wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 10:35 PM GMT

PDB ID	:	4XYC
Title	:	NANOMOLAR INHIBITORS OF MYCOBACTERIUM TUBERCULOSIS GLUTAMINE SYNTHETASE 1: SYNTHESIS, BIOLOGICAL EVALUATION AND X-RAY CRYSTALLOGRAPHIC STUDIES
Authors	:	Couturier, C.; Silve, S.; Morales, R.; Ppessegue, B.; Llopart, S.; Nair, A.; Bauer, A.; Scheiper, B.; poeverlein, c.; Ganzhorn, A.; Lagrange, S.; Bacque, E.
Deposited on	:	2015-02-02
Resolution	:	3.30 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

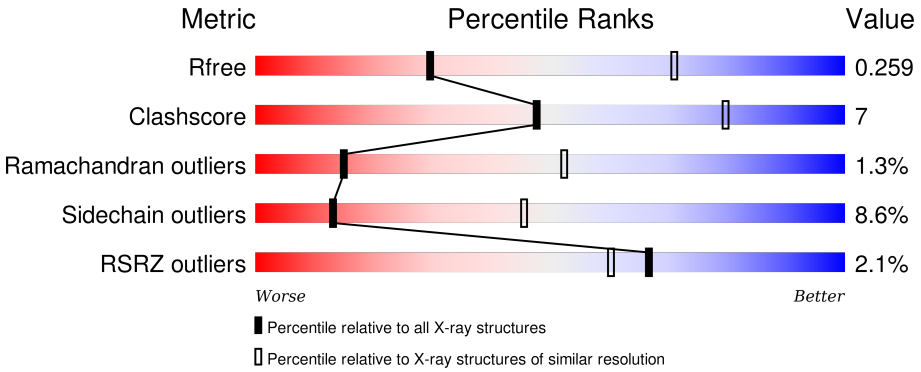
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 3.30 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	2060 (3.40-3.20)
Clashscore	102246	1058 (3.38-3.22)
Ramachandran outliers	100387	1038 (3.38-3.22)
Sidechain outliers	100360	1037 (3.38-3.22)
RSRZ outliers	91569	2070 (3.40-3.20)

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	478	4% 78% 18%
1	B	478	3% 70% 26%
1	C	478	% 73% 19%
1	D	478	2% 76% 20%
1	E	478	% 76% 19%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	478	71% 23%
1	G	478	3% 71% 24%
1	H	478	3% 73% 22%
1	I	478	76% 19%
1	J	478	78% 18%
1	K	478	77% 17%
1	L	478	73% 22%
1	M	478	3% 77% 17%
1	N	478	74% 20%
1	O	478	73% 21%
1	P	478	74% 21%
1	Q	478	71% 23%
1	R	478	78% 16%
1	S	478	77% 19%
1	T	478	76% 19%
1	U	478	76% 19%
1	V	478	79% 17%
1	W	478	75% 20%
1	X	478	75% 20%

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	2K9	O	900	-	-	-	X
2	2K9	U	900	-	-	-	X

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 88848 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 1.

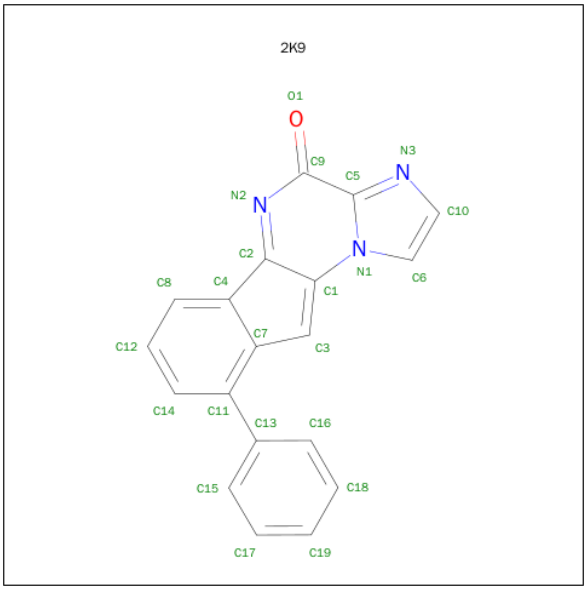
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	466	Total	C	N	O	S	0	0	0
			3693	2354	618	709	12			
1	B	464	Total	C	N	O	S	0	0	0
			3674	2342	613	707	12			
1	C	460	Total	C	N	O	S	0	0	0
			3645	2325	609	699	12			
1	D	466	Total	C	N	O	S	0	0	0
			3690	2352	616	710	12			
1	E	463	Total	C	N	O	S	0	0	0
			3670	2341	613	704	12			
1	F	467	Total	C	N	O	S	0	0	0
			3702	2359	619	712	12			
1	G	466	Total	C	N	O	S	0	0	0
			3693	2354	618	709	12			
1	H	464	Total	C	N	O	S	0	0	0
			3674	2342	613	707	12			
1	I	460	Total	C	N	O	S	0	0	0
			3645	2325	609	699	12			
1	J	466	Total	C	N	O	S	0	0	0
			3690	2352	616	710	12			
1	K	463	Total	C	N	O	S	0	0	0
			3670	2341	613	704	12			
1	L	467	Total	C	N	O	S	0	0	0
			3702	2359	619	712	12			
1	M	466	Total	C	N	O	S	0	0	0
			3693	2354	618	709	12			
1	N	464	Total	C	N	O	S	0	0	0
			3674	2342	613	707	12			
1	O	460	Total	C	N	O	S	0	0	0
			3645	2325	609	699	12			
1	P	466	Total	C	N	O	S	0	0	0
			3690	2352	616	710	12			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Q	463	Total	C	N	O	S	0	0	0
			3670	2341	613	704	12			
1	R	467	Total	C	N	O	S	0	0	0
			3702	2359	619	712	12			
1	S	466	Total	C	N	O	S	0	0	0
			3693	2354	618	709	12			
1	T	464	Total	C	N	O	S	0	0	0
			3674	2342	613	707	12			
1	U	460	Total	C	N	O	S	0	0	0
			3645	2325	609	699	12			
1	V	466	Total	C	N	O	S	0	0	0
			3690	2352	616	710	12			
1	W	463	Total	C	N	O	S	0	0	0
			3670	2341	613	704	12			
1	X	467	Total	C	N	O	S	0	0	0
			3702	2359	619	712	12			

- Molecule 2 is 9-phenyl-4H-imidazo[1,2-a]indeno[1,2-e]pyrazin-4-one (three-letter code: 2K9) (formula: $C_{19}H_{11}N_3O$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			23	19	3	1		
2	B	1	Total	C	N	O	0	0
			23	19	3	1		
2	C	1	Total	C	N	O	0	0
			23	19	3	1		

Continued on next page...

Continued from previous page...

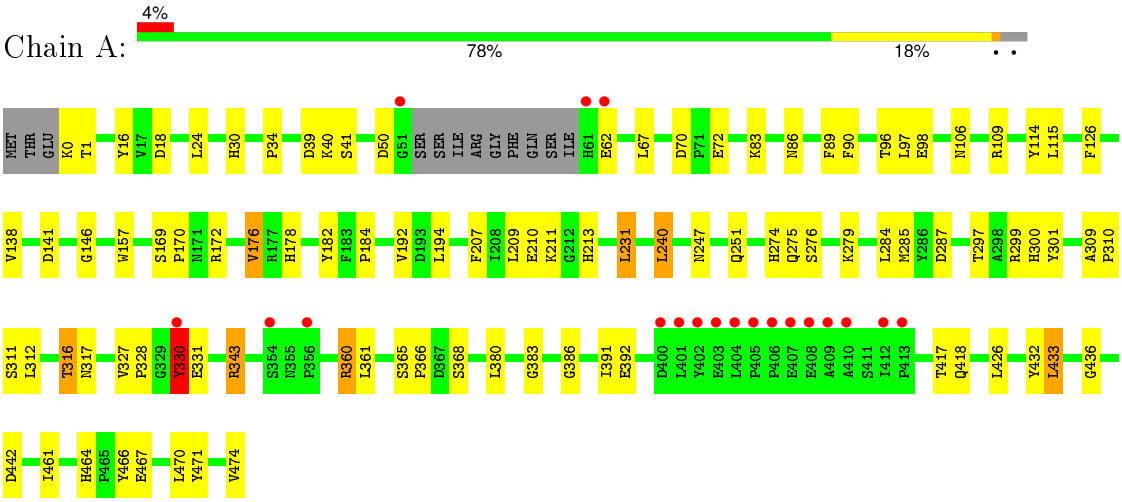
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	D	1	Total	C	N	O	0	0
			23	19	3	1		
2	E	1	Total	C	N	O	0	0
			23	19	3	1		
2	F	1	Total	C	N	O	0	0
			23	19	3	1		
2	G	1	Total	C	N	O	0	0
			23	19	3	1		
2	H	1	Total	C	N	O	0	0
			23	19	3	1		
2	I	1	Total	C	N	O	0	0
			23	19	3	1		
2	J	1	Total	C	N	O	0	0
			23	19	3	1		
2	K	1	Total	C	N	O	0	0
			23	19	3	1		
2	L	1	Total	C	N	O	0	0
			23	19	3	1		
2	M	1	Total	C	N	O	0	0
			23	19	3	1		
2	N	1	Total	C	N	O	0	0
			23	19	3	1		
2	O	1	Total	C	N	O	0	0
			23	19	3	1		
2	P	1	Total	C	N	O	0	0
			23	19	3	1		
2	Q	1	Total	C	N	O	0	0
			23	19	3	1		
2	R	1	Total	C	N	O	0	0
			23	19	3	1		
2	S	1	Total	C	N	O	0	0
			23	19	3	1		
2	T	1	Total	C	N	O	0	0
			23	19	3	1		
2	U	1	Total	C	N	O	0	0
			23	19	3	1		
2	V	1	Total	C	N	O	0	0
			23	19	3	1		
2	W	1	Total	C	N	O	0	0
			23	19	3	1		
2	X	1	Total	C	N	O	0	0
			23	19	3	1		

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 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.

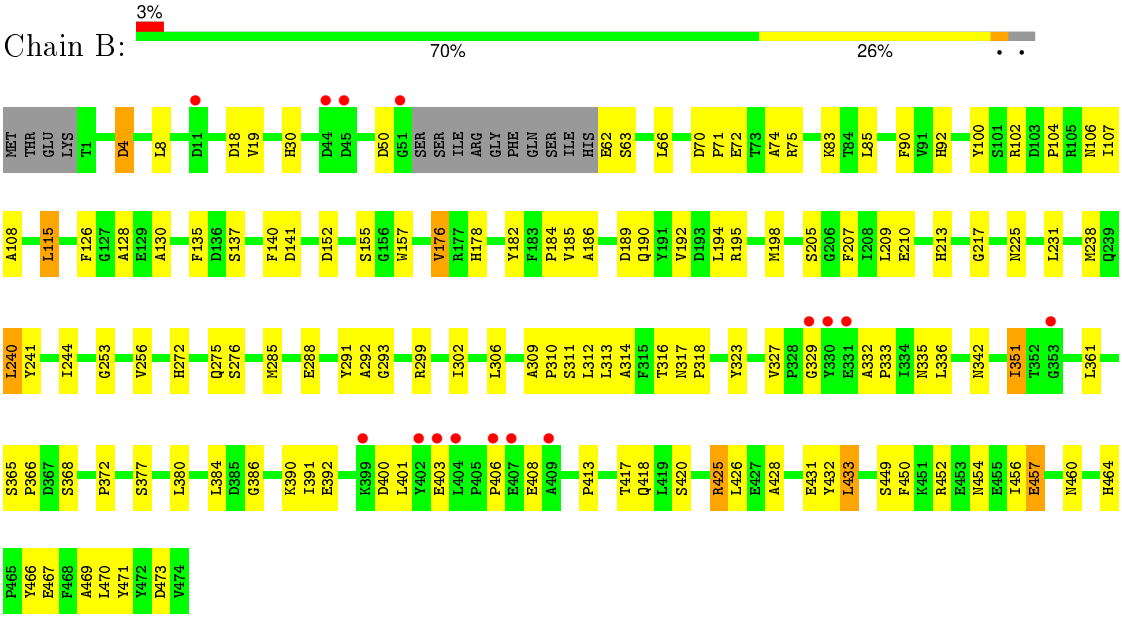
- Molecule 1: Glutamine synthetase 1

Chain A:



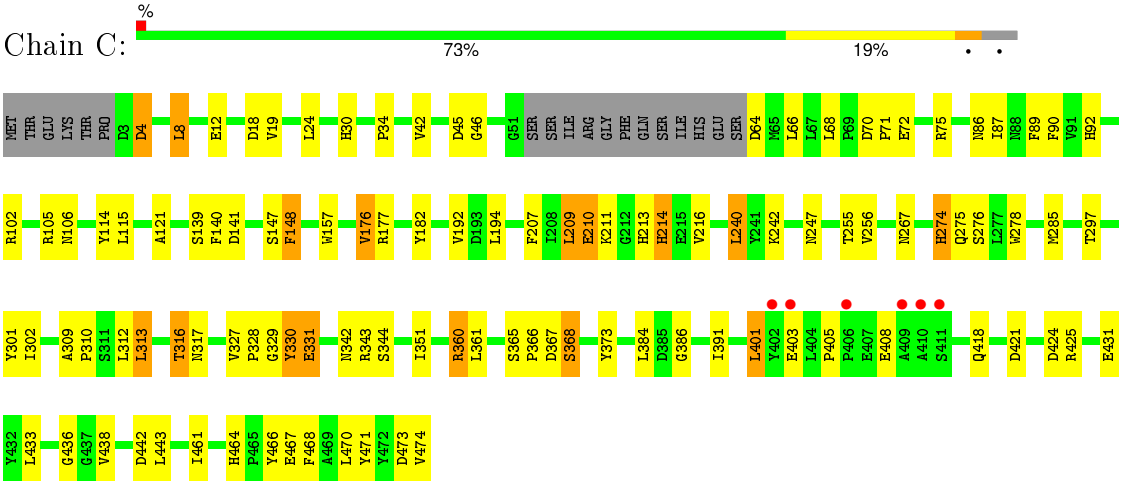
- Molecule 1: Glutamine synthetase 1

Chain B:



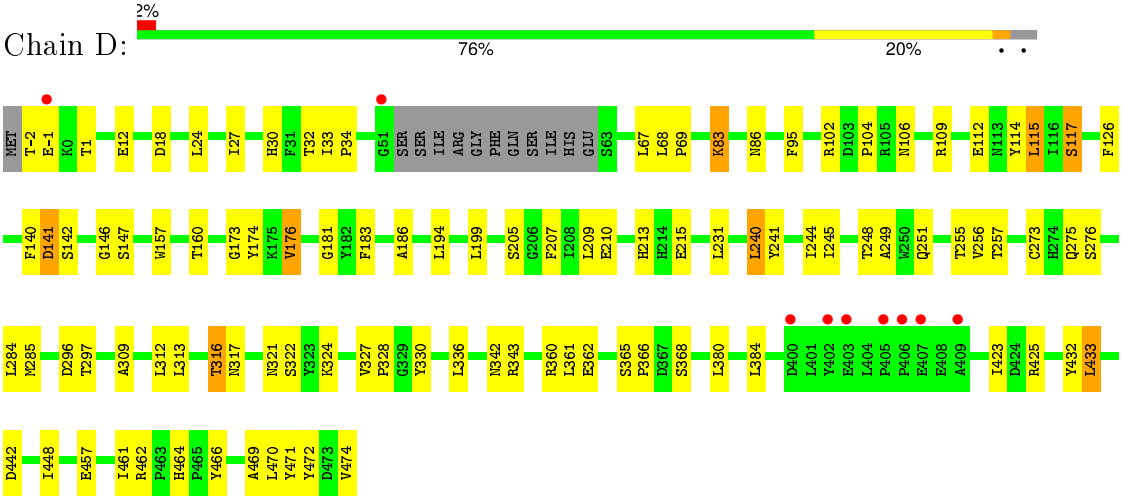
- Molecule 1: Glutamine synthetase 1

Chain C:



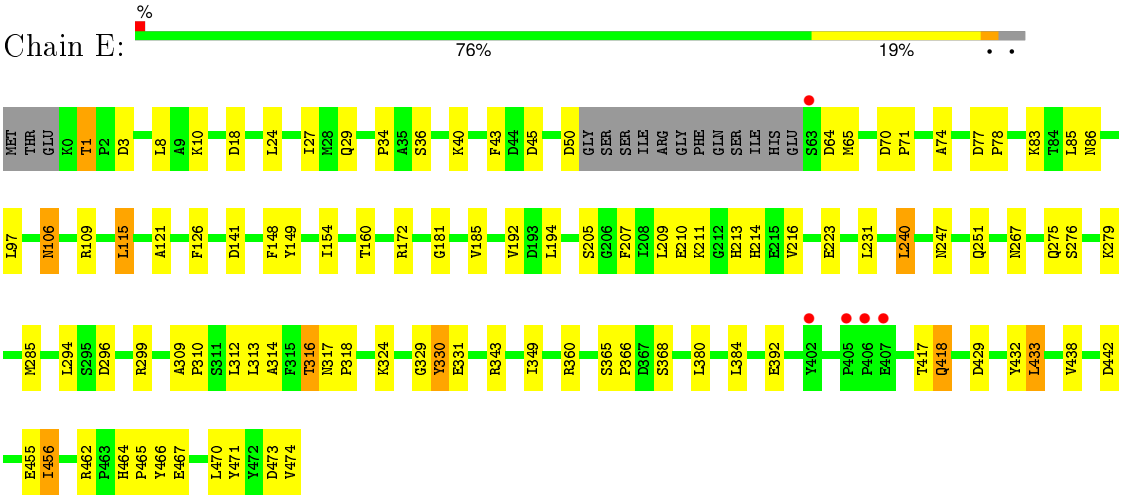
• Molecule 1: Glutamine synthetase 1

Chain D:



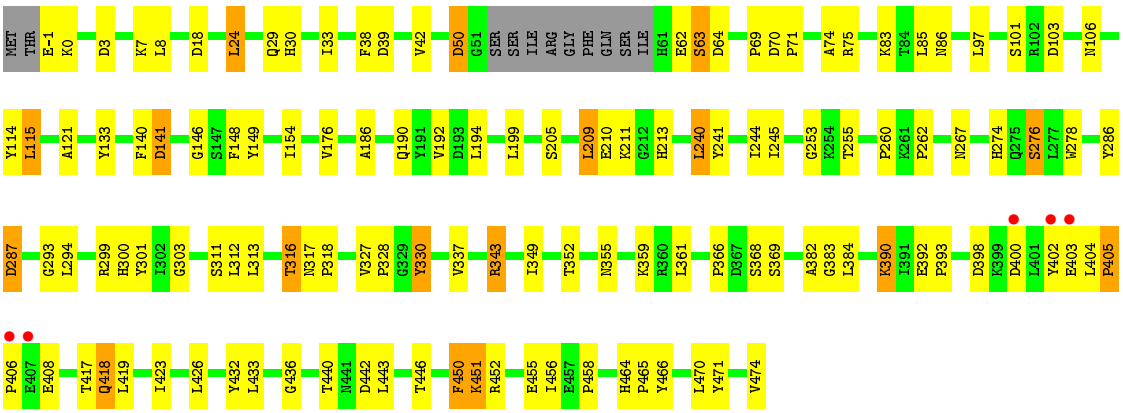
• Molecule 1: Glutamine synthetase 1

Chain E:



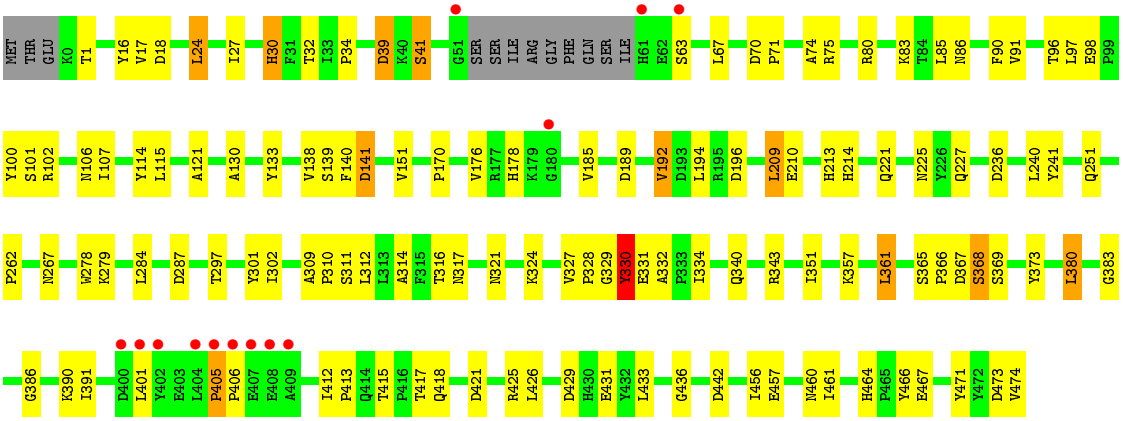
• Molecule 1: Glutamine synthetase 1

Chain F:



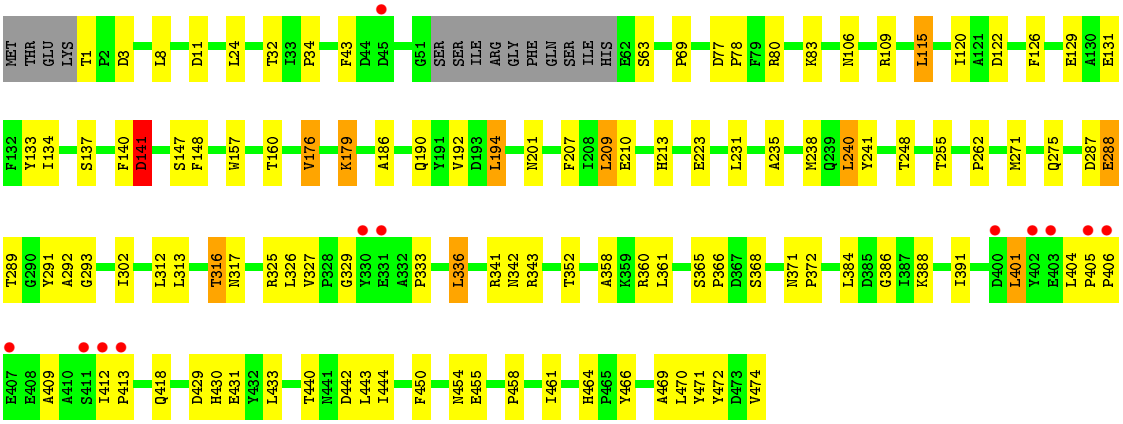
• Molecule 1: Glutamine synthetase 1

Chain G:



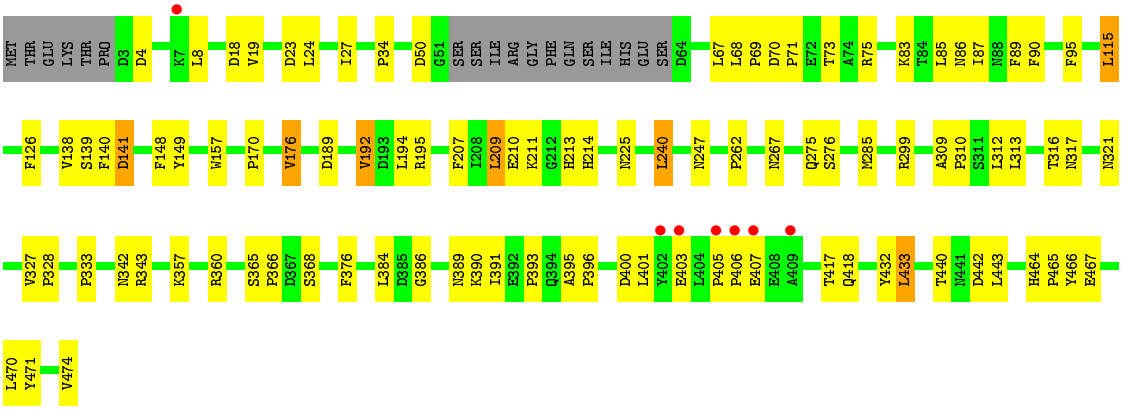
• Molecule 1: Glutamine synthetase 1

Chain H:



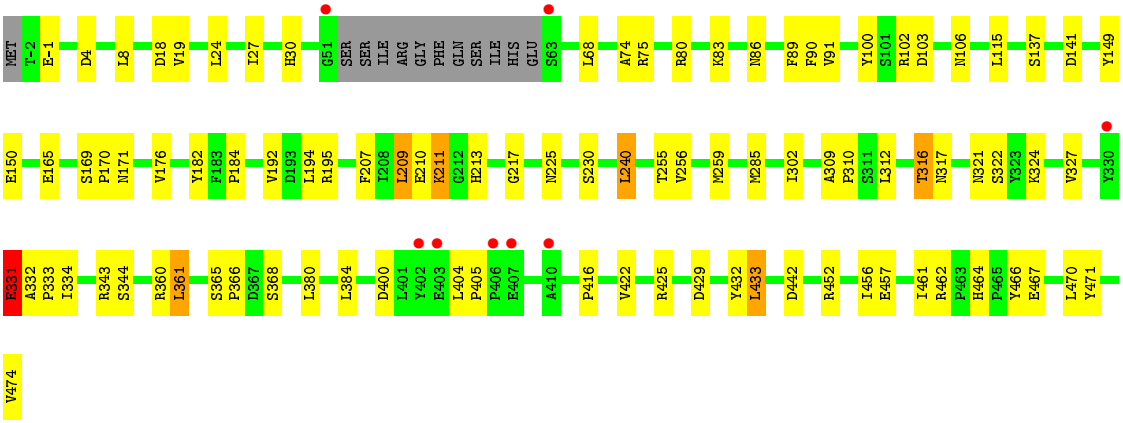
• Molecule 1: Glutamine synthetase 1

Chain I:



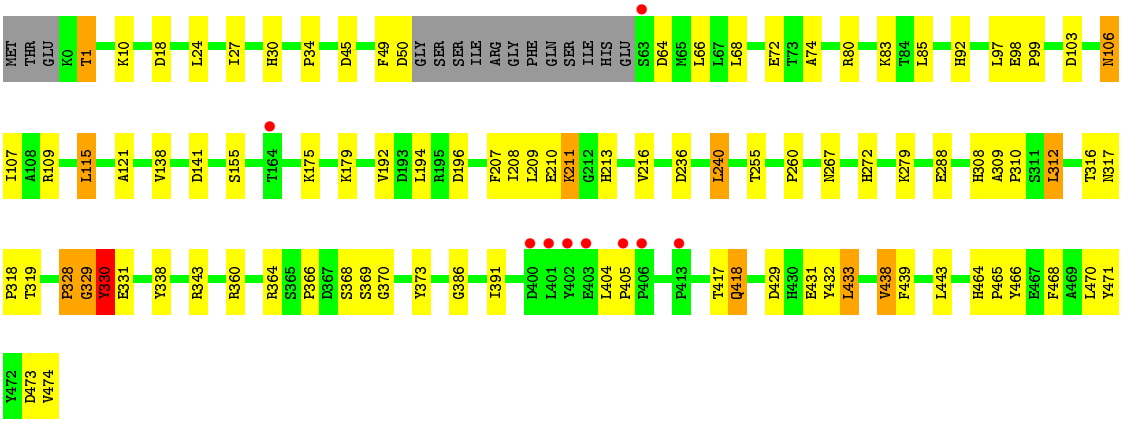
• Molecule 1: Glutamine synthetase 1

Chain J:



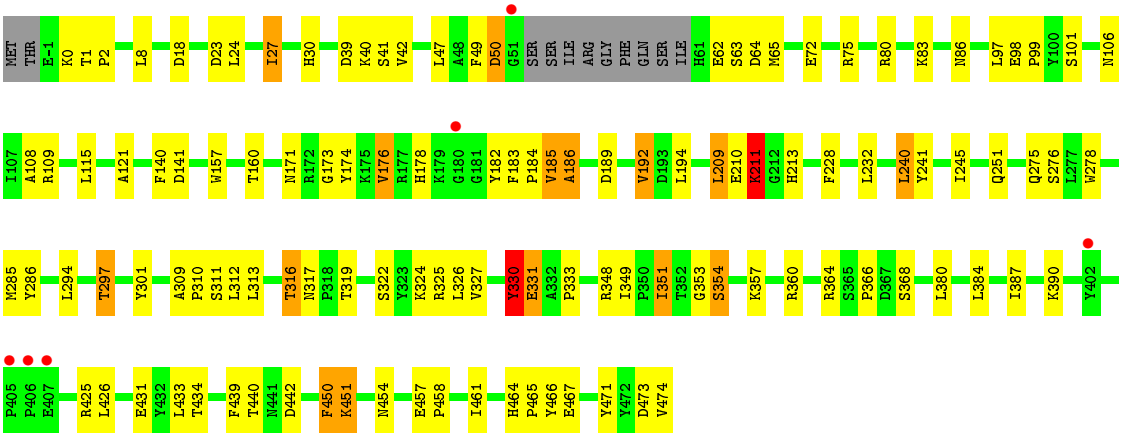
• Molecule 1: Glutamine synthetase 1

Chain K:



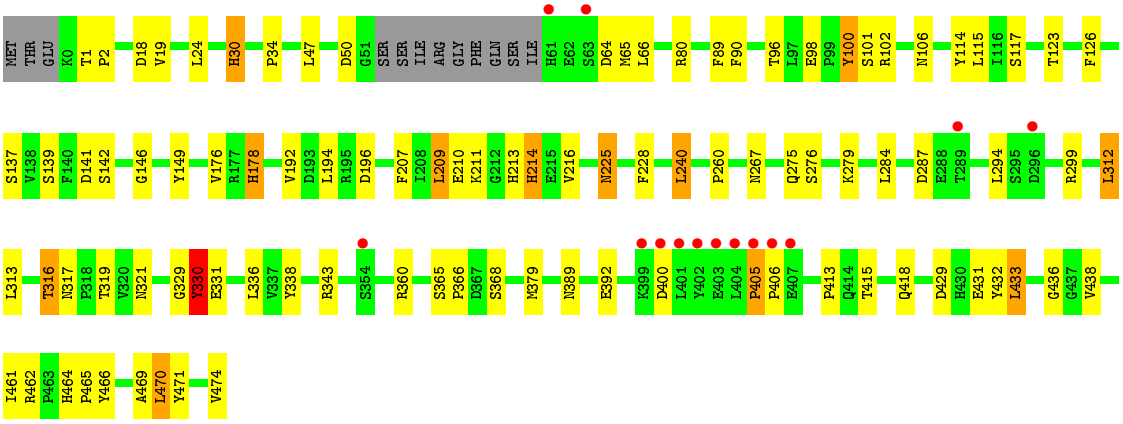
• Molecule 1: Glutamine synthetase 1

Chain L:



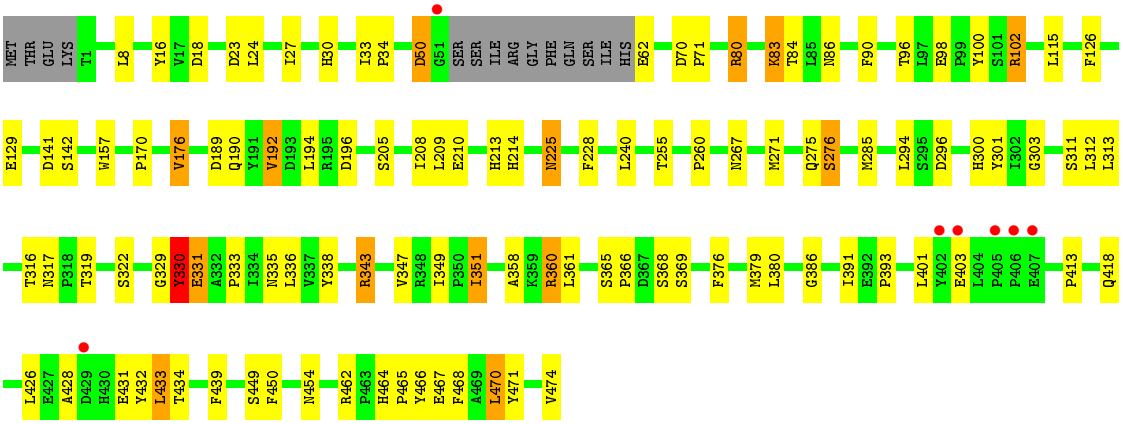
• Molecule 1: Glutamine synthetase 1

Chain M:



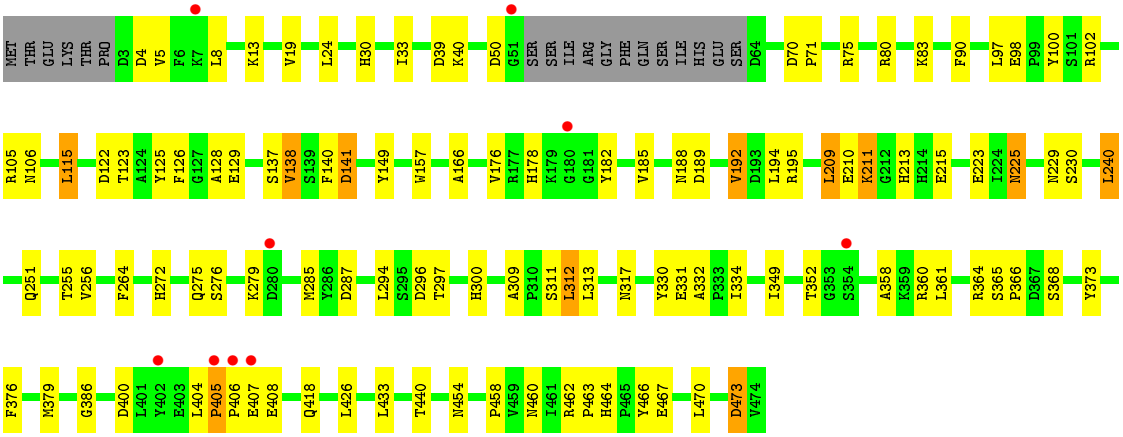
• Molecule 1: Glutamine synthetase 1

Chain N:



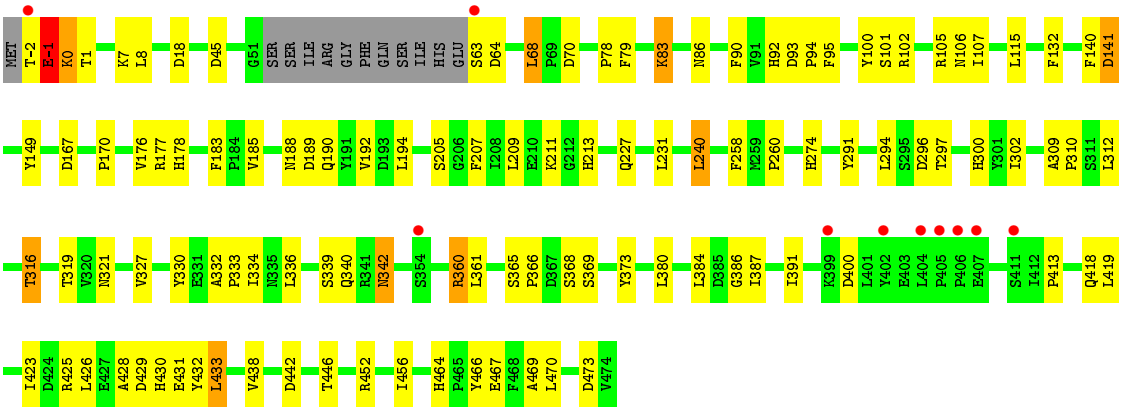
• Molecule 1: Glutamine synthetase 1

Chain O:



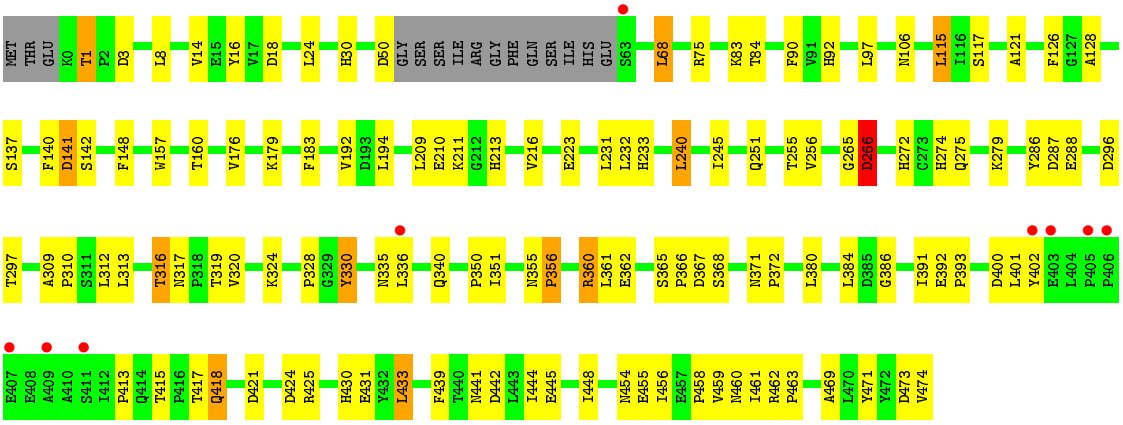
• Molecule 1: Glutamine synthetase 1

Chain P:



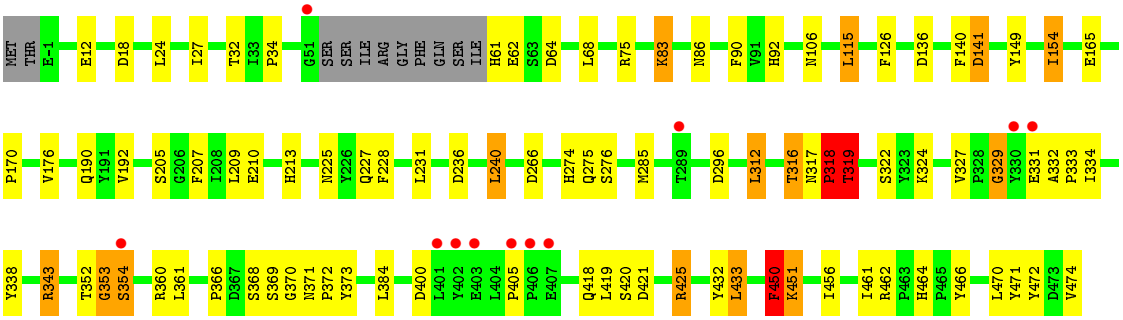
• Molecule 1: Glutamine synthetase 1

Chain Q:



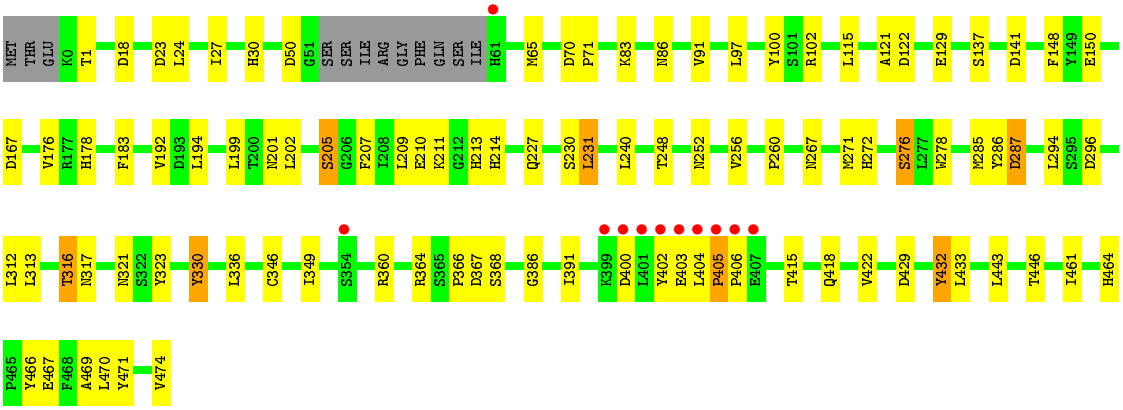
• Molecule 1: Glutamine synthetase 1

Chain R:



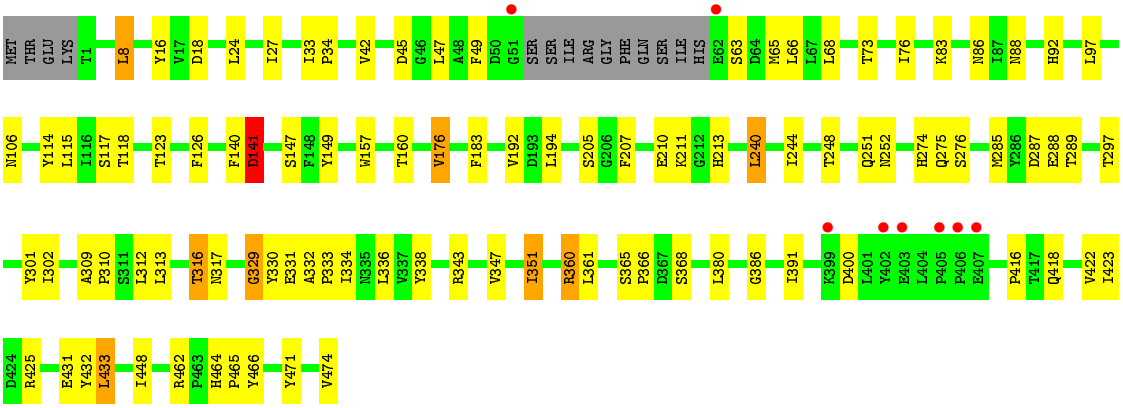
• Molecule 1: Glutamine synthetase 1

Chain S:



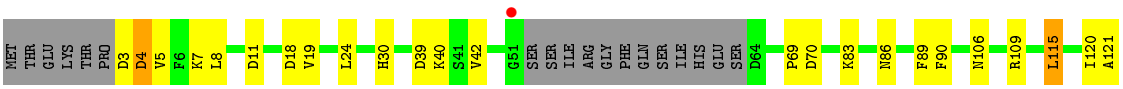
• Molecule 1: Glutamine synthetase 1

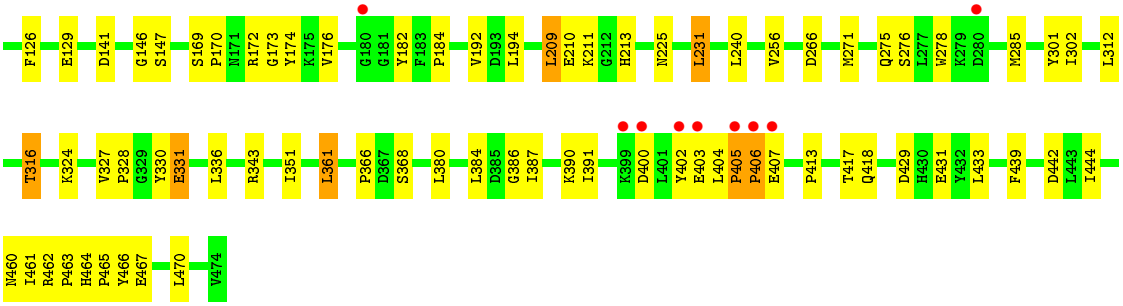
Chain T:



• Molecule 1: Glutamine synthetase 1

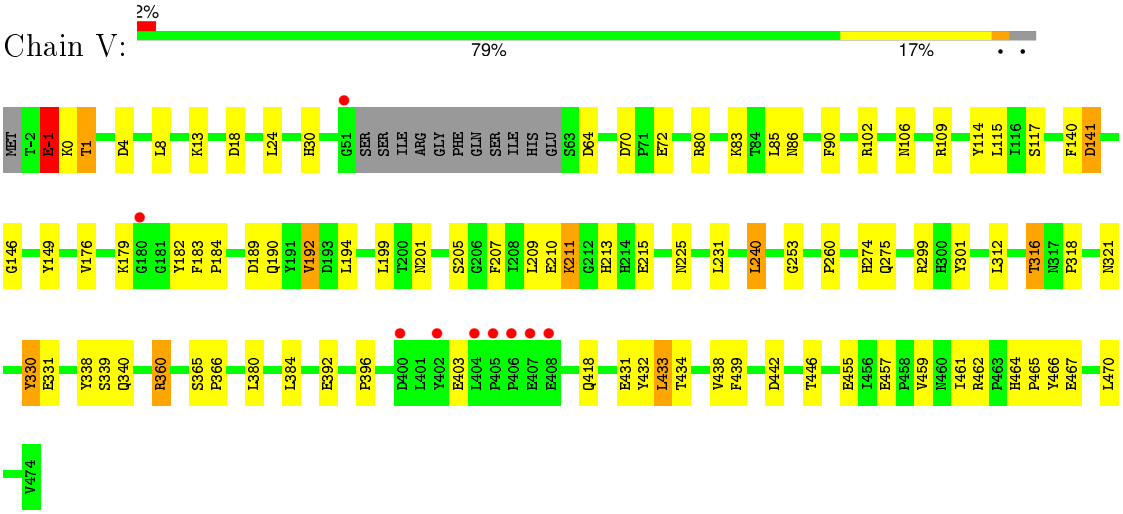
Chain U:





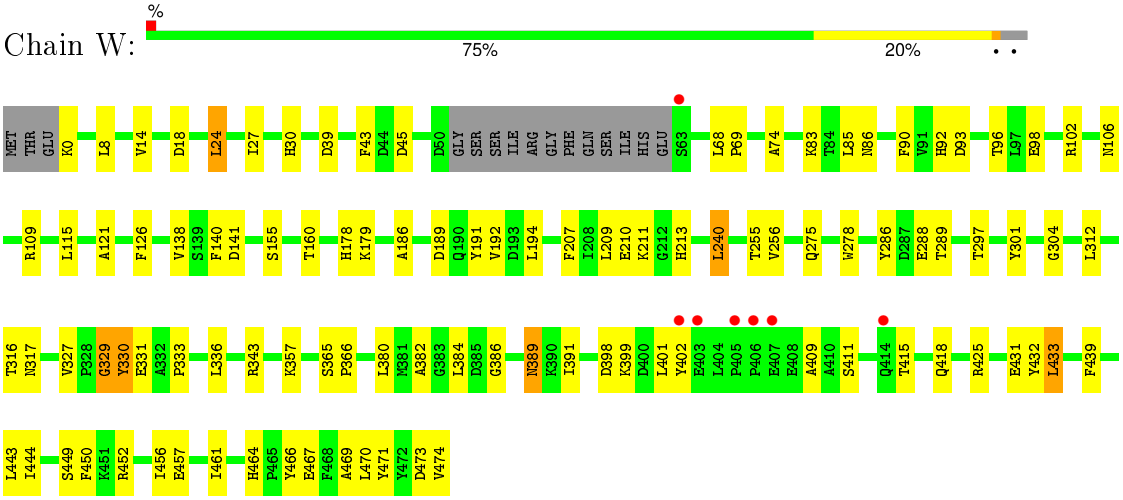
• Molecule 1: Glutamine synthetase 1

Chain V:



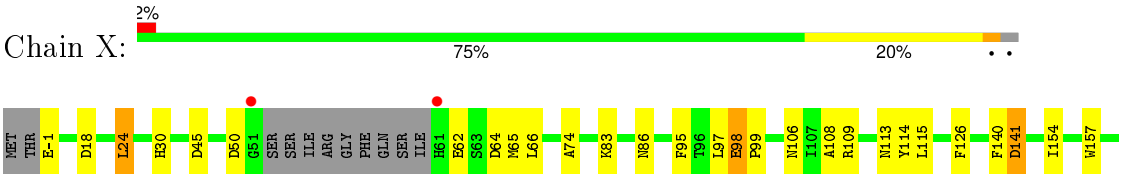
• Molecule 1: Glutamine synthetase 1

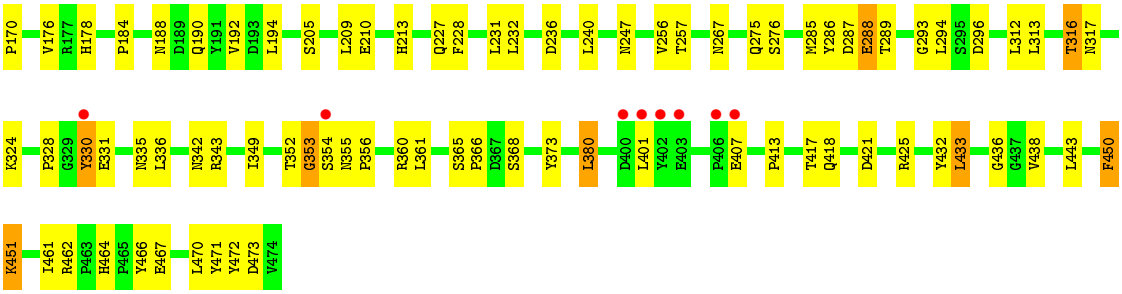
Chain W:



• Molecule 1: Glutamine synthetase 1

Chain X:




P170 V176 R177 H178 P184 N188 D189 Q190 Y191 V192 D193 L194 S205 L209 E210 H213 Q227 F228 L231 L232 D236 L240 N247 V256 T257 N267 Q275 S276 M285 Y286 D287 E288 T289 G293 L294 S295 D296 L312 L313 T316 N317
K324 P328 G329 Y330 E331 N335 L336 N342 R343 I349 T352 G353 S354 N355 P356 R360 L361 S365 P366 D367 S368 Y373 L380 D400 L401 Y402 E403 P406 E407 P413 T417 Q418 D421 R425 Y432 L433 G436 G437 V438 L443 F450
K451 I461 R462 P463 H464 P465 Y466 E467 L470 Y471 Y472 D473 V474

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	260.11Å 273.50Å 207.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.30 49.62 – 3.30	Depositor EDS
% Data completeness (in resolution range)	99.8 (50.00-3.30) 99.9 (49.62-3.30)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.80 (at 3.33Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.200 , 0.265 0.197 , 0.259	Depositor DCC
R_{free} test set	11105 reflections (5.28%)	DCC
Wilson B-factor (Å^2)	62.5	Xtriage
Anisotropy	0.025	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 23.2	EDS
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Outliers	3 of 221317 reflections (0.001%)	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	88848	wwPDB-VP
Average B, all atoms (Å^2)	57.0	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 34.86 % 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 6.3847e-04.The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 2K9

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	0.48	0/3797	0.59	0/5161
1	B	0.59	0/3777	0.62	0/5135
1	C	0.50	0/3747	0.62	0/5093
1	D	0.47	0/3793	0.61	0/5156
1	E	0.47	0/3773	0.62	0/5129
1	F	0.52	0/3806	0.63	1/5173 (0.0%)
1	G	0.49	0/3797	0.61	1/5161 (0.0%)
1	H	0.52	0/3777	0.63	0/5135
1	I	0.49	0/3747	0.61	1/5093 (0.0%)
1	J	0.46	0/3793	0.62	1/5156 (0.0%)
1	K	0.47	0/3773	0.61	0/5129
1	L	0.49	0/3806	0.65	1/5173 (0.0%)
1	M	0.53	0/3797	0.61	1/5161 (0.0%)
1	N	0.50	0/3777	0.61	1/5135 (0.0%)
1	O	0.51	0/3747	0.62	1/5093 (0.0%)
1	P	0.51	0/3793	0.62	0/5156
1	Q	0.57	0/3773	0.63	1/5129 (0.0%)
1	R	0.48	0/3806	0.65	5/5173 (0.1%)
1	S	0.47	0/3797	0.61	0/5161
1	T	0.51	0/3777	0.60	1/5135 (0.0%)
1	U	0.50	0/3747	0.62	0/5093
1	V	0.48	0/3793	0.61	0/5156
1	W	0.49	0/3773	0.61	0/5129
1	X	0.46	0/3806	0.63	2/5173 (0.0%)
All	All	0.50	0/90772	0.62	17/123388 (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	L	0	1
1	R	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	R	318	PRO	N-CA-C	7.49	131.58	112.10
1	R	318	PRO	C-N-CA	7.06	139.35	121.70
1	R	318	PRO	CA-C-N	5.83	130.02	117.20
1	X	380	LEU	CA-CB-CG	5.83	128.70	115.30
1	N	380	LEU	CA-CB-CG	5.74	128.50	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	L	185	VAL	Peptide
1	R	318	PRO	Peptide

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	3693	0	3523	53	0
1	B	3674	0	3503	81	0
1	C	3645	0	3475	61	0
1	D	3690	0	3523	58	0
1	E	3670	0	3507	59	0
1	F	3702	0	3529	78	0
1	G	3693	0	3523	73	0
1	H	3674	0	3503	62	0
1	I	3645	0	3475	56	0
1	J	3690	0	3523	50	0
1	K	3670	0	3507	46	0
1	L	3702	0	3529	69	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	M	3693	0	3523	61	0
1	N	3674	0	3503	63	0
1	O	3645	0	3475	59	0
1	P	3690	0	3523	55	0
1	Q	3670	0	3507	64	0
1	R	3702	0	3529	62	0
1	S	3693	0	3523	47	0
1	T	3674	0	3503	58	0
1	U	3645	0	3475	50	0
1	V	3690	0	3523	49	0
1	W	3670	0	3507	56	0
1	X	3702	0	3529	58	0
2	A	23	0	11	0	0
2	B	23	0	11	1	0
2	C	23	0	11	0	0
2	D	23	0	11	0	0
2	E	23	0	11	0	0
2	F	23	0	11	2	0
2	G	23	0	11	0	0
2	H	23	0	11	0	0
2	I	23	0	11	0	0
2	J	23	0	11	0	0
2	K	23	0	11	0	0
2	L	23	0	11	0	0
2	M	23	0	11	2	0
2	N	23	0	11	1	0
2	O	23	0	11	0	0
2	P	23	0	11	1	0
2	Q	23	0	11	0	0
2	R	23	0	11	1	0
2	S	23	0	11	0	0
2	T	23	0	11	0	0
2	U	23	0	11	0	0
2	V	23	0	11	0	0
2	W	23	0	11	0	0
2	X	23	0	11	0	0
All	All	88848	0	84504	1283	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 7.

The worst 5 of 1283 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:H:313:LEU:HG	1:H:317:ASN:HD22	1.29	0.98
1:P:68:LEU:HD23	1:P:92:HIS:CD2	1.99	0.96
1:R:318:PRO:HG2	1:R:370:GLY:HA2	1.48	0.96
1:J:464:HIS:HD2	1:J:466:TYR:H	1.14	0.95
1:I:210:GLU:H	1:I:213:HIS:HD2	1.12	0.95

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	462/478 (97%)	432 (94%)	25 (5%)	5 (1%)	17 57
1	B	460/478 (96%)	414 (90%)	40 (9%)	6 (1%)	15 52
1	C	456/478 (95%)	408 (90%)	44 (10%)	4 (1%)	21 60
1	D	462/478 (97%)	430 (93%)	30 (6%)	2 (0%)	39 76
1	E	459/478 (96%)	421 (92%)	32 (7%)	6 (1%)	15 52
1	F	463/478 (97%)	419 (90%)	38 (8%)	6 (1%)	15 52
1	G	462/478 (97%)	426 (92%)	28 (6%)	8 (2%)	11 47
1	H	460/478 (96%)	424 (92%)	30 (6%)	6 (1%)	15 52
1	I	456/478 (95%)	421 (92%)	30 (7%)	5 (1%)	17 57
1	J	462/478 (97%)	432 (94%)	26 (6%)	4 (1%)	21 60
1	K	459/478 (96%)	418 (91%)	34 (7%)	7 (2%)	13 49
1	L	463/478 (97%)	424 (92%)	29 (6%)	10 (2%)	8 41
1	M	462/478 (97%)	425 (92%)	32 (7%)	5 (1%)	17 57
1	N	460/478 (96%)	422 (92%)	34 (7%)	4 (1%)	21 60
1	O	456/478 (95%)	421 (92%)	29 (6%)	6 (1%)	15 52
1	P	462/478 (97%)	418 (90%)	35 (8%)	9 (2%)	10 45

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	Q	459/478 (96%)	420 (92%)	31 (7%)	8 (2%)	11 47
1	R	463/478 (97%)	422 (91%)	32 (7%)	9 (2%)	10 45
1	S	462/478 (97%)	424 (92%)	32 (7%)	6 (1%)	15 52
1	T	460/478 (96%)	420 (91%)	34 (7%)	6 (1%)	15 52
1	U	456/478 (95%)	425 (93%)	27 (6%)	4 (1%)	21 60
1	V	462/478 (97%)	429 (93%)	26 (6%)	7 (2%)	13 49
1	W	459/478 (96%)	426 (93%)	29 (6%)	4 (1%)	21 60
1	X	463/478 (97%)	428 (92%)	28 (6%)	7 (2%)	13 49
All	All	11048/11472 (96%)	10149 (92%)	755 (7%)	144 (1%)	15 52

5 of 144 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	211	LYS
1	C	330	TYR
1	D	-1	GLU
1	F	287	ASP
1	F	451	LYS

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	394/405 (97%)	368 (93%)	26 (7%)	21 59
1	B	392/405 (97%)	360 (92%)	32 (8%)	14 48
1	C	388/405 (96%)	349 (90%)	39 (10%)	9 35
1	D	394/405 (97%)	361 (92%)	33 (8%)	14 46
1	E	392/405 (97%)	359 (92%)	33 (8%)	14 46
1	F	395/405 (98%)	351 (89%)	44 (11%)	8 31
1	G	394/405 (97%)	361 (92%)	33 (8%)	14 46
1	H	392/405 (97%)	358 (91%)	34 (9%)	13 45

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	I	388/405 (96%)	363 (94%)	25 (6%)	22	60
1	J	394/405 (97%)	363 (92%)	31 (8%)	15	49
1	K	392/405 (97%)	356 (91%)	36 (9%)	11	40
1	L	395/405 (98%)	363 (92%)	32 (8%)	15	48
1	M	394/405 (97%)	360 (91%)	34 (9%)	13	45
1	N	392/405 (97%)	353 (90%)	39 (10%)	10	37
1	O	388/405 (96%)	351 (90%)	37 (10%)	11	38
1	P	394/405 (97%)	359 (91%)	35 (9%)	12	43
1	Q	392/405 (97%)	357 (91%)	35 (9%)	12	43
1	R	395/405 (98%)	364 (92%)	31 (8%)	16	50
1	S	394/405 (97%)	360 (91%)	34 (9%)	13	45
1	T	392/405 (97%)	362 (92%)	30 (8%)	16	51
1	U	388/405 (96%)	350 (90%)	38 (10%)	10	37
1	V	394/405 (97%)	364 (92%)	30 (8%)	16	51
1	W	392/405 (97%)	358 (91%)	34 (9%)	13	45
1	X	395/405 (98%)	360 (91%)	35 (9%)	12	43
All	All	9420/9720 (97%)	8610 (91%)	810 (9%)	13	45

5 of 810 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	8	LEU
1	N	255	THR
1	W	83	LYS
1	L	160	THR
1	M	192	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 288 such sidechains are listed below:

Mol	Chain	Res	Type
1	K	213	HIS
1	N	213	HIS
1	V	464	HIS
1	K	342	ASN
1	M	30	HIS

5.3.3 RNA

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

24 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	2K9	A	900	-	25,27,27	1.82	5 (20%)	27,40,40	2.04	4 (14%)
2	2K9	B	900	-	25,27,27	1.80	5 (20%)	27,40,40	1.91	3 (11%)
2	2K9	C	900	-	25,27,27	1.87	6 (24%)	27,40,40	2.08	6 (22%)
2	2K9	D	900	-	25,27,27	1.79	5 (20%)	27,40,40	2.04	4 (14%)
2	2K9	E	900	-	25,27,27	1.84	5 (20%)	27,40,40	2.04	3 (11%)
2	2K9	F	900	-	25,27,27	1.69	4 (16%)	27,40,40	2.17	4 (14%)
2	2K9	G	900	-	25,27,27	1.86	5 (20%)	27,40,40	2.08	6 (22%)
2	2K9	H	900	-	25,27,27	1.88	5 (20%)	27,40,40	2.03	4 (14%)
2	2K9	I	900	-	25,27,27	1.92	5 (20%)	27,40,40	2.06	3 (11%)
2	2K9	J	900	-	25,27,27	1.76	5 (20%)	27,40,40	1.99	4 (14%)
2	2K9	K	900	-	25,27,27	1.74	6 (24%)	27,40,40	1.93	3 (11%)
2	2K9	L	900	-	25,27,27	1.74	6 (24%)	27,40,40	1.89	2 (7%)
2	2K9	M	900	-	25,27,27	1.83	5 (20%)	27,40,40	1.99	4 (14%)
2	2K9	N	900	-	25,27,27	1.89	5 (20%)	27,40,40	1.94	3 (11%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	2K9	O	900	-	25,27,27	1.83	5 (20%)	27,40,40	2.02	5 (18%)
2	2K9	P	900	-	25,27,27	1.91	7 (28%)	27,40,40	1.78	2 (7%)
2	2K9	Q	900	-	25,27,27	1.72	6 (24%)	27,40,40	1.92	3 (11%)
2	2K9	R	900	-	25,27,27	1.77	5 (20%)	27,40,40	2.01	2 (7%)
2	2K9	S	900	-	25,27,27	1.87	5 (20%)	27,40,40	1.92	4 (14%)
2	2K9	T	900	-	25,27,27	1.86	5 (20%)	27,40,40	1.82	2 (7%)
2	2K9	U	900	-	25,27,27	1.88	5 (20%)	27,40,40	2.03	5 (18%)
2	2K9	V	900	-	25,27,27	1.86	5 (20%)	27,40,40	1.91	3 (11%)
2	2K9	W	900	-	25,27,27	1.76	6 (24%)	27,40,40	1.87	3 (11%)
2	2K9	X	900	-	25,27,27	1.77	5 (20%)	27,40,40	1.93	2 (7%)

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	2K9	A	900	-	-	0/4/12/12	0/5/5/5
2	2K9	B	900	-	-	0/4/12/12	0/5/5/5
2	2K9	C	900	-	-	0/4/12/12	0/5/5/5
2	2K9	D	900	-	-	0/4/12/12	0/5/5/5
2	2K9	E	900	-	-	0/4/12/12	0/5/5/5
2	2K9	F	900	-	-	0/4/12/12	0/5/5/5
2	2K9	G	900	-	-	0/4/12/12	0/5/5/5
2	2K9	H	900	-	-	0/4/12/12	0/5/5/5
2	2K9	I	900	-	-	0/4/12/12	0/5/5/5
2	2K9	J	900	-	-	0/4/12/12	0/5/5/5
2	2K9	K	900	-	-	0/4/12/12	0/5/5/5
2	2K9	L	900	-	-	0/4/12/12	0/5/5/5
2	2K9	M	900	-	-	0/4/12/12	0/5/5/5
2	2K9	N	900	-	-	0/4/12/12	0/5/5/5
2	2K9	O	900	-	-	0/4/12/12	0/5/5/5
2	2K9	P	900	-	-	0/4/12/12	0/5/5/5
2	2K9	Q	900	-	-	0/4/12/12	0/5/5/5
2	2K9	R	900	-	-	0/4/12/12	0/5/5/5
2	2K9	S	900	-	-	0/4/12/12	0/5/5/5
2	2K9	T	900	-	-	0/4/12/12	0/5/5/5
2	2K9	U	900	-	-	0/4/12/12	0/5/5/5
2	2K9	V	900	-	-	0/4/12/12	0/5/5/5
2	2K9	W	900	-	-	0/4/12/12	0/5/5/5

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	2K9	X	900	-	-	0/4/12/12	0/5/5/5

The worst 5 of 126 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	900	2K9	C4-C2	-5.62	1.37	1.46
2	O	900	2K9	C4-C2	-5.45	1.37	1.46
2	H	900	2K9	C4-C2	-5.35	1.37	1.46
2	I	900	2K9	C4-C2	-5.27	1.37	1.46
2	E	900	2K9	C4-C2	-5.23	1.37	1.46

The worst 5 of 84 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	H	900	2K9	C1-C3-C7	-7.92	99.65	107.93
2	O	900	2K9	C1-C3-C7	-7.58	100.01	107.93
2	E	900	2K9	C1-C3-C7	-7.46	100.13	107.93
2	I	900	2K9	C1-C3-C7	-7.32	100.28	107.93
2	A	900	2K9	C1-C3-C7	-7.29	100.31	107.93

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	900	2K9	1	0
2	F	900	2K9	2	0
2	M	900	2K9	2	0
2	N	900	2K9	1	0
2	P	900	2K9	1	0
2	R	900	2K9	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues

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, 95^{th} 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	466/478 (97%)	-0.16	19 (4%)	41	34	31, 54, 85, 128	0
1	B	464/478 (97%)	-0.09	15 (3%)	51	44	33, 55, 85, 125	0
1	C	460/478 (96%)	-0.18	6 (1%)	79	74	32, 53, 83, 124	0
1	D	466/478 (97%)	-0.15	9 (1%)	70	63	31, 52, 82, 124	0
1	E	463/478 (96%)	-0.19	5 (1%)	82	78	31, 52, 82, 124	0
1	F	467/478 (97%)	-0.17	5 (1%)	82	78	31, 54, 85, 124	0
1	G	466/478 (97%)	-0.16	13 (2%)	56	50	31, 54, 84, 124	0
1	H	464/478 (97%)	-0.18	12 (2%)	59	53	32, 55, 84, 124	0
1	I	460/478 (96%)	-0.24	7 (1%)	76	71	32, 53, 82, 124	0
1	J	466/478 (97%)	-0.23	8 (1%)	73	67	31, 52, 82, 124	0
1	K	463/478 (96%)	-0.13	9 (1%)	70	63	31, 53, 82, 124	0
1	L	467/478 (97%)	-0.23	6 (1%)	79	74	32, 53, 84, 124	0
1	M	466/478 (97%)	-0.07	14 (3%)	54	47	33, 54, 85, 124	0
1	N	464/478 (97%)	-0.21	7 (1%)	76	71	32, 53, 83, 124	0
1	O	460/478 (96%)	-0.18	9 (1%)	68	62	32, 54, 83, 124	0
1	P	466/478 (97%)	-0.21	10 (2%)	67	60	32, 54, 83, 124	0
1	Q	463/478 (96%)	-0.15	9 (1%)	70	63	34, 55, 83, 125	0
1	R	467/478 (97%)	-0.18	11 (2%)	62	55	32, 53, 84, 124	0
1	S	466/478 (97%)	-0.15	11 (2%)	62	55	33, 54, 84, 124	0
1	T	464/478 (97%)	-0.20	8 (1%)	73	67	32, 54, 83, 124	0
1	U	460/478 (96%)	-0.20	10 (2%)	65	59	32, 53, 83, 124	0
1	V	466/478 (97%)	-0.26	9 (1%)	70	63	31, 53, 82, 124	0
1	W	463/478 (96%)	-0.23	7 (1%)	76	71	33, 54, 82, 124	0
1	X	467/478 (97%)	-0.20	10 (2%)	67	60	31, 53, 84, 124	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
All	All	11144/11472 (97%)	-0.18	229 (2%)	67	60	31, 54, 84, 128	0

The worst 5 of 229 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	R	402	TYR	6.7
1	O	406	PRO	5.8
1	M	405	PRO	4.9
1	M	406	PRO	4.7
1	A	61	HIS	4.7

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	2K9	U	900	23/23	0.93	0.32	3.20	69,73,75,77	0
2	2K9	O	900	23/23	0.89	0.44	3.11	82,85,88,88	0
2	2K9	G	900	23/23	0.94	0.26	1.77	54,54,56,57	0
2	2K9	R	900	23/23	0.95	0.32	1.68	57,58,59,59	0
2	2K9	P	900	23/23	0.93	0.27	1.36	64,65,66,67	0
2	2K9	X	900	23/23	0.95	0.28	1.34	60,61,61,62	0
2	2K9	E	900	23/23	0.97	0.25	1.30	43,44,46,46	0
2	2K9	T	900	23/23	0.96	0.24	1.11	75,75,76,76	0
2	2K9	A	900	23/23	0.96	0.27	1.01	58,60,62,62	0
2	2K9	H	900	23/23	0.94	0.27	0.88	90,90,91,92	0
2	2K9	F	900	23/23	0.90	0.27	0.70	64,67,67,68	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
2	2K9	M	900	23/23	0.91	0.27	0.67	85,86,87,88	0
2	2K9	Q	900	23/23	0.92	0.24	0.42	113,114,114,114	0
2	2K9	J	900	23/23	0.96	0.24	0.39	56,58,60,60	0
2	2K9	B	900	23/23	0.93	0.26	0.36	100,102,103,104	0
2	2K9	N	900	23/23	0.94	0.24	0.35	58,60,61,62	0
2	2K9	L	900	23/23	0.96	0.24	0.18	50,51,52,53	0
2	2K9	S	900	23/23	0.93	0.22	0.05	67,69,71,72	0
2	2K9	D	900	23/23	0.95	0.22	-0.05	47,51,53,54	0
2	2K9	W	900	23/23	0.92	0.22	-0.07	76,77,77,78	0
2	2K9	V	900	23/23	0.95	0.23	-0.15	39,43,45,47	0
2	2K9	C	900	23/23	0.96	0.20	-0.17	47,48,49,50	0
2	2K9	K	900	23/23	0.95	0.20	-0.40	46,48,49,50	0
2	2K9	I	900	23/23	0.95	0.19	-0.44	54,55,57,57	0

6.5 Other polymers

There are no such residues in this entry.