



# wwPDB X-ray Structure Validation Summary Report

Feb 22, 2024 – 12:15 PM EST

PDB ID : 4S0R  
Title : Structure of GS-ThrA complex  
Authors : Schumacher, M.A.; Chinnam, N.G.; Cuthbert, B.; Tonthat, N.K.  
Deposited on : 2015-01-04  
Resolution : 3.50 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](mailto: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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

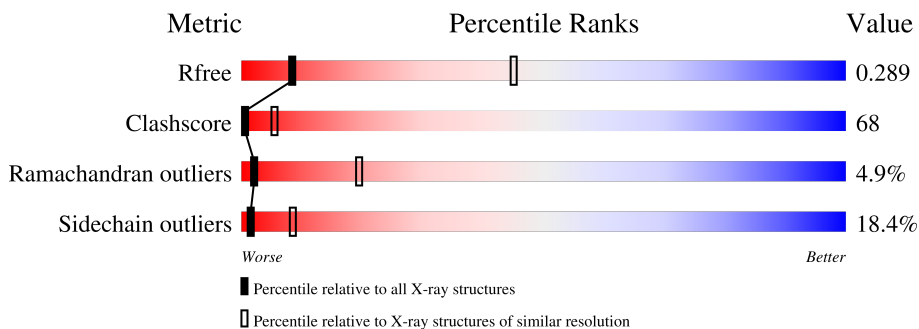
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.50 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.




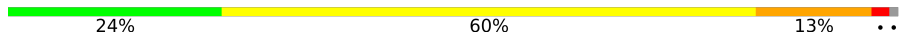


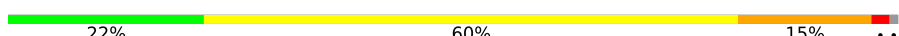
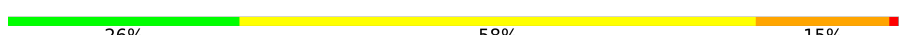
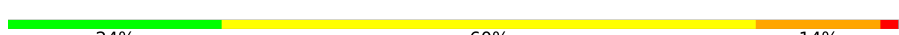
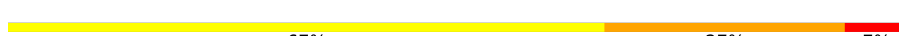



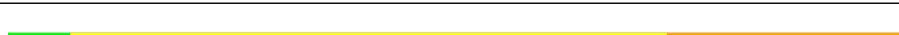

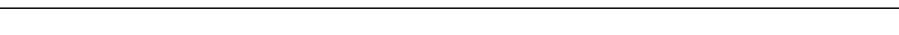





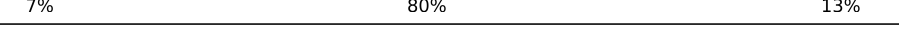

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)

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  $\geq 3$ , 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	447	
1	B	447	
1	C	447	
1	D	447	
1	E	447	
1	F	447	
1	G	447	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	H	447	
1	I	447	
1	J	447	
1	K	447	
1	L	447	
1	M	447	
1	N	447	
2	1	15	
2	2	15	
2	O	15	
2	P	15	
2	Q	15	
2	R	15	
2	S	15	
2	T	15	
2	U	15	
2	V	15	
2	W	15	
2	X	15	
2	Y	15	
2	Z	15	

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
3	GLN	B	501	-	-	X	-

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 51555 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 Glutamine synthetase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	443	3535	2259	590	670	16	0	0	0
1	B	443	3535	2259	590	670	16	0	0	0
1	C	443	3535	2259	590	670	16	0	0	0
1	D	447	3563	2275	596	675	17	0	0	0
1	E	443	3535	2259	590	670	16	0	0	0
1	F	443	3535	2259	590	670	16	0	0	0
1	G	443	3535	2259	590	670	16	0	0	0
1	H	443	3535	2259	590	670	16	0	0	0
1	I	443	3535	2259	590	670	16	0	0	0
1	J	443	3535	2259	590	670	16	0	0	0
1	K	443	3535	2259	590	670	16	0	0	0
1	L	443	3535	2259	590	670	16	0	0	0
1	M	443	3535	2259	590	670	16	0	0	0
1	N	443	3535	2259	590	670	16	0	0	0

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP P12425

*Continued on next page...*

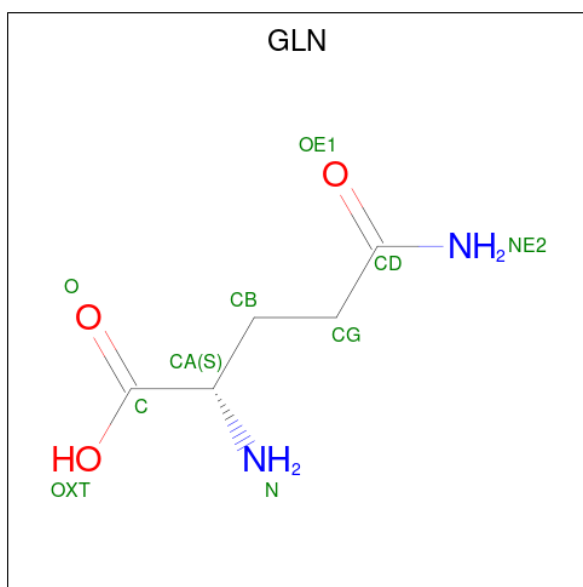
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	SER	-	expression tag	UNP P12425
A	0	HIS	-	expression tag	UNP P12425
B	-2	GLY	-	expression tag	UNP P12425
B	-1	SER	-	expression tag	UNP P12425
B	0	HIS	-	expression tag	UNP P12425
C	-2	GLY	-	expression tag	UNP P12425
C	-1	SER	-	expression tag	UNP P12425
C	0	HIS	-	expression tag	UNP P12425
D	-2	GLY	-	expression tag	UNP P12425
D	-1	SER	-	expression tag	UNP P12425
D	0	HIS	-	expression tag	UNP P12425
E	-2	GLY	-	expression tag	UNP P12425
E	-1	SER	-	expression tag	UNP P12425
E	0	HIS	-	expression tag	UNP P12425
F	-2	GLY	-	expression tag	UNP P12425
F	-1	SER	-	expression tag	UNP P12425
F	0	HIS	-	expression tag	UNP P12425
G	-2	GLY	-	expression tag	UNP P12425
G	-1	SER	-	expression tag	UNP P12425
G	0	HIS	-	expression tag	UNP P12425
H	-2	GLY	-	expression tag	UNP P12425
H	-1	SER	-	expression tag	UNP P12425
H	0	HIS	-	expression tag	UNP P12425
I	-2	GLY	-	expression tag	UNP P12425
I	-1	SER	-	expression tag	UNP P12425
I	0	HIS	-	expression tag	UNP P12425
J	-2	GLY	-	expression tag	UNP P12425
J	-1	SER	-	expression tag	UNP P12425
J	0	HIS	-	expression tag	UNP P12425
K	-2	GLY	-	expression tag	UNP P12425
K	-1	SER	-	expression tag	UNP P12425
K	0	HIS	-	expression tag	UNP P12425
L	-2	GLY	-	expression tag	UNP P12425
L	-1	SER	-	expression tag	UNP P12425
L	0	HIS	-	expression tag	UNP P12425
M	-2	GLY	-	expression tag	UNP P12425
M	-1	SER	-	expression tag	UNP P12425
M	0	HIS	-	expression tag	UNP P12425
N	-2	GLY	-	expression tag	UNP P12425
N	-1	SER	-	expression tag	UNP P12425
N	0	HIS	-	expression tag	UNP P12425

- Molecule 2 is a protein called TnrA peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	O	15	Total 133	C 82	N 28	O 22	S 1	0	0	0
2	P	15	Total 133	C 82	N 28	O 22	S 1	0	0	0
2	Q	15	Total 133	C 82	N 28	O 22	S 1	0	0	0
2	R	15	Total 133	C 82	N 28	O 22	S 1	0	0	0
2	S	15	Total 133	C 82	N 28	O 22	S 1	0	0	0
2	T	15	Total 133	C 82	N 28	O 22	S 1	0	0	0
2	U	15	Total 133	C 82	N 28	O 22	S 1	0	0	0
2	V	15	Total 133	C 82	N 28	O 22	S 1	0	0	0
2	W	15	Total 133	C 82	N 28	O 22	S 1	0	0	0
2	X	15	Total 133	C 82	N 28	O 22	S 1	0	0	0
2	Y	15	Total 133	C 82	N 28	O 22	S 1	0	0	0
2	Z	15	Total 133	C 82	N 28	O 22	S 1	0	0	0
2	1	15	Total 132	C 82	N 28	O 21	S 1	0	0	0
2	2	15	Total 133	C 82	N 28	O 22	S 1	0	0	0

- Molecule 3 is GLUTAMINE (three-letter code: GLN) (formula: C<sub>5</sub>H<sub>10</sub>N<sub>2</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	Total 9	C 5	N 2	O 2	0	0
3	B	1	Total 9	C 5	N 2	O 2	0	0
3	C	1	Total 10	C 5	N 2	O 3	0	0
3	D	1	Total 9	C 5	N 2	O 2	0	0
3	E	1	Total 9	C 5	N 2	O 2	0	0
3	F	1	Total 9	C 5	N 2	O 2	0	0
3	G	1	Total 9	C 5	N 2	O 2	0	0
3	H	1	Total 9	C 5	N 2	O 2	0	0
3	I	1	Total 9	C 5	N 2	O 2	0	0
3	J	1	Total 9	C 5	N 2	O 2	0	0
3	K	1	Total 9	C 5	N 2	O 2	0	0
3	L	1	Total 10	C 5	N 2	O 3	0	0
3	M	1	Total 10	C 5	N 2	O 3	0	0
3	N	1	Total 9	C 5	N 2	O 2	0	0

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

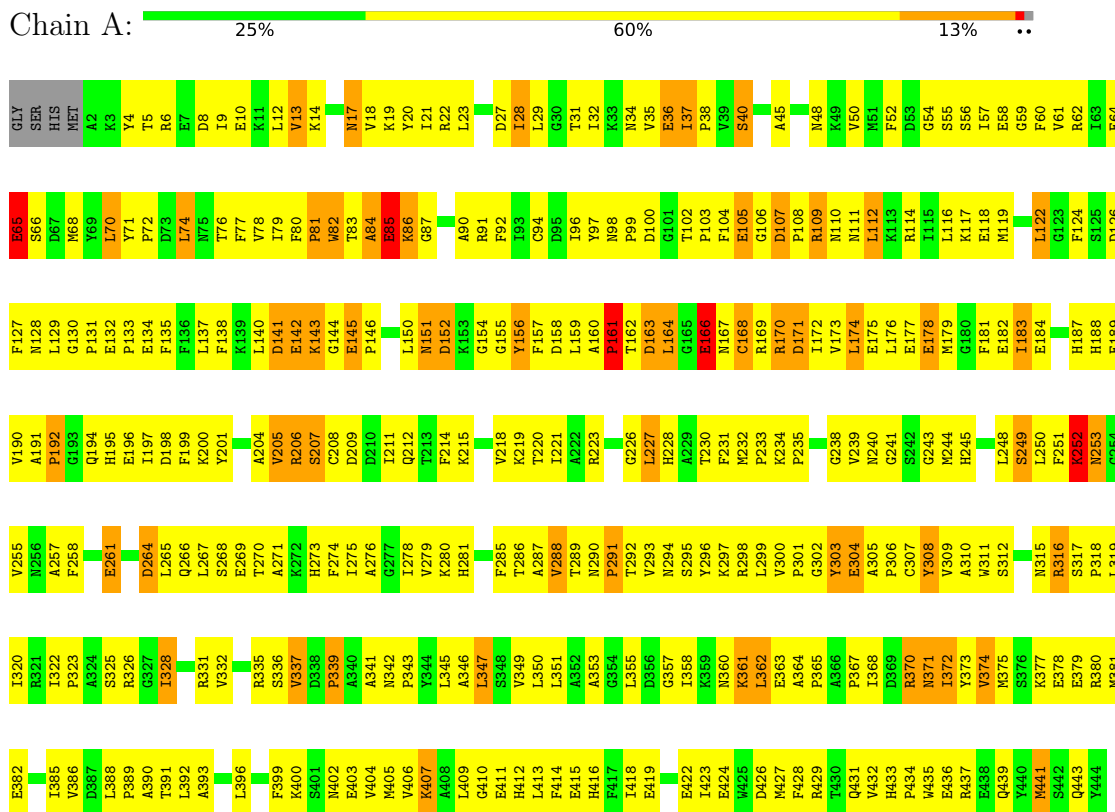
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	4	Total 4	Mg 4	0	0
4	B	3	Total 3	Mg 3	0	0
4	C	7	Total 7	Mg 7	0	0
4	D	8	Total 8	Mg 8	0	0
4	E	2	Total 2	Mg 2	0	0
4	F	3	Total 3	Mg 3	0	0
4	G	1	Total 1	Mg 1	0	0
4	H	3	Total 3	Mg 3	0	0
4	I	2	Total 2	Mg 2	0	0
4	J	3	Total 3	Mg 3	0	0
4	K	3	Total 3	Mg 3	0	0
4	L	3	Total 3	Mg 3	0	0
4	M	3	Total 3	Mg 3	0	0
4	N	2	Total 2	Mg 2	0	0



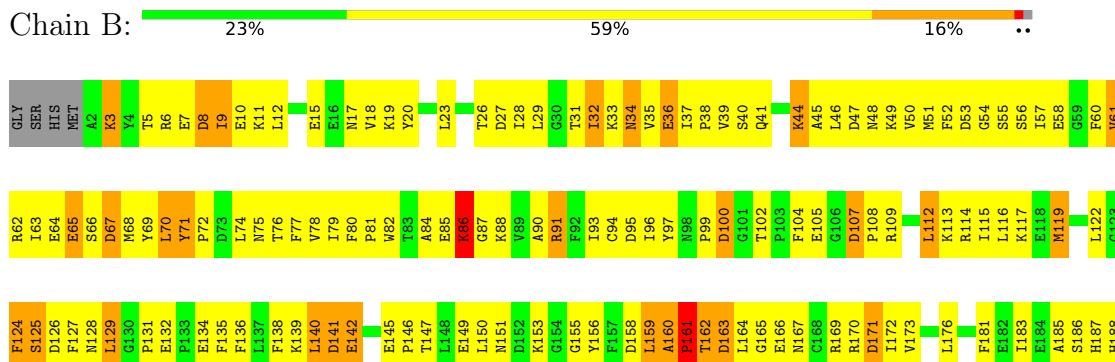
### 3 Residue-property plots

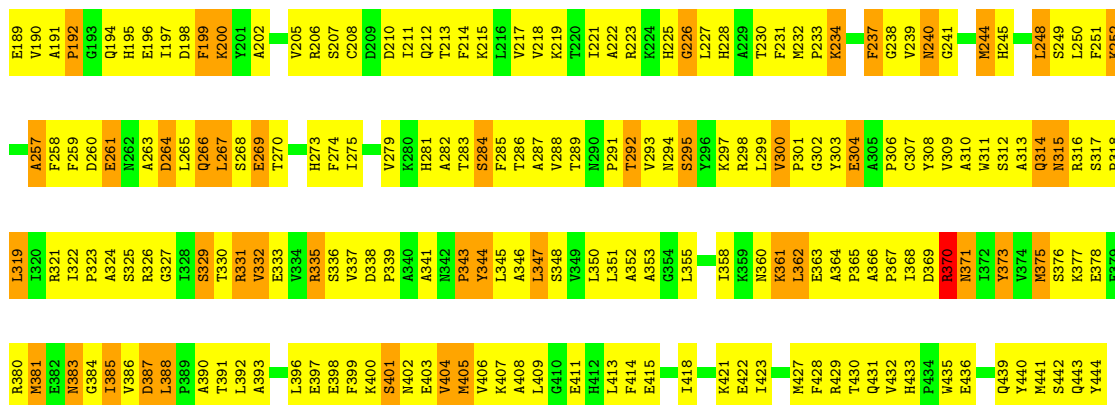
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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.

- Molecule 1: Glutamine synthetase

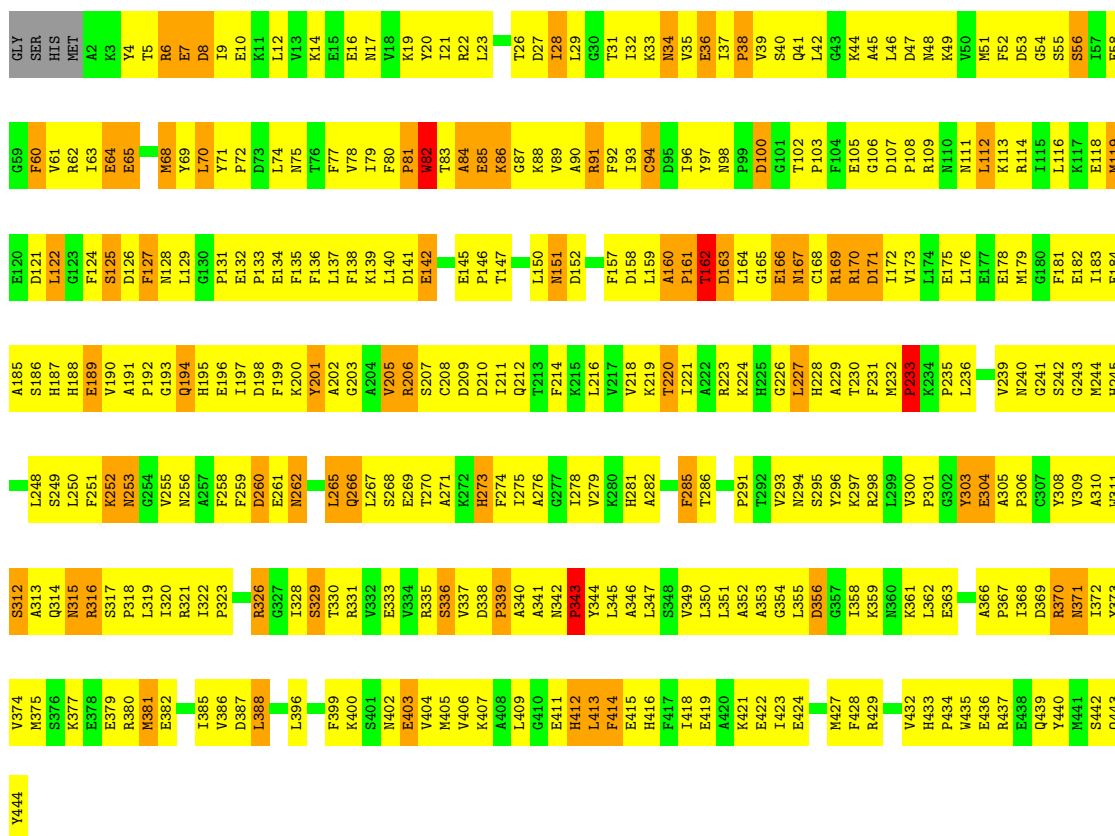


- Molecule 1: Glutamine synthetase

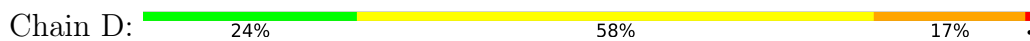


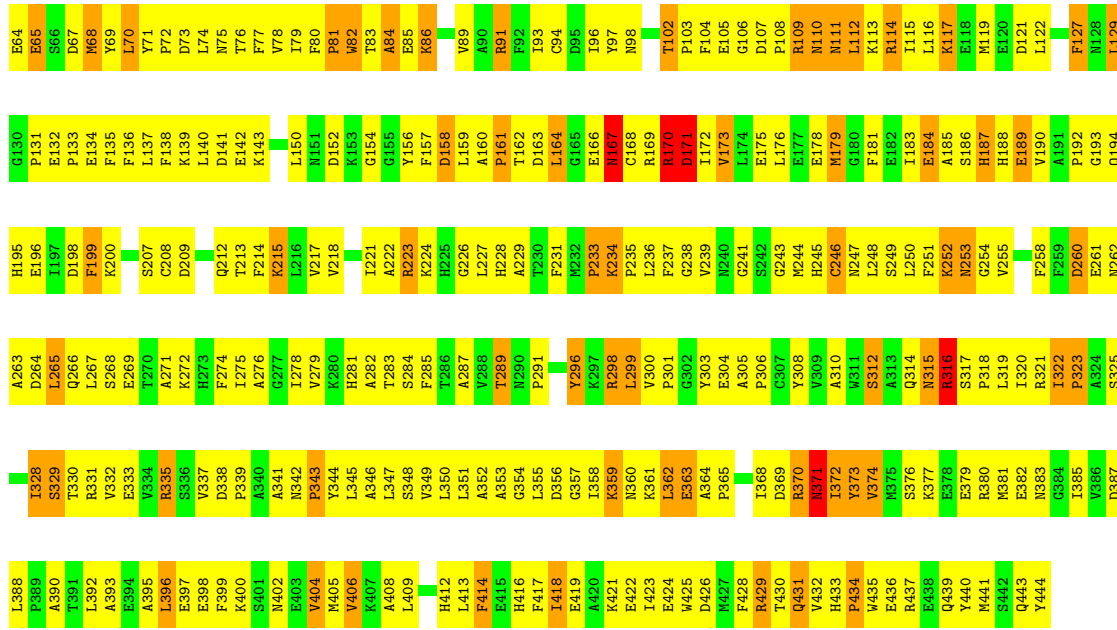


• Molecule 1: Glutamine synthetase

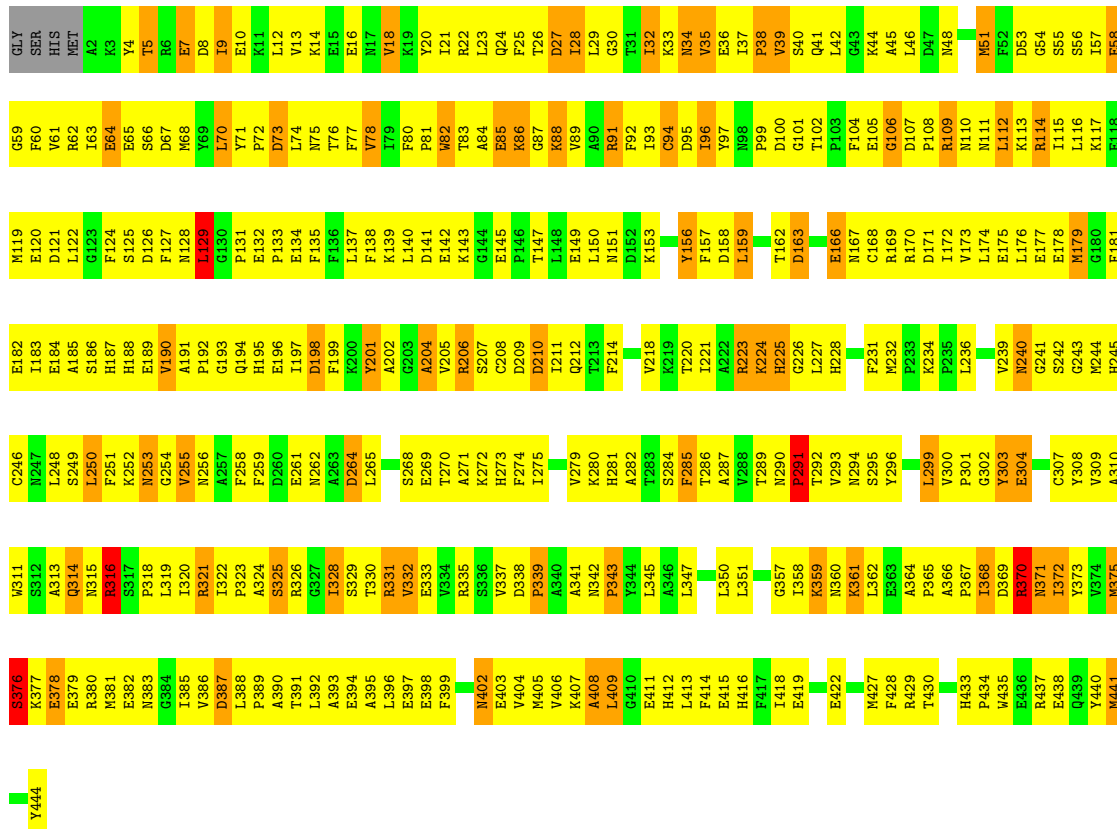


• Molecule 1: Glutamine synthetase

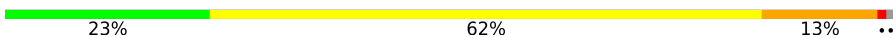




• Molecule 1: Glutamine synthetase



• Molecule 1: Glutamine synthetase

Chain F:  23% 62% 13% ..

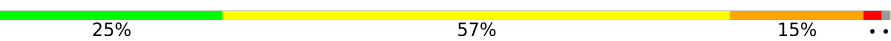
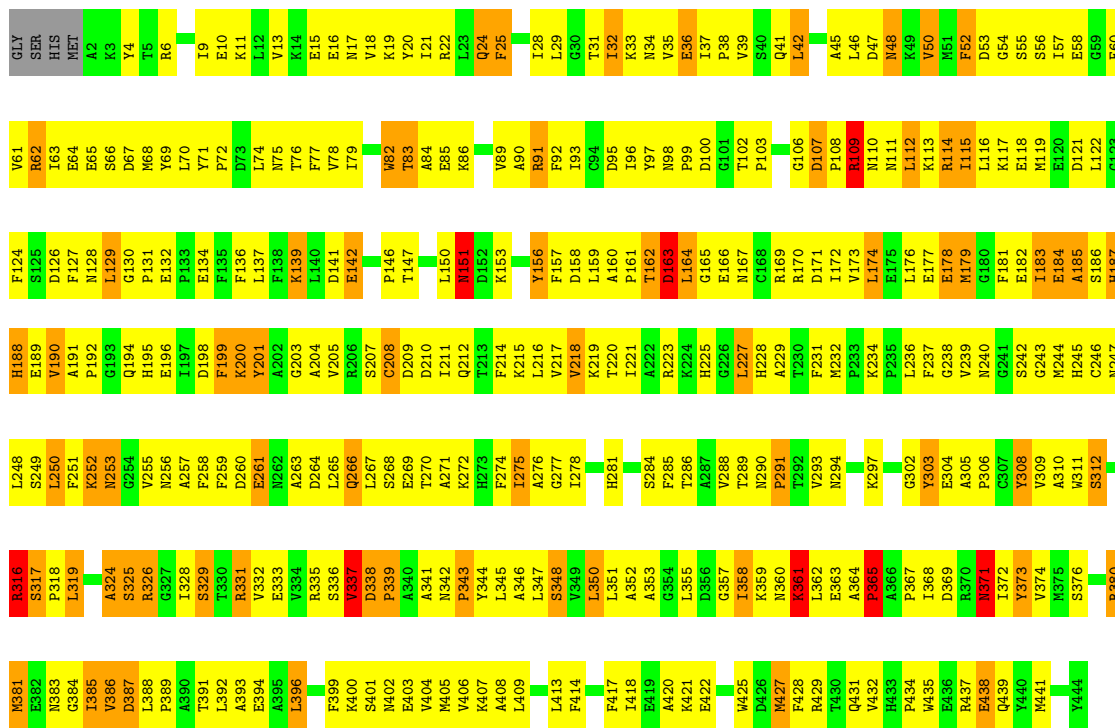
GLY	SER	HIS	MET	A2	K3	Y4	T5	R6	L70	E7	D8	I9	E10	K11	L12	V13	K14	E15	E16	M17	V18	F25	Y20	I21	R22	L23	Q24	R24	F26	T26	D27	L28	L29	N34	V35	E36	I37	P38	V39	S40	Q41	L42	G43	K44	A45	L46	D47	N48	K49	V50	H51	F52	D53	G54	G55	S56	L57	F60	V61							
R62	I63	E64	E65	S66	D67	M68	V69	L70	Y71	F72	D73	L74	N75	F76	F77	W78	I79	F80	F81	W82	T83	A84	I85	E86	K87	L88	R89	A90	R91	F92	I93	C94	D95	I96	Y97	B100	G101	T102	P103	F104	E105	G106	D107	P108	R109	N110	N111	L112	K113	F114	I115	G116	F117	S118	M119	E120	G123									
F124	S125	D126	F127	N128	L129	G130	P131	E132	P133	E134	F135	L136	L137	F138	K139	L140	D141	F80	E142	K143	G144	E145	P146	L147	L148	E149	L150	N151	G155	F156	F157	L158	L159	A160	P181	T182	D183	E186	N187	C188	R189	G190	L191	N192	C193	P233	R170	G234	P235	L236	F237	L174	E175	N111	L112	M179	I115	F181	E182	K117	L118	E184	A185	S186	H187	L250
H188	E189	V190	A191	G192	G193	Q194	H195	E196	I197	D198	F199	K200	A204	V205	R206	S207	C208	D209	D210	I211	Q212	T213	F214	K215	L216	V217	V218	A90	R219	T220	L221	A222	R223	K224	H225	G226	L227	T230	F231	N232	C233	P233	K234	P235	L236	F237	G238	V239	M240	G241	G242	G243	N244	N245	C246	N247	L248	S249	L250							
F251	K252	F258	D260	F261	A262	N263	Q266	L267	S268	E269	T270	A271	K272	R273	H274	I275	A281	H282	A283	T283	S284	F285	V286	L287	A341	N342	N343	Y344	L345	A346	L347	S348	V349	L351	K352	L299	V300	P301	G302	Y303	E304	A305	P306	C307	Y308	V309	K310	S311	S312	G313	A314	N315	K316	S317	F318	L319	I320									
R321	I322	P323	S325	D326	G327	S328	I329	T330	R331	V332	F333	V334	K335	S336	V337	D338	P339	A340	A341	N342	N343	Y344	L345	A346	L347	S348	V349	L351	K352	L299	V300	P301	G302	Y303	E304	A305	P306	C307	Y308	V309	K310	S311	S312	G313	A314	N315	K316	S317	F318	L319	I320															
E382	N383	G384	V385	V386	D387	L388	P389	A390	E391	A393	E394	A395	L396	E397	E398	F399	K400	S401	M402	E403	V404	N405	V406	L409	G410	E411	H412	L413	F414	E415	H416	F417	L418	H419	F420	L423	L424	N425	D426	M427	H428	N429	Y429	P430	H433	P434	H435	E436	R437	Y440	M441	S442	Q443	Y444												

• Molecule 1: Glutamine synthetase

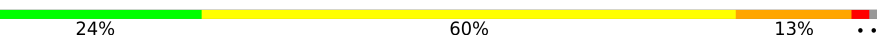
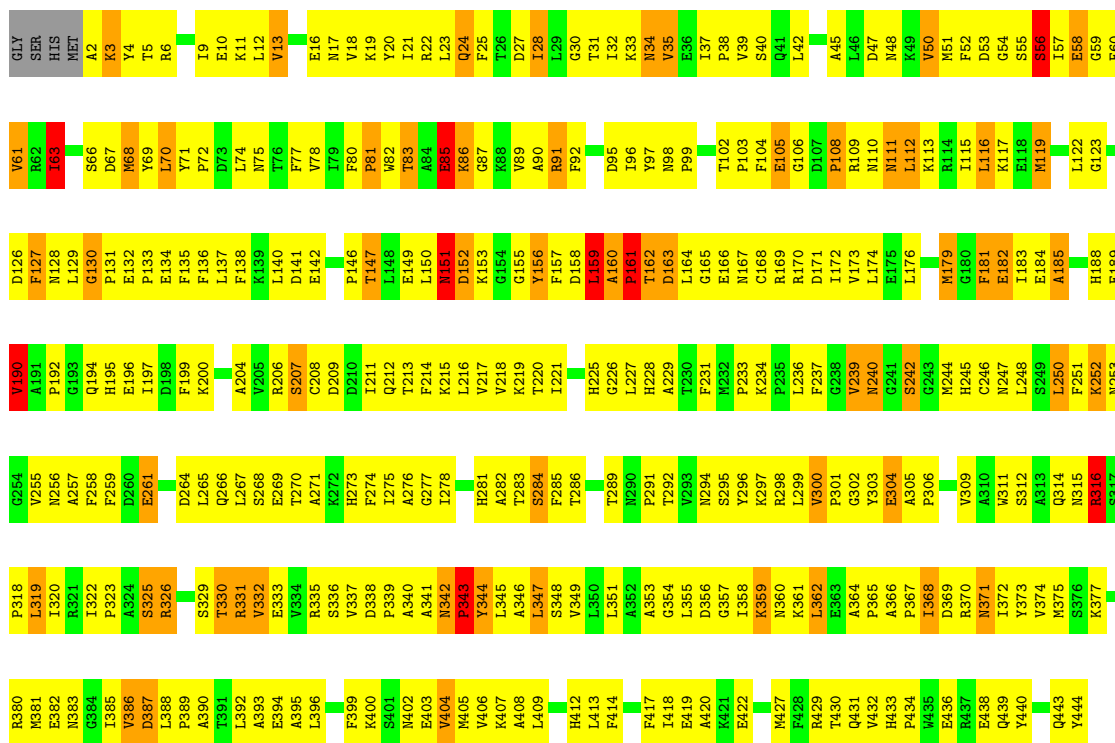
Chain G:  23% 59% 16% ..

GLY	SER	HIS	MET	A2	K3	Y4	T5	R6	L70	E7	D8	I9	E10	K11	L12	V13	K14	E15	E16	M17	V18	F25	Y20	I21	R22	L23	Q24	R24	F26	T26	D27	L28	L29	N34	V35	E36	I37	P38	V39	S40	Q41	L42	G43	K44	A45	L46	D47	N48	K49	V50	H51	F52	D53	G54	G55	S56	L57	F60	V61							
F60	V61	R62	I63	E64	S66	D67	M68	V69	L70	Y71	L74	F77	W78	I79	F80	F81	W82	T83	A84	I85	E86	K87	L88	R89	A90	R91	F92	I93	C94	D95	I96	Y97	N98	G101	T102	P103	F104	E105	G106	D107	P108	R109	N110	N111	L112	K113	F114	I115	G116	F117	S118	M119	E120	G123												
F124	S125	D126	F127	N128	L129	G130	P131	E132	P133	E134	F135	L136	L137	F138	K139	L140	D141	F80	E142	K143	G144	E145	P146	L147	L148	E149	L150	N151	G155	F156	F157	L158	L159	A160	P181	T182	D183	E186	N187	C188	R189	G190	L191	N192	C193	P233	R170	G234	P235	L236	F237	L174	E175	N111	L112	M179	I115	F181	E182	K117	L118	E184	A185	S186	H187	L250
E189	V190	A191	G192	G193	Q194	H195	E196	I197	D198	F199	K200	A202	V205	R206	S207	C208	D209	D210	I211	Q212	T213	F214	K215	L216	V217	V218	A90	R219	T220	L221	A222	R223	K224	H225	G226	L227	T230	F231	N232	C233	P233	K234	P235	L236	F237	G238	V239	M240	G241	G242	G243	N244	N245	C246	N247	L248	S249	L250	K252							
N253	G254	V256	L257	A258	F259	D260	E261	N262	F263	D264	L265	S268	A271	K272	R273	F274	I275	A276	G277	I278	V279	K280	H281	F285	V286	N289	P291	T292	N293	N294	S295	Y296	K297	R298	L299	V300	P301	G302	Y303	E304	A305	P306	C307	Y308	V309	K310	S311	S312	G313	A314	N315	K316	S317	F318	L319	I320										
L319	I320	R321	L322	P323	A324	S325	R326	G327	I328	R331	V332	E333	V334	R335	S336	V337	D338	P339	A340	A341	N342	P343	Y344	L345	A346	L347	S348	V349	L351	K352	L299	V300	P301	G302	Y303	E304	A305	P306	C307	Y308	V309	K310	S311	S312	G313	A314	N315	K316	S317	F318	L319	I320														
V386	D387	L388	P389	A390	T391	L392	A393	E394	A395	L396	K400	S401	N402	E403	V404	M405	V406	K407	E411	H412	L413	F414	E415	H416	F417	L418	E419	A420	E422	I423	E424	M425	D426	M427	R428	T429	A430	Q431	V432	M435	E436	R437	Y438	Q439	Y440	M441	S442	Q443	R444	S445	E446															

## ● Molecule 1: Glutamine synthetase

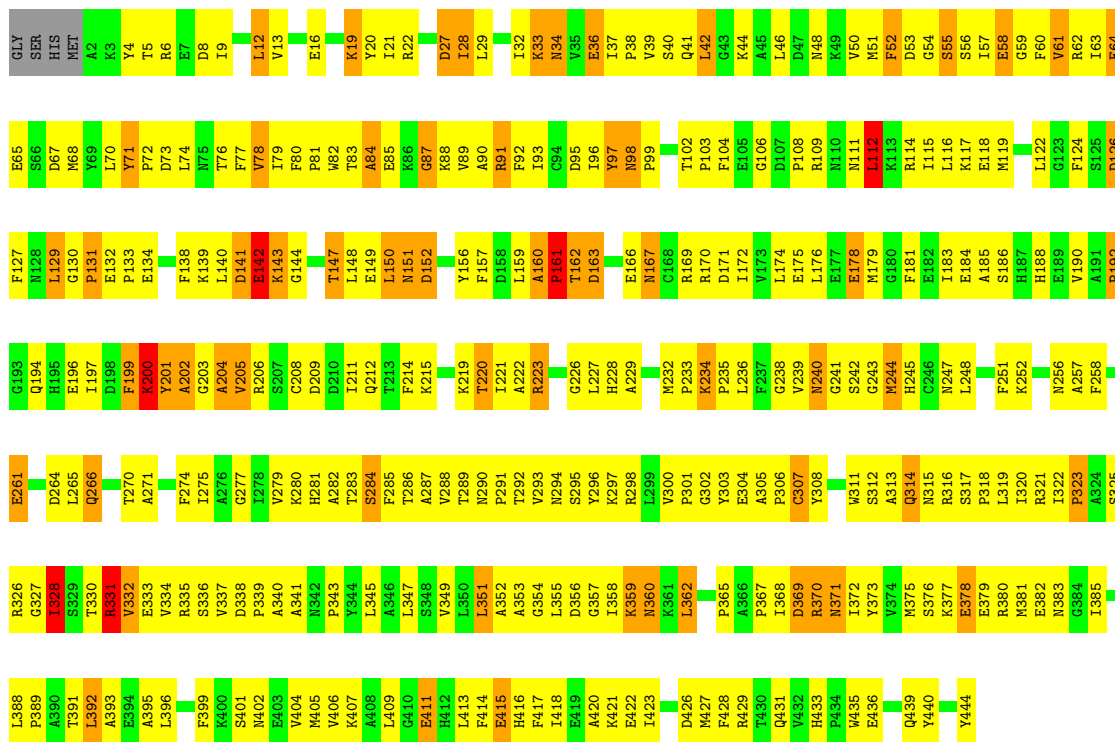
Chain H:  25% 57% 15% ..

## ● Molecule 1: Glutamine synthetase

Chain I:  24% 60% 13% ..

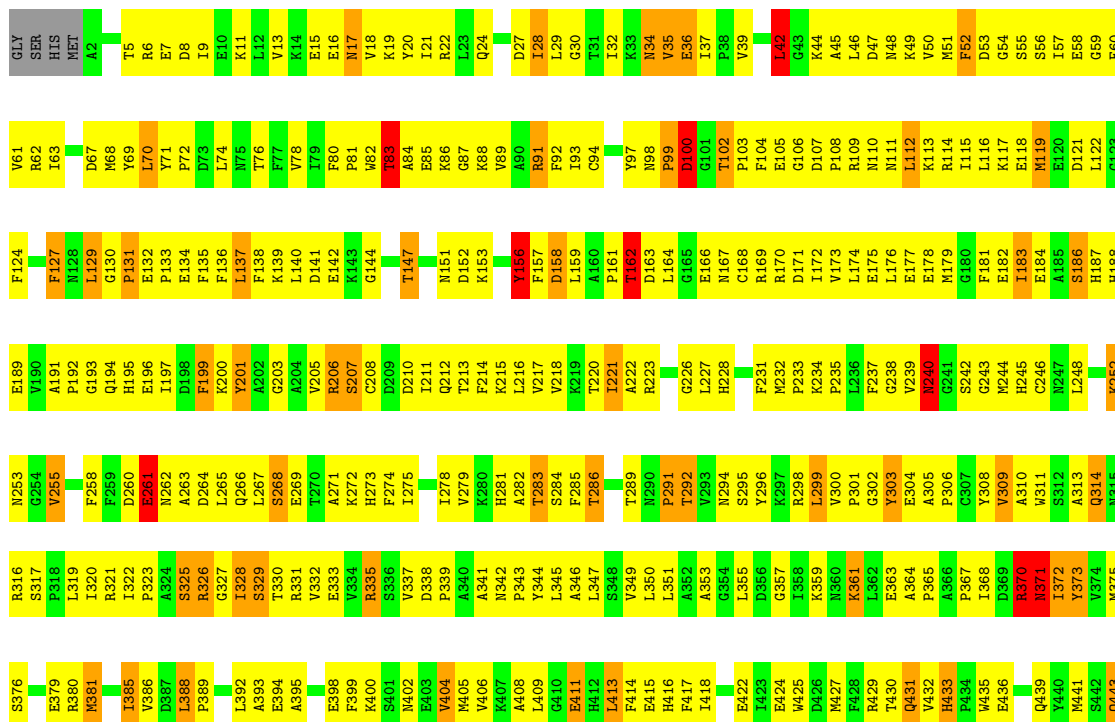
• Molecule 1: Glutamine synthetase

Chain J: 28% 56% 14%



• Molecule 1: Glutamine synthetase

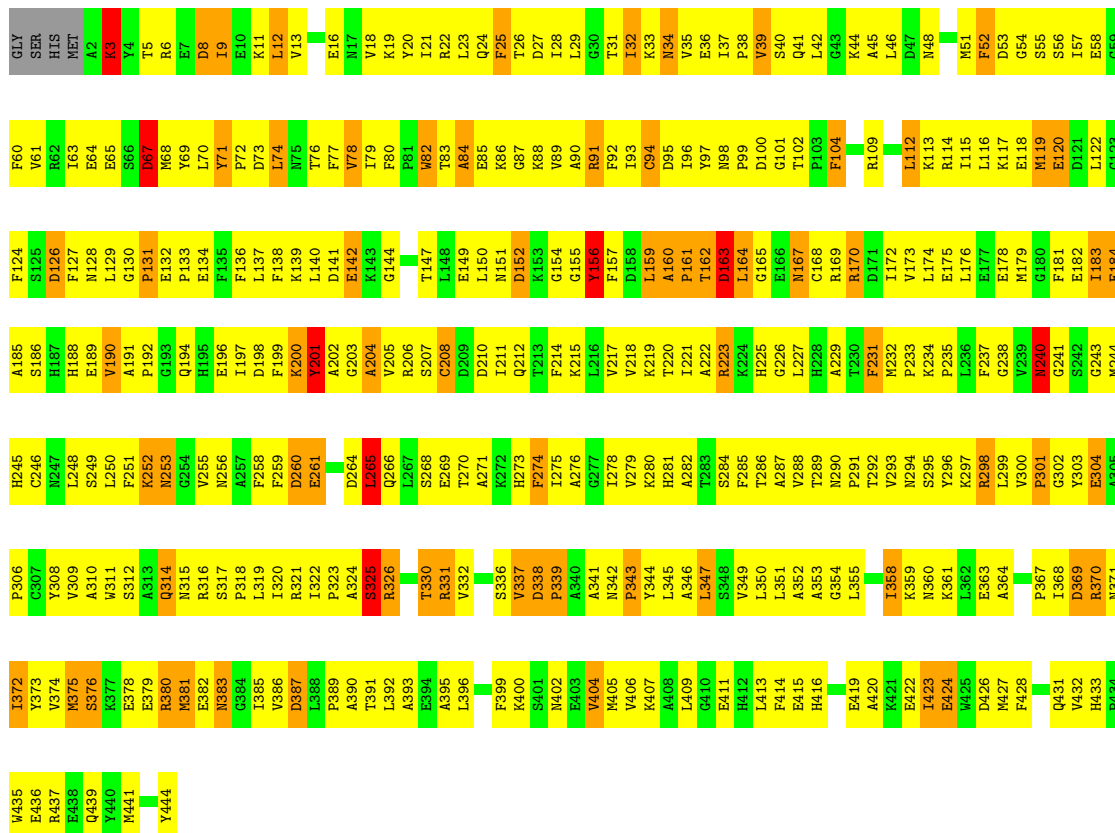
Chain K: 26% 59% 12%



Y444

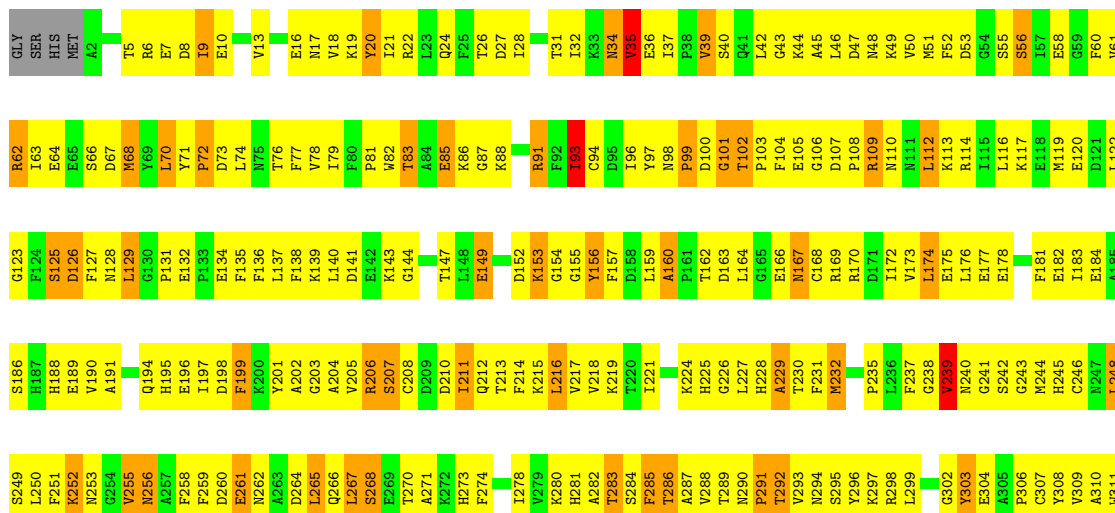
- Molecule 1: Glutamine synthetase

Chain L: 22% 60% 15%



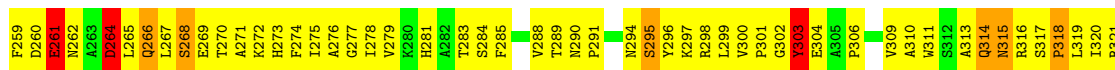
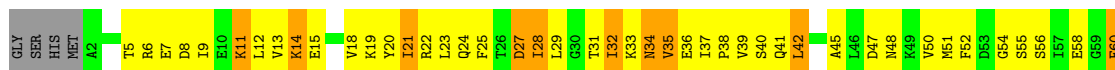
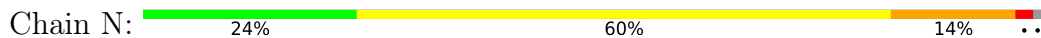
- Molecule 1: Glutamine synthetase

Chain M: 26% 58% 15%

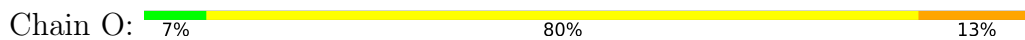




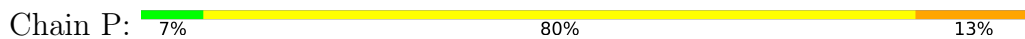
• Molecule 1: Glutamine synthetase



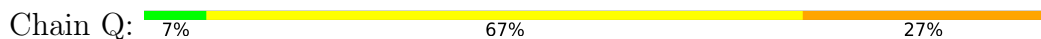
• Molecule 2: ThrA peptide



• Molecule 2: ThrA peptide



• Molecule 2: ThrA peptide







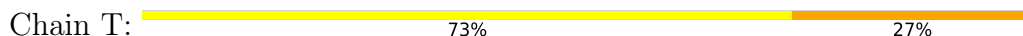
- Molecule 2: ThrA peptide



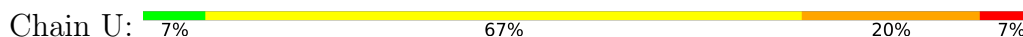
- Molecule 2: ThrA peptide



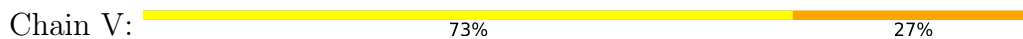
- Molecule 2: ThrA peptide



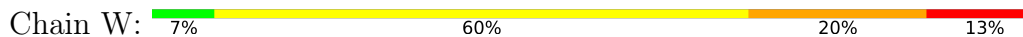
- Molecule 2: ThrA peptide



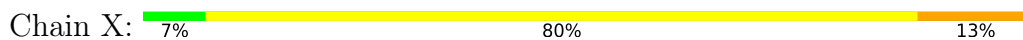
- Molecule 2: ThrA peptide



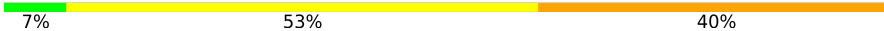
- Molecule 2: ThrA peptide



- Molecule 2: ThrA peptide

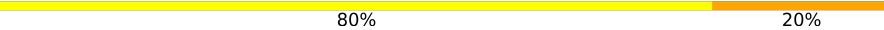


- Molecule 2: ThrA peptide

Chain Y:  7% 53% 40%



● Molecule 2: TnrA peptide

Chain Z:  80% 20%



● Molecule 2: TnrA peptide

Chain 1:  67% 27% 7%



● Molecule 2: TnrA peptide

Chain 2:  73% 27%



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	295.80Å 295.80Å 103.60Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	147.90 – 3.50 147.90 – 3.50	Depositor EDS
% Data completeness (in resolution range)	(Not available) (147.90-3.50) 95.8 (147.90-3.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.29 (at 3.49Å)	Xtrriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.243 , 0.284 0.258 , 0.289	Depositor DCC
$R_{free}$ test set	14819 reflections (12.09%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	57.1	Xtrriage
Anisotropy	0.227	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 18.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.40$ , $\langle L^2 \rangle = 0.22$	Xtrriage
Estimated twinning fraction	0.380 for -h,-k,l 0.387 for h,-h-k,-l 0.387 for -k,-h,-l	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	51555	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	57.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.20% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: MG

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	0.59	0/3618	0.97	9/4895 (0.2%)
1	B	0.58	0/3618	0.98	6/4895 (0.1%)
1	C	0.56	0/3618	0.94	4/4895 (0.1%)
1	D	0.59	0/3647	0.97	4/4933 (0.1%)
1	E	0.58	1/3618 (0.0%)	0.91	5/4895 (0.1%)
1	F	0.58	0/3618	0.91	3/4895 (0.1%)
1	G	0.56	0/3618	0.94	8/4895 (0.2%)
1	H	0.56	1/3618 (0.0%)	0.92	3/4895 (0.1%)
1	I	0.57	0/3618	0.95	6/4895 (0.1%)
1	J	0.55	0/3618	0.96	9/4895 (0.2%)
1	K	0.53	0/3618	0.91	5/4895 (0.1%)
1	L	0.57	1/3618 (0.0%)	0.94	5/4895 (0.1%)
1	M	0.60	1/3618 (0.0%)	0.99	15/4895 (0.3%)
1	N	0.59	0/3618	0.98	6/4895 (0.1%)
2	1	0.63	0/134	1.37	3/175 (1.7%)
2	2	0.67	0/135	1.07	1/175 (0.6%)
2	O	0.55	0/135	0.84	0/175
2	P	0.55	0/135	1.01	0/175
2	Q	0.72	0/135	1.09	0/175
2	R	0.68	0/135	1.27	1/175 (0.6%)
2	S	0.72	0/135	1.15	1/175 (0.6%)
2	T	0.60	0/135	1.12	1/175 (0.6%)
2	U	0.61	0/135	1.19	1/175 (0.6%)
2	V	1.06	1/135 (0.7%)	1.32	2/175 (1.1%)
2	W	0.81	0/135	1.39	3/175 (1.7%)
2	X	0.63	0/135	1.19	1/175 (0.6%)
2	Y	0.89	0/135	1.27	1/175 (0.6%)
2	Z	0.75	0/135	1.06	1/175 (0.6%)
All	All	0.58	5/52570 (0.0%)	0.96	104/71018 (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	3
1	B	0	1
1	D	0	1
1	F	0	1
1	J	0	1
All	All	0	7

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	408	ALA	CA-CB	8.77	1.70	1.52
2	V	760	ARG	C-O	8.06	1.38	1.23
1	H	208	CYS	CB-SG	-5.99	1.72	1.81
1	M	435	TRP	CB-CG	-5.13	1.41	1.50
1	L	201	TYR	CB-CG	-5.10	1.44	1.51

The worst 5 of 104 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	V	747	MET	N-CA-C	-10.28	83.24	111.00
1	J	331	ARG	NE-CZ-NH1	9.93	125.27	120.30
1	J	331	ARG	NE-CZ-NH2	-9.24	115.68	120.30
2	W	747	MET	N-CA-C	-7.96	89.52	111.00
2	1	747	MET	N-CA-C	-7.57	90.55	111.00

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	201	TYR	Sidechain
1	A	296	TYR	Sidechain
1	A	4	TYR	Sidechain
1	B	303	TYR	Sidechain
1	D	296	TYR	Sidechain

## 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	3535	0	3466	464	0
1	B	3535	0	3466	483	0
1	C	3535	0	3466	473	0
1	D	3563	0	3493	527	0
1	E	3535	0	3466	567	0
1	F	3535	0	3466	506	0
1	G	3535	0	3466	526	0
1	H	3535	0	3466	484	0
1	I	3535	0	3466	444	0
1	J	3535	0	3466	445	0
1	K	3535	0	3466	479	0
1	L	3535	0	3466	521	0
1	M	3535	0	3466	494	0
1	N	3535	0	3466	525	0
2	1	132	0	130	51	0
2	2	133	0	130	49	0
2	O	133	0	130	45	0
2	P	133	0	130	46	0
2	Q	133	0	130	49	0
2	R	133	0	130	36	0
2	S	133	0	130	55	0
2	T	133	0	130	56	0
2	U	133	0	130	46	0
2	V	133	0	130	47	0
2	W	133	0	130	73	0
2	X	133	0	130	41	0
2	Y	133	0	130	60	0
2	Z	133	0	130	65	0
3	A	9	0	7	5	0
3	B	9	0	7	6	0
3	C	10	0	7	3	0
3	D	9	0	7	3	0
3	E	9	0	7	2	0
3	F	9	0	7	5	0
3	G	9	0	7	2	0
3	H	9	0	7	1	0
3	I	9	0	7	2	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	J	9	0	7	3	0
3	K	9	0	7	0	0
3	L	10	0	7	4	0
3	M	10	0	7	0	0
3	N	9	0	7	1	0
4	A	4	0	0	0	0
4	B	3	0	0	0	0
4	C	7	0	0	0	0
4	D	8	0	0	0	0
4	E	2	0	0	0	0
4	F	3	0	0	0	0
4	G	1	0	0	0	0
4	H	3	0	0	0	0
4	I	2	0	0	0	0
4	J	3	0	0	0	0
4	K	3	0	0	0	0
4	L	3	0	0	0	0
4	M	3	0	0	0	0
4	N	2	0	0	0	0
All	All	51555	0	50469	6970	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 68.

The worst 5 of 6970 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:163:ASP:HA	1:A:170:ARG:NH2	1.36	1.38
2:W:758:LYS:HD2	2:W:758:LYS:O	1.23	1.26
1:C:113:LYS:HA	1:C:116:LEU:HD12	1.22	1.19
1:A:163:ASP:CA	1:A:170:ARG:HH22	1.56	1.18
1:M:329:SER:O	1:M:331:ARG:HD3	1.42	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	441/447 (99%)	330 (75%)	96 (22%)	15 (3%)	3	28
1	B	441/447 (99%)	338 (77%)	80 (18%)	23 (5%)	2	18
1	C	441/447 (99%)	325 (74%)	91 (21%)	25 (6%)	1	16
1	D	445/447 (100%)	333 (75%)	84 (19%)	28 (6%)	1	14
1	E	441/447 (99%)	331 (75%)	93 (21%)	17 (4%)	3	25
1	F	441/447 (99%)	335 (76%)	92 (21%)	14 (3%)	4	29
1	G	441/447 (99%)	327 (74%)	83 (19%)	31 (7%)	1	12
1	H	441/447 (99%)	335 (76%)	83 (19%)	23 (5%)	2	18
1	I	441/447 (99%)	333 (76%)	78 (18%)	30 (7%)	1	13
1	J	441/447 (99%)	338 (77%)	83 (19%)	20 (4%)	2	21
1	K	441/447 (99%)	339 (77%)	85 (19%)	17 (4%)	3	25
1	L	441/447 (99%)	340 (77%)	83 (19%)	18 (4%)	3	23
1	M	441/447 (99%)	336 (76%)	89 (20%)	16 (4%)	3	26
1	N	441/447 (99%)	317 (72%)	94 (21%)	30 (7%)	1	13
2	1	13/15 (87%)	12 (92%)	0	1 (8%)	1	10
2	2	13/15 (87%)	12 (92%)	1 (8%)	0	100	100
2	O	13/15 (87%)	12 (92%)	1 (8%)	0	100	100
2	P	13/15 (87%)	11 (85%)	1 (8%)	1 (8%)	1	10
2	Q	13/15 (87%)	10 (77%)	3 (23%)	0	100	100
2	R	13/15 (87%)	12 (92%)	1 (8%)	0	100	100
2	S	13/15 (87%)	10 (77%)	3 (23%)	0	100	100
2	T	13/15 (87%)	12 (92%)	1 (8%)	0	100	100
2	U	13/15 (87%)	8 (62%)	5 (38%)	0	100	100
2	V	13/15 (87%)	11 (85%)	2 (15%)	0	100	100
2	W	13/15 (87%)	11 (85%)	1 (8%)	1 (8%)	1	10

*Continued on next page...*



Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	X	13/15 (87%)	8 (62%)	5 (38%)	0	100	100
2	Y	13/15 (87%)	9 (69%)	4 (31%)	0	100	100
2	Z	13/15 (87%)	13 (100%)	0	0	100	100
All	All	6360/6468 (98%)	4808 (76%)	1242 (20%)	310 (5%)	2	19

5 of 310 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	161	PRO
1	A	371	ASN
1	B	86	LYS
1	B	161	PRO
1	B	162	THR

### 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	382/385 (99%)	322 (84%)	60 (16%)	2	15
1	B	382/385 (99%)	313 (82%)	69 (18%)	1	9
1	C	382/385 (99%)	313 (82%)	69 (18%)	1	9
1	D	385/385 (100%)	310 (80%)	75 (20%)	1	7
1	E	382/385 (99%)	300 (78%)	82 (22%)	1	5
1	F	382/385 (99%)	314 (82%)	68 (18%)	2	10
1	G	382/385 (99%)	312 (82%)	70 (18%)	1	8
1	H	382/385 (99%)	303 (79%)	79 (21%)	1	6
1	I	382/385 (99%)	314 (82%)	68 (18%)	2	10
1	J	382/385 (99%)	317 (83%)	65 (17%)	2	12
1	K	382/385 (99%)	316 (83%)	66 (17%)	2	11
1	L	382/385 (99%)	307 (80%)	75 (20%)	1	7
1	M	382/385 (99%)	313 (82%)	69 (18%)	1	9

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	N	382/385 (99%)	321 (84%)	61 (16%)	2 14
2	1	13/13 (100%)	11 (85%)	2 (15%)	2 16
2	2	13/13 (100%)	10 (77%)	3 (23%)	1 4
2	O	13/13 (100%)	11 (85%)	2 (15%)	2 16
2	P	13/13 (100%)	12 (92%)	1 (8%)	13 42
2	Q	13/13 (100%)	9 (69%)	4 (31%)	0 2
2	R	13/13 (100%)	8 (62%)	5 (38%)	0 1
2	S	13/13 (100%)	11 (85%)	2 (15%)	2 16
2	T	13/13 (100%)	10 (77%)	3 (23%)	1 4
2	U	13/13 (100%)	9 (69%)	4 (31%)	0 2
2	V	13/13 (100%)	11 (85%)	2 (15%)	2 16
2	W	13/13 (100%)	9 (69%)	4 (31%)	0 2
2	X	13/13 (100%)	12 (92%)	1 (8%)	13 42
2	Y	13/13 (100%)	8 (62%)	5 (38%)	0 1
2	Z	13/13 (100%)	11 (85%)	2 (15%)	2 16
All	All	5533/5572 (99%)	4517 (82%)	1016 (18%)	1 8

5 of 1016 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	237	PHE
1	M	270	THR
1	I	24	GLN
1	M	225	HIS
1	N	179	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 163 such sidechains are listed below:

Mol	Chain	Res	Type
1	L	247	ASN
2	Q	751	GLN
1	M	17	ASN
1	M	290	ASN
2	T	759	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 61 ligands modelled in this entry, 47 are monoatomic - leaving 14 for Mogul analysis.

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	$\# Z  > 2$
3	GLN	J	501	-	7,8,9	0.51	0	4,9,11	0.21	0
3	GLN	M	503	-	8,9,9	0.73	0	10,11,11	0.52	0
3	GLN	A	501	4	7,8,9	0.79	0	4,9,11	0.20	0
3	GLN	G	501	-	7,8,9	0.52	0	4,9,11	0.18	0
3	GLN	C	503	-	8,9,9	0.77	0	10,11,11	0.50	0
3	GLN	E	501	-	7,8,9	0.48	0	4,9,11	0.12	0
3	GLN	K	501	4	7,8,9	0.69	0	4,9,11	0.45	0
3	GLN	D	501	-	7,8,9	0.51	0	4,9,11	0.07	0
3	GLN	B	501	4	7,8,9	0.67	0	4,9,11	0.14	0
3	GLN	L	501	4	8,9,9	0.74	0	10,11,11	0.64	0
3	GLN	I	501	4	7,8,9	0.48	0	4,9,11	0.12	0
3	GLN	F	501	-	7,8,9	0.73	0	4,9,11	0.26	0
3	GLN	H	501	4	7,8,9	0.52	0	4,9,11	0.08	0
3	GLN	N	501	4	7,8,9	0.45	0	4,9,11	0.18	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
3	GLN	J	501	-	-	1/6/7/9	-
3	GLN	M	503	-	-	4/9/9/9	-
3	GLN	A	501	4	-	1/6/7/9	-
3	GLN	G	501	-	-	0/6/7/9	-
3	GLN	C	503	-	-	8/9/9/9	-
3	GLN	E	501	-	-	0/6/7/9	-
3	GLN	K	501	4	-	0/6/7/9	-
3	GLN	D	501	-	-	2/6/7/9	-
3	GLN	B	501	4	-	3/6/7/9	-
3	GLN	L	501	4	-	4/9/9/9	-
3	GLN	I	501	4	-	3/6/7/9	-
3	GLN	F	501	-	-	2/6/7/9	-
3	GLN	H	501	4	-	2/6/7/9	-
3	GLN	N	501	4	-	2/6/7/9	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 32 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	503	GLN	O-C-CA-N
3	C	503	GLN	N-CA-CB-CG
3	C	503	GLN	C-CA-CB-CG
3	D	501	GLN	N-CA-CB-CG
3	D	501	GLN	C-CA-CB-CG

There are no ring outliers.

12 monomers are involved in 37 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	J	501	GLN	3	0
3	A	501	GLN	5	0
3	G	501	GLN	2	0
3	C	503	GLN	3	0
3	E	501	GLN	2	0
3	D	501	GLN	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	501	GLN	6	0
3	L	501	GLN	4	0
3	I	501	GLN	2	0
3	F	501	GLN	5	0
3	H	501	GLN	1	0
3	N	501	GLN	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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.