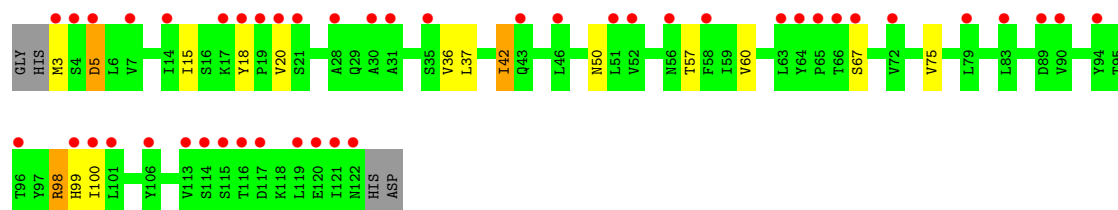
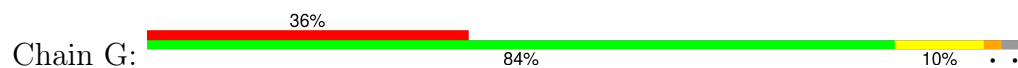
- Molecule 1: GU4 nucleic-binding protein 1



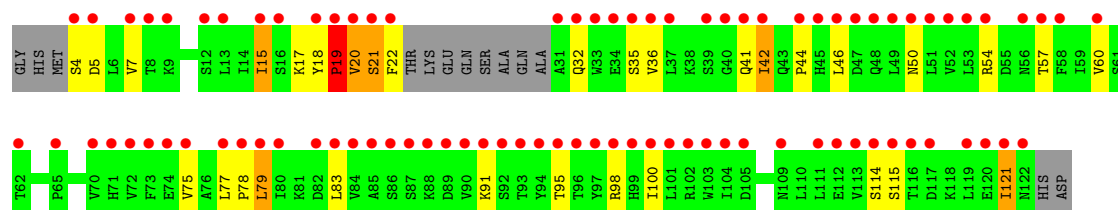
- Molecule 1: GU4 nucleic-binding protein 1



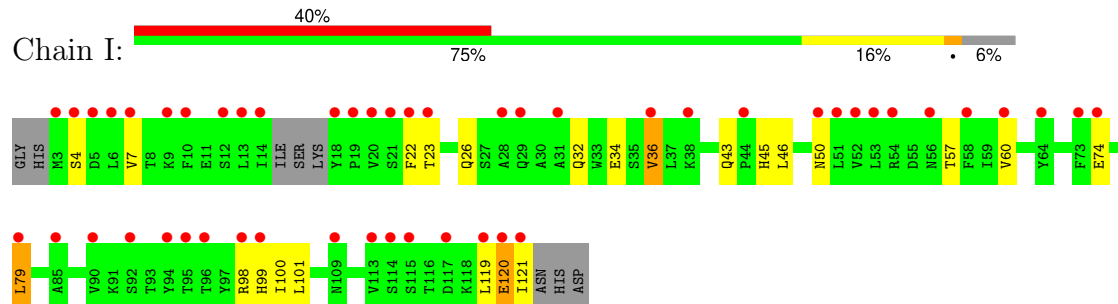
- Molecule 1: GU4 nucleic-binding protein 1



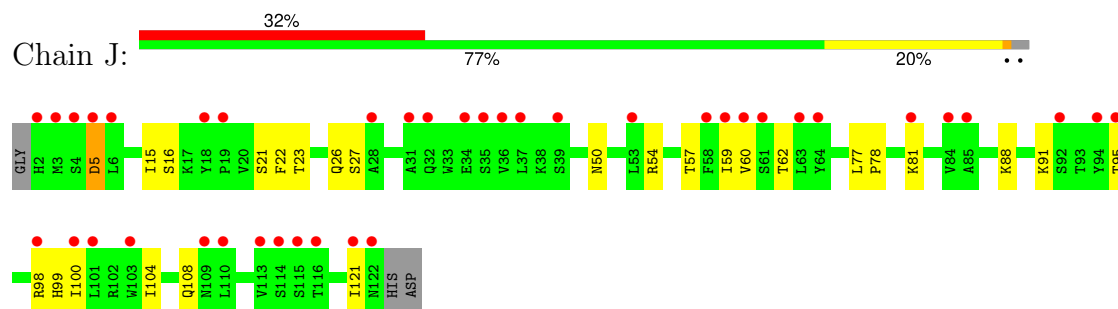
- Molecule 1: GU4 nucleic-binding protein 1



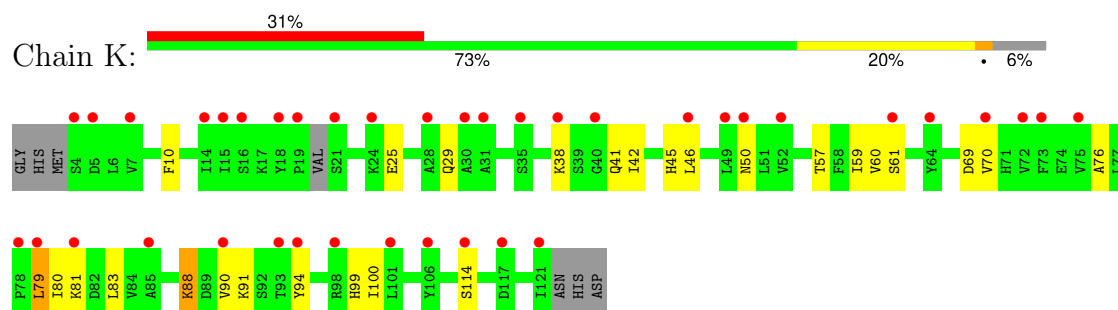
- Molecule 1: GU4 nucleic-binding protein 1



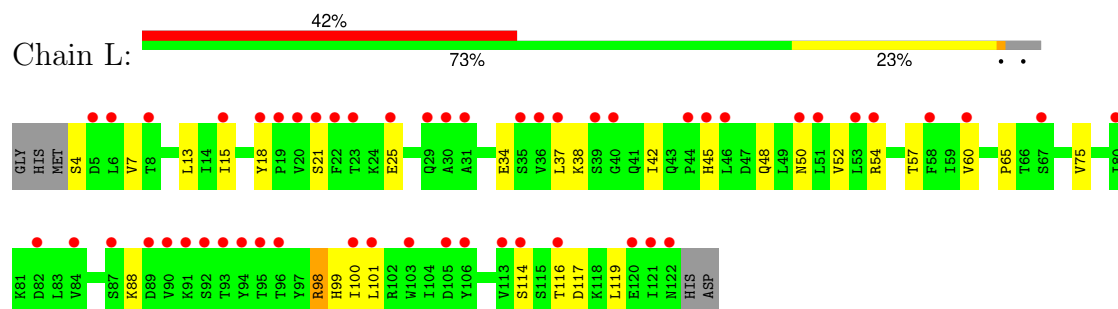
- Molecule 1: GU4 nucleic-binding protein 1



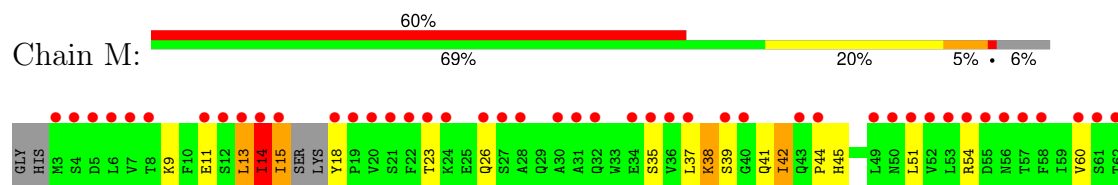
- Molecule 1: GU4 nucleic-binding protein 1

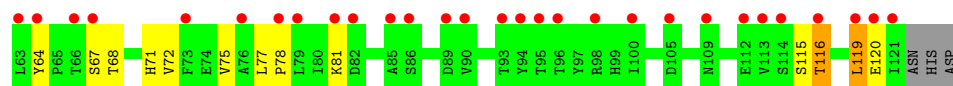


- Molecule 1: GU4 nucleic-binding protein 1

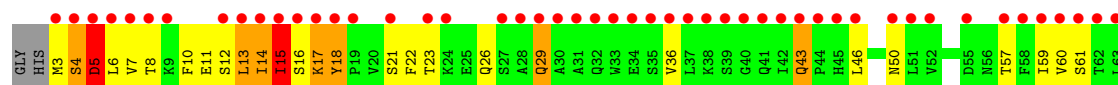
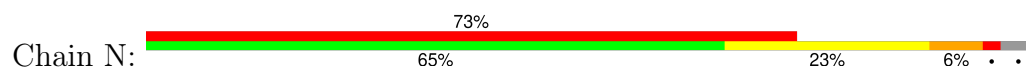


- Molecule 1: GU4 nucleic-binding protein 1

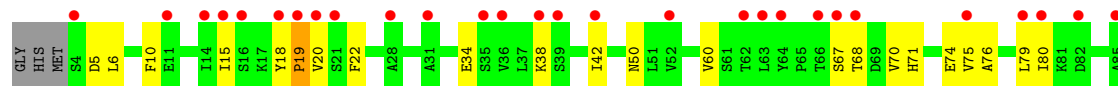




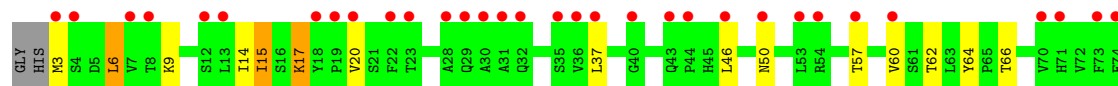
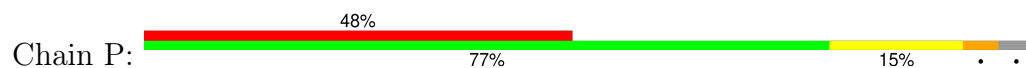
- Molecule 1: GU4 nucleic-binding protein 1



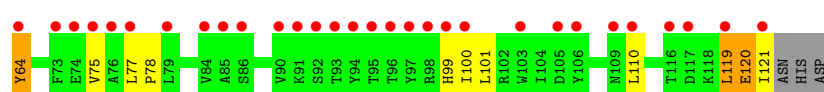
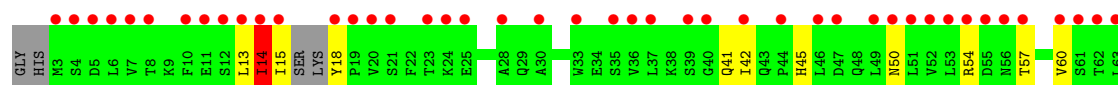
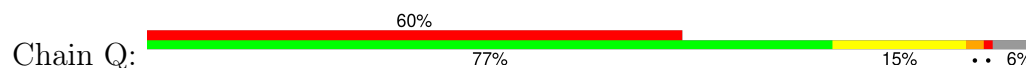
- Molecule 1: GU4 nucleic-binding protein 1



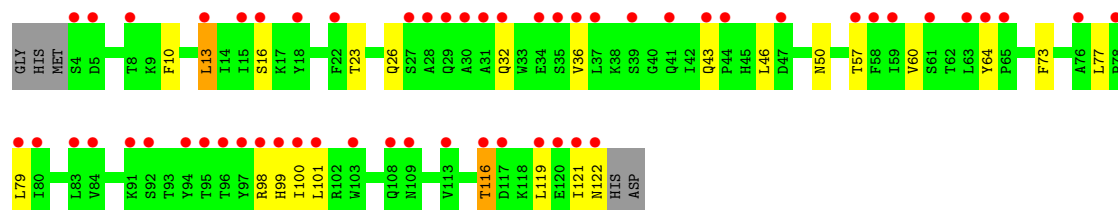
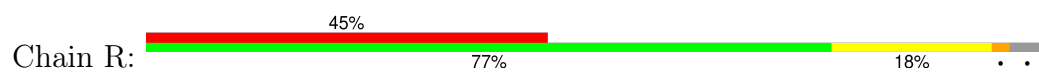
- Molecule 1: GU4 nucleic-binding protein 1



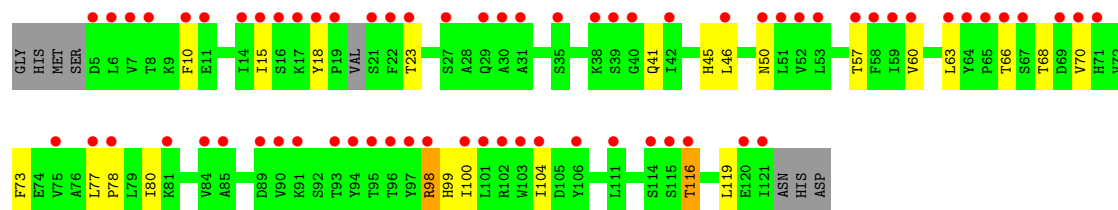
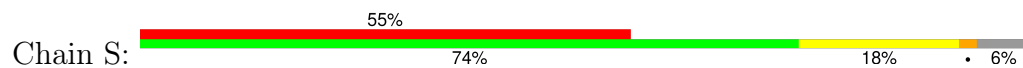
- Molecule 1: GU4 nucleic-binding protein 1



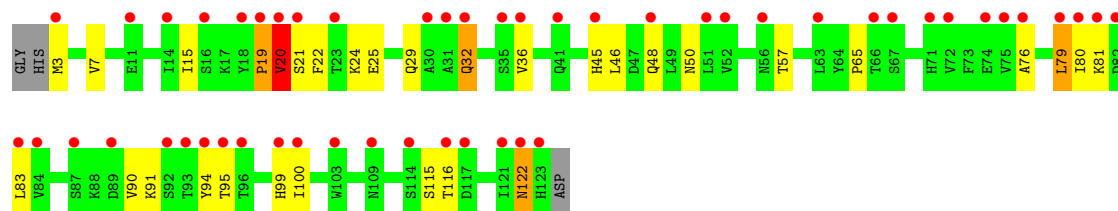
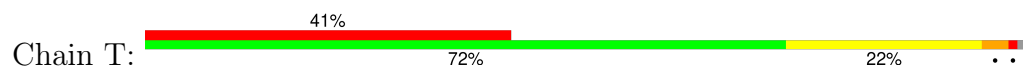
- Molecule 1: GU4 nucleic-binding protein 1



- Molecule 1: GU4 nucleic-binding protein 1



- Molecule 1: GU4 nucleic-binding protein 1



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	222.32Å 89.46Å 126.79Å 90.00° 99.39° 90.00°	Depositor
Resolution (Å)	50.00 – 1.90 50.00 – 1.90	Depositor EDS
% Data completeness (in resolution range)	97.0 (50.00-1.90) 97.0 (50.00-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.04	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.09 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.209 , 0.262 0.221 , 0.271	Depositor DCC
R_{free} test set	9383 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	30.5	Xtriage
Anisotropy	0.035	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 56.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	19976	wwPDB-VP
Average B, all atoms (Å ²)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 52.55 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 4.7859e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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

5.1 Standard geometry

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	0.90	0/928	0.99	0/1265
1	B	0.97	0/973	1.01	0/1327
1	C	0.82	0/903	0.95	1/1233 (0.1%)
1	D	0.92	0/976	0.91	0/1331
1	E	0.90	1/919 (0.1%)	0.93	1/1255 (0.1%)
1	F	0.91	0/972	0.96	1/1326 (0.1%)
1	G	0.96	1/958 (0.1%)	0.97	0/1307
1	H	0.83	0/881	0.90	0/1204
1	I	0.78	0/918	0.93	0/1252
1	J	0.91	0/974	0.95	0/1329
1	K	0.73	0/936	0.86	0/1275
1	L	0.80	0/969	0.98	0/1320
1	M	0.76	0/956	0.94	0/1301
1	N	0.83	1/940 (0.1%)	0.95	0/1285
1	O	0.92	0/953	1.04	1/1298 (0.1%)
1	P	0.81	0/963	0.87	0/1313
1	Q	0.92	0/944	1.02	2/1285 (0.2%)
1	R	0.99	0/960	1.00	0/1310
1	S	0.88	1/912 (0.1%)	0.94	0/1246
1	T	0.92	0/984	0.92	1/1340 (0.1%)
All	All	0.88	4/18919 (0.0%)	0.95	7/25802 (0.0%)

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	2
1	D	0	1
1	E	0	2

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
1	G	0	2
1	H	0	1
1	M	0	2
1	N	0	2
1	O	0	1
1	S	0	1
1	T	0	3
All	All	0	17

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	N	36	VAL	C-O	6.25	1.31	1.24
1	E	114	SER	C-O	6.05	1.32	1.23
1	S	104	ILE	CA-C	-5.76	1.45	1.52
1	G	98	ARG	C-O	-5.00	1.18	1.24

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	T	122	ASN	N-CA-C	-5.76	106.90	114.04
1	O	19	PRO	N-CA-CB	5.74	109.28	103.25
1	E	59	ILE	N-CA-C	5.54	116.27	110.62
1	C	116	THR	N-CA-C	-5.42	108.44	114.62
1	Q	64	TYR	CA-C-N	-5.34	114.28	119.78

There are no chirality outliers.

5 of 17 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	114	SER	Peptide
1	A	120	GLU	Peptide
1	D	3	MET	Peptide
1	E	18	TYR	Peptide
1	E	4	SER	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	911	0	902	27	0
1	B	954	0	952	13	0
1	C	887	0	861	23	0
1	D	957	0	956	13	0
1	E	902	0	877	24	0
1	F	953	0	948	37	0
1	G	940	0	929	13	0
1	H	864	0	846	28	0
1	I	902	0	876	15	0
1	J	955	0	948	19	0
1	K	919	0	913	18	0
1	L	950	0	952	25	0
1	M	938	0	940	35	0
1	N	921	0	898	50	0
1	O	936	0	932	32	0
1	P	944	0	941	19	0
1	Q	927	0	929	26	0
1	R	941	0	934	17	0
1	S	895	0	868	17	0
1	T	965	0	970	36	0
2	A	5	0	0	0	0
2	B	5	0	0	1	0
2	E	10	0	0	1	0
2	I	5	0	0	0	0
2	J	5	0	0	0	0
2	M	5	0	0	0	0
2	N	5	0	0	0	0
2	Q	5	0	0	0	0
2	S	5	0	0	0	0
3	A	75	0	0	0	0
3	B	97	0	0	1	0
3	C	40	0	0	0	0
3	D	82	0	0	0	0
3	E	61	0	0	0	0
3	F	55	0	0	3	0
3	G	86	0	0	0	0
3	H	44	0	0	1	0
3	I	57	0	0	0	0
3	J	77	0	0	1	0
3	K	49	0	0	3	0
3	L	56	0	0	0	0
3	M	64	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	N	52	0	0	1	0
3	O	80	0	0	4	0
3	P	77	0	0	1	0
3	Q	84	0	0	2	0
3	R	93	0	0	2	0
3	S	62	0	0	1	0
3	T	74	0	0	0	0
All	All	19976	0	18372	446	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 12.

The worst 5 of 446 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:T:20:VAL:HG12	1:T:21:SER:CA	1.37	1.54
1:H:20:VAL:HA	1:H:21:SER:CB	1.46	1.38
1:T:19:PRO:O	1:T:20:VAL:HG23	1.23	1.27
1:N:10:PHE:O	1:N:13:LEU:HD23	1.39	1.19
1:N:15:ILE:N	1:N:15:ILE:HD12	1.60	1.15

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	111/124 (90%)	106 (96%)	2 (2%)	3 (3%)	41
1	B	118/124 (95%)	116 (98%)	2 (2%)	0	100100
1	C	110/124 (89%)	103 (94%)	6 (6%)	1 (1%)	146
1	D	118/124 (95%)	114 (97%)	4 (3%)	0	100100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	112/124 (90%)	107 (96%)	4 (4%)	1 (1%)	14	6
1	F	118/124 (95%)	116 (98%)	1 (1%)	1 (1%)	16	8
1	G	118/124 (95%)	114 (97%)	3 (2%)	1 (1%)	16	8
1	H	107/124 (86%)	102 (95%)	2 (2%)	3 (3%)	4	0
1	I	112/124 (90%)	109 (97%)	3 (3%)	0	100	100
1	J	119/124 (96%)	115 (97%)	4 (3%)	0	100	100
1	K	113/124 (91%)	110 (97%)	3 (3%)	0	100	100
1	L	117/124 (94%)	114 (97%)	3 (3%)	0	100	100
1	M	113/124 (91%)	106 (94%)	4 (4%)	3 (3%)	4	1
1	N	117/124 (94%)	109 (93%)	4 (3%)	4 (3%)	3	0
1	O	117/124 (94%)	112 (96%)	2 (2%)	3 (3%)	4	1
1	P	117/124 (94%)	113 (97%)	4 (3%)	0	100	100
1	Q	113/124 (91%)	110 (97%)	1 (1%)	2 (2%)	6	1
1	R	117/124 (94%)	115 (98%)	2 (2%)	0	100	100
1	S	112/124 (90%)	107 (96%)	4 (4%)	1 (1%)	14	6
1	T	119/124 (96%)	114 (96%)	4 (3%)	1 (1%)	16	8
All	All	2298/2480 (93%)	2212 (96%)	62 (3%)	24 (1%)	12	5

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	7	VAL
1	A	120	GLU
1	C	22	PHE
1	E	5	ASP
1	F	3	MET

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	105/117 (90%)	97 (92%)	8 (8%)	12	5
1	B	110/117 (94%)	105 (96%)	5 (4%)	24	17
1	C	98/117 (84%)	90 (92%)	8 (8%)	10	5
1	D	111/117 (95%)	107 (96%)	4 (4%)	31	23
1	E	100/117 (86%)	98 (98%)	2 (2%)	48	46
1	F	110/117 (94%)	103 (94%)	7 (6%)	16	8
1	G	106/117 (91%)	103 (97%)	3 (3%)	38	32
1	H	95/117 (81%)	86 (90%)	9 (10%)	8	3
1	I	100/117 (86%)	95 (95%)	5 (5%)	22	14
1	J	109/117 (93%)	102 (94%)	7 (6%)	16	8
1	K	104/117 (89%)	98 (94%)	6 (6%)	18	10
1	L	110/117 (94%)	105 (96%)	5 (4%)	24	17
1	M	109/117 (93%)	101 (93%)	8 (7%)	13	6
1	N	102/117 (87%)	92 (90%)	10 (10%)	7	3
1	O	106/117 (91%)	104 (98%)	2 (2%)	50	47
1	P	108/117 (92%)	99 (92%)	9 (8%)	10	4
1	Q	107/117 (92%)	104 (97%)	3 (3%)	38	32
1	R	107/117 (92%)	104 (97%)	3 (3%)	38	32
1	S	97/117 (83%)	94 (97%)	3 (3%)	35	29
1	T	112/117 (96%)	106 (95%)	6 (5%)	20	12
All	All	2106/2340 (90%)	1993 (95%)	113 (5%)	20	12

5 of 113 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	J	121	ILE
1	T	81	LYS
1	M	15	ILE
1	T	79	LEU
1	Q	15	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 63 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	50	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	R	43	GLN
1	L	29	GLN
1	R	41	GLN
1	T	32	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

10 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	E	2004	-	4,4,4	0.33	0	6,6,6	0.48	0
2	SO4	S	2010	-	4,4,4	0.18	0	6,6,6	0.70	0
2	SO4	Q	2009	-	4,4,4	0.21	0	6,6,6	0.39	0
2	SO4	I	2005	-	4,4,4	0.24	0	6,6,6	0.41	0
2	SO4	A	2002	-	4,4,4	0.27	0	6,6,6	0.55	0
2	SO4	J	2006	-	4,4,4	0.26	0	6,6,6	0.46	0
2	SO4	M	2008	-	4,4,4	0.28	0	6,6,6	0.34	0
2	SO4	N	2007	-	4,4,4	0.20	0	6,6,6	0.67	0
2	SO4	B	2001	-	4,4,4	0.14	0	6,6,6	0.41	0
2	SO4	E	2003	-	4,4,4	0.29	0	6,6,6	0.49	0

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.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	2004	SO4	1	0
2	B	2001	SO4	1	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

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	115/124 (92%)	2.34	68 (59%)	0 0	24, 46, 56, 57	0
1	B	120/124 (96%)	1.88	44 (36%)	1 0	42, 46, 52, 60	0
1	C	114/124 (91%)	1.95	57 (50%)	0 0	39, 44, 50, 55	0
1	D	120/124 (96%)	1.67	40 (33%)	1 1	40, 45, 56, 62	0
1	E	116/124 (93%)	2.55	77 (66%)	0 0	20, 45, 51, 59	0
1	F	120/124 (96%)	2.17	57 (47%)	0 0	24, 47, 53, 60	0
1	G	120/124 (96%)	2.05	45 (37%)	1 0	25, 46, 58, 64	0
1	H	111/124 (89%)	3.07	88 (79%)	0 0	27, 44, 55, 62	0
1	I	116/124 (93%)	2.16	50 (43%)	0 0	40, 46, 52, 60	0
1	J	121/124 (97%)	1.86	40 (33%)	1 1	20, 44, 51, 57	0
1	K	117/124 (94%)	1.76	39 (33%)	1 1	40, 46, 53, 60	0
1	L	119/124 (95%)	2.02	52 (43%)	0 0	35, 44, 51, 60	0
1	M	117/124 (94%)	2.54	74 (63%)	0 0	38, 43, 53, 60	0
1	N	119/124 (95%)	2.92	91 (76%)	0 0	21, 45, 52, 60	0
1	O	119/124 (95%)	1.85	46 (38%)	1 0	40, 45, 57, 66	0
1	P	119/124 (95%)	2.16	59 (49%)	0 0	40, 44, 53, 58	0
1	Q	117/124 (94%)	2.41	74 (63%)	0 0	40, 45, 56, 62	0
1	R	119/124 (95%)	2.15	56 (47%)	0 0	42, 47, 53, 68	0
1	S	116/124 (93%)	2.23	68 (58%)	0 0	40, 45, 56, 62	0
1	T	121/124 (97%)	1.94	51 (42%)	0 0	25, 46, 54, 65	0
All	All	2356/2480 (95%)	2.18	1176 (49%)	0 0	20, 45, 54, 68	0

The worst 5 of 1176 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	N	14	ILE	10.6
1	J	122	ASN	9.3
1	E	4	SER	8.9
1	E	18	TYR	8.8
1	H	4	SER	8.4

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 oligosaccharides 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($\text{\AA}^2$)	Q<0.9
2	SO4	M	2008	5/5	0.79	0.18	66,68,69,69	0
2	SO4	Q	2009	5/5	0.82	0.14	71,71,72,72	0
2	SO4	J	2006	5/5	0.83	0.18	72,73,74,74	0
2	SO4	A	2002	5/5	0.85	0.14	60,64,65,65	0
2	SO4	N	2007	5/5	0.86	0.16	64,65,67,67	0
2	SO4	B	2001	5/5	0.87	0.16	61,63,64,66	0
2	SO4	S	2010	5/5	0.87	0.16	61,62,63,64	0
2	SO4	I	2005	5/5	0.91	0.12	68,68,68,69	0
2	SO4	E	2003	5/5	0.91	0.16	54,58,58,58	0
2	SO4	E	2004	5/5	0.92	0.12	65,65,66,67	0

6.5 Other polymers [i](#)

There are no such residues in this entry.