wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 11:42 PM GMT

PDB ID	:	1YAE
Title	:	Structure of the Kainate Receptor Subunit GluR6 Agonist Binding Domain Complexed with Domoic Acid
Authors	:	Nanao, M.H.; Green, T.; Stern-Bach, Y.; Heinemann, S.F.; Choe, S.
Deposited on	:	2004-12-17
Resolution	:	3.11 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:

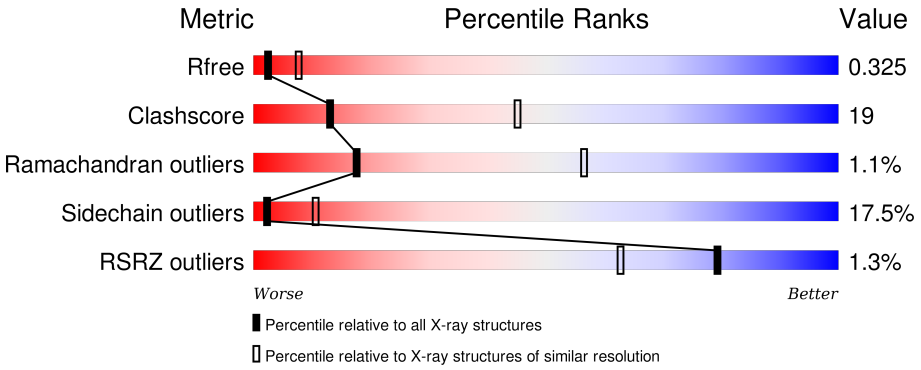
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 3.11 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	1112 (3.16-3.08)
Clashscore	102246	1218 (3.16-3.08)
Ramachandran outliers	100387	1175 (3.16-3.08)
Sidechain outliers	100360	1175 (3.16-3.08)
RSRZ outliers	91569	1114 (3.16-3.08)

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	312	50% 28% 5% 16%
1	B	312	47% 29% 8% • 16%
1	C	312	49% 29% • • 18%
1	D	312	43% 34% 6% • 17%
1	E	312	2% 46% 28% 7% • 18%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	312	2% 29% 31% 13% 28%

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	NAG	A	1	X	-	-	-
3	NAG	C	1302	X	-	-	-
3	NAG	D	1401	X	-	-	-
3	NAG	D	1402	X	-	-	-
3	NAG	E	1502	X	-	-	-
3	NAG	F	1601	X	-	-	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 12512 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 Glutamate receptor, ionotropic kainate 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	261	Total	C	N	O	S	0	0	0
			2079	1324	346	397	12			
1	B	261	Total	C	N	O	S	0	0	0
			2082	1325	347	398	12			
1	C	257	Total	C	N	O	S	0	0	0
			2053	1309	341	392	11			
1	D	259	Total	C	N	O	S	0	0	0
			2067	1316	344	395	12			
1	E	257	Total	C	N	O	S	0	0	0
			2056	1312	341	392	11			
1	F	226	Total	C	N	O	S	0	0	0
			1815	1163	297	344	11			

There are 54 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	653	GLY	-	LINKER	GB 56280
A	654	GLY	-	LINKER	GB 56280
A	655	SER	-	LINKER	GB 56280
A	656	LEU	-	LINKER	GB 56280
A	657	VAL	-	LINKER	GB 56280
A	658	PRO	-	LINKER	GB 56280
A	659	ARG	-	LINKER	GB 56280
A	660	GLY	-	LINKER	GB 56280
A	661	SER	-	LINKER	GB 56280
B	653	GLY	-	LINKER	GB 56280
B	654	GLY	-	LINKER	GB 56280
B	655	SER	-	LINKER	GB 56280
B	656	LEU	-	LINKER	GB 56280
B	657	VAL	-	LINKER	GB 56280
B	658	PRO	-	LINKER	GB 56280
B	659	ARG	-	LINKER	GB 56280
B	660	GLY	-	LINKER	GB 56280

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	661	SER	-	LINKER	GB 56280
C	653	GLY	-	LINKER	GB 56280
C	654	GLY	-	LINKER	GB 56280
C	655	SER	-	LINKER	GB 56280
C	656	LEU	-	LINKER	GB 56280
C	657	VAL	-	LINKER	GB 56280
C	658	PRO	-	LINKER	GB 56280
C	659	ARG	-	LINKER	GB 56280
C	660	GLY	-	LINKER	GB 56280
C	661	SER	-	LINKER	GB 56280
D	653	GLY	-	LINKER	GB 56280
D	654	GLY	-	LINKER	GB 56280
D	655	SER	-	LINKER	GB 56280
D	656	LEU	-	LINKER	GB 56280
D	657	VAL	-	LINKER	GB 56280
D	658	PRO	-	LINKER	GB 56280
D	659	ARG	-	LINKER	GB 56280
D	660	GLY	-	LINKER	GB 56280
D	661	SER	-	LINKER	GB 56280
E	653	GLY	-	LINKER	GB 56280
E	654	GLY	-	LINKER	GB 56280
E	655	SER	-	LINKER	GB 56280
E	656	LEU	-	LINKER	GB 56280
E	657	VAL	-	LINKER	GB 56280
E	658	PRO	-	LINKER	GB 56280
E	659	ARG	-	LINKER	GB 56280
E	660	GLY	-	LINKER	GB 56280
E	661	SER	-	LINKER	GB 56280
F	653	GLY	-	LINKER	GB 56280
F	654	GLY	-	LINKER	GB 56280
F	655	SER	-	LINKER	GB 56280
F	656	LEU	-	LINKER	GB 56280
F	657	VAL	-	LINKER	GB 56280
F	658	PRO	-	LINKER	GB 56280
F	659	ARG	-	LINKER	GB 56280
F	660	GLY	-	LINKER	GB 56280
F	661	SER	-	LINKER	GB 56280

- Molecule 2 is a polymer of unknown type called SUGAR (NAG-FUC).

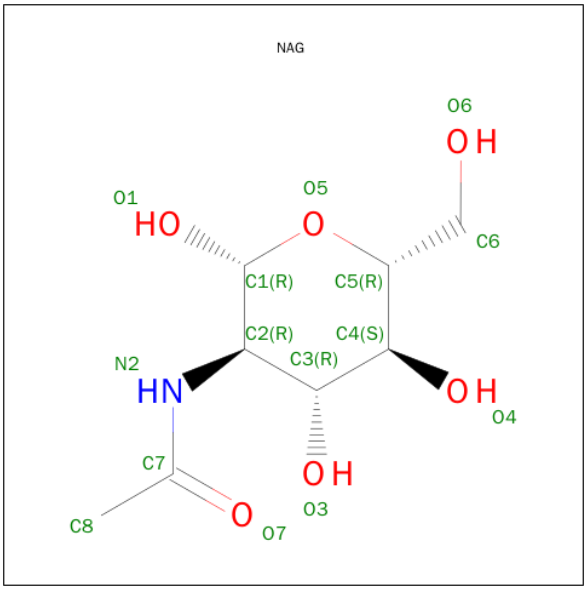
Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf

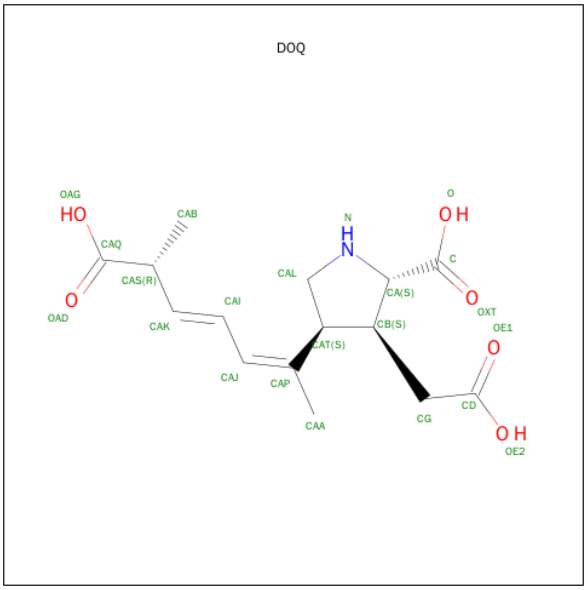
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	2	Total C N O 24 14 1 9	0	0

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total C N O 14 8 1 5	0	0
3	B	1	Total C N O 14 8 1 5	0	0
3	C	1	Total C N O 14 8 1 5	0	0
3	C	1	Total C N O 14 8 1 5	0	0
3	D	1	Total C N O 14 8 1 5	0	0
3	D	1	Total C N O 14 8 1 5	0	0
3	E	1	Total C N O 14 8 1 5	0	0
3	E	1	Total C N O 14 8 1 5	0	0
3	F	1	Total C N O 14 8 1 5	0	0

- Molecule 4 is (2S,3S,4S)-2-CARBOXY-4-[(1Z,3E,5R)-5-CARBOXY-1-METHYL-1,3-HEXADIENYL]-3-PYRROLIDINEACETIC ACID (three-letter code: DOQ) (formula: $C_{15}H_{21}NO_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O 22 15 1 6	0	0
4	B	1	Total C N O 22 15 1 6	0	0
4	C	1	Total C N O 22 15 1 6	0	0
4	D	1	Total C N O 22 15 1 6	0	0
4	E	1	Total C N O 22 15 1 6	0	0
4	F	1	Total C N O 22 15 1 6	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	13	Total O 13 13	0	0
5	B	19	Total O 19 19	0	0
5	C	14	Total O 14 14	0	0
5	D	16	Total O 16 16	0	0

Continued on next page...

Continued from previous page...

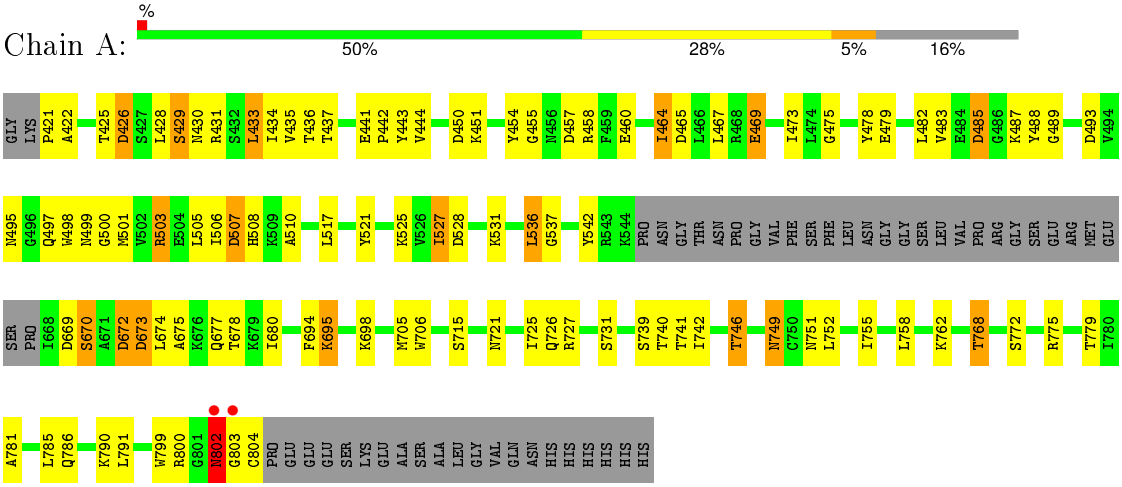
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	E	9	Total O 9 9	0	0
5	F	7	Total O 7 7	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

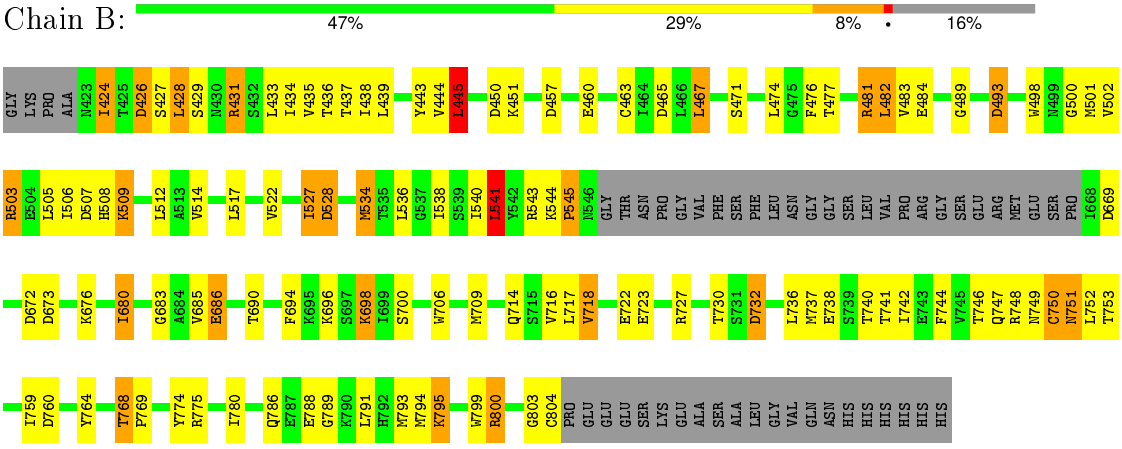
- Molecule 1: Glutamate receptor, ionotropic kainate 2

Chain A:



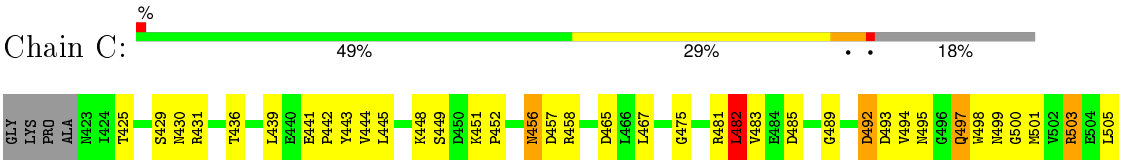
- Molecule 1: Glutamate receptor, ionotropic kainate 2

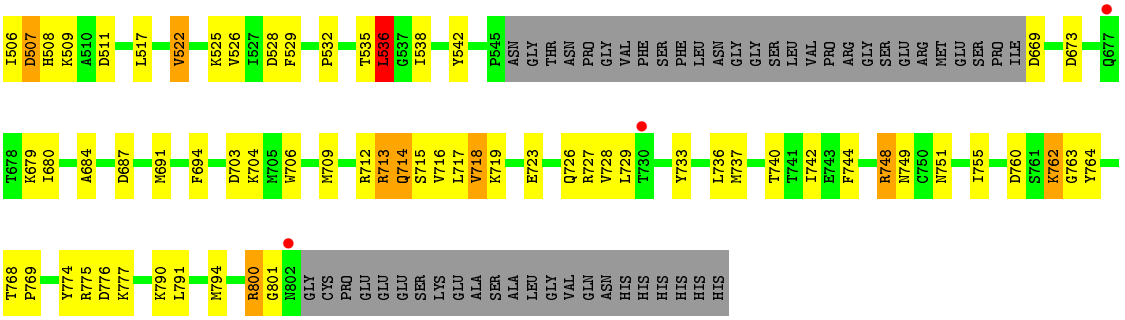
Chain B:



- Molecule 1: Glutamate receptor, ionotropic kainate 2

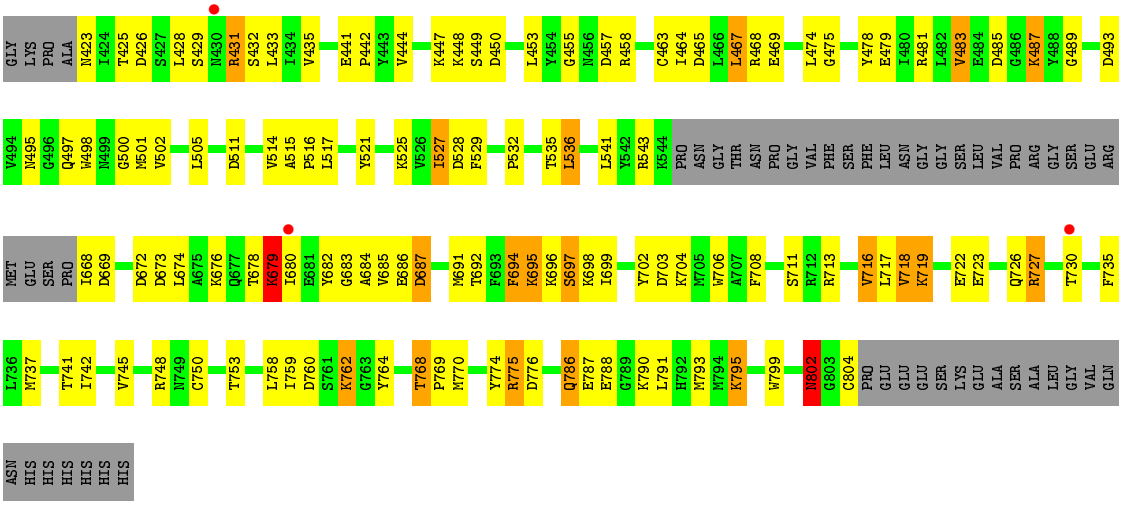
Chain C:





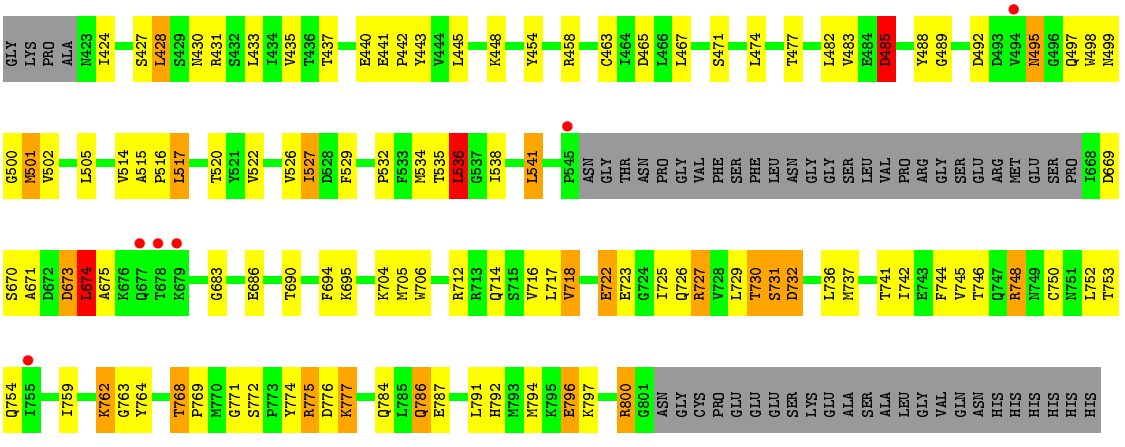
• Molecule 1: Glutamate receptor, ionotropic kainate 2

Chain D:



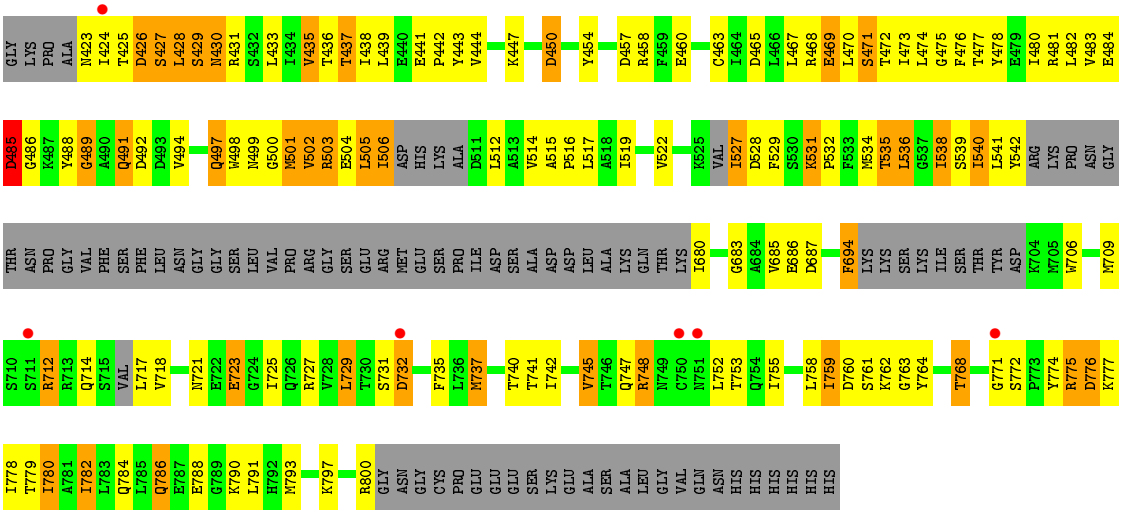
• Molecule 1: Glutamate receptor, ionotropic kainate 2

Chain E:



• Molecule 1: Glutamate receptor, ionotropic kainate 2

Chain F:


GLY LYS PRO ALA N423 I424 T425 D426 S427 L428 S429 N430 R431 S432 L433 I434 V435 T436 T437 I438 L439 E440 E441 P442 Y443 V444 K447 D450 Y454 D457 R458 F459 E460 C463 I464 D465 L466 L467 R468 E469 L470 S471 T472 I473 L474 G475 F476 T477 Y478 E479 I480 R481 L482 V483 E484
D485 G486 K487 Y488 G489 A490 Q491 D492 D493 V494 Q497 W498 N499 G500 M501 V502 R503 E504 L505 I506 ASP HIS LYS ALA D511 L512 A513 V514 A515 P516 L517 A518 I519 V522 R525 VAL I527 D528 F529 S530 K531 P532 F533 M534 T535 L536 G537 I538 S539 I540 L541 Y542 ARG LYS PRO ASN GLY
THR ASN PRO GLY VAL PHE SER PHE LEU ASN GLY GLY SER LEU VAL PRO ARG GLY SER GLU ARG MET GLU SER PRO ILE ASP SER ALA ASP ASP LEU ALA LYS GLN THR LYS I680 G683 A684 V685 E686 D687 F694 LYS LYS SER LYS ILE SER THR TYR ASP K704 M705 W706 M709
S710 S711 R712 R713 Q714 S715 VAL L717 V718 N721 E722 E723 G724 I725 Q726 R727 Y728 L729 T730 S731 D732 F735 L736 M737 T740 T741 I742 V745 T746 Q747 R748 W749 C750 N751 L752 T753 Q754 I755 L758 I759 D760 S761 K762 G763 Y764 T768 G771 S772 P773 Y774 R775 D776 K777
I778 T779 I780 A781 I782 L783 Q784 L785 Q786 E787 E788 G789 K790 L791 H792 M793 K797 R800 GLY ASN GLY CYS PRO GLU GLU GLU GLU SER LYS GLU ALA SER ALA LEU GLY VAL GLN ASN HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	246.36Å 106.57Å 172.69Å 90.00° 133.19° 90.00°	Depositor
Resolution (Å)	45.00 – 3.11 44.49 – 3.11	Depositor EDS
% Data completeness (in resolution range)	94.3 (45.00-3.11) 94.4 (44.49-3.11)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.12	Depositor
$< I/\sigma(I) >$ [1]	1.83 (at 3.12Å)	Xtriage
Refinement program	REFMAC 5.2.0003	Depositor
R, R_{free}	0.275 , 0.334 0.269 , 0.325	Depositor DCC
R_{free} test set	2817 reflections (5.33%)	DCC
Wilson B-factor (Å^2)	69.4	Xtriage
Anisotropy	0.176	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 48.5	EDS
Estimated twinning fraction	0.006 for h+2*l,k,-h-l 0.008 for h,-k,-h-l 0.021 for -h-2*l,-k,l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Outliers	0 of 55645 reflections	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	12512	wwPDB-VP
Average B, all atoms (Å^2)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.17% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: DOQ, NAG, FUC

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	0.47	0/2116	0.83	6/2850 (0.2%)
1	B	0.46	0/2119	0.87	10/2855 (0.4%)
1	C	0.46	0/2090	0.85	10/2816 (0.4%)
1	D	0.58	2/2103 (0.1%)	0.87	11/2832 (0.4%)
1	E	0.51	0/2093	0.85	8/2820 (0.3%)
1	F	0.63	0/1845	0.86	4/2482 (0.2%)
All	All	0.52	2/12366 (0.0%)	0.86	49/16655 (0.3%)

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
2	A	1	0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	697	SER	CB-OG	-11.88	1.26	1.42
1	D	696	LYS	CB-CG	7.02	1.71	1.52

The worst 5 of 49 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	545	PRO	CA-N-CD	-12.18	94.44	111.50
1	B	732	ASP	CB-CG-OD2	6.29	123.96	118.30
1	B	493	ASP	CB-CG-OD2	6.19	123.87	118.30
1	D	694	PHE	CB-CG-CD1	-6.05	116.57	120.80
1	D	750	CYS	CA-CB-SG	-5.97	103.25	114.00

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	1	NAG	C1

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2079	0	2084	66	0
1	B	2082	0	2082	90	0
1	C	2053	0	2054	57	0
1	D	2067	0	2069	71	0
1	E	2056	0	2065	79	0
1	F	1815	0	1809	121	0
2	A	24	0	22	1	0
3	B	28	0	26	3	0
3	C	28	0	26	0	0
3	D	28	0	26	2	0
3	E	28	0	26	0	0
3	F	14	0	13	1	0
4	A	22	0	18	2	0
4	B	22	0	18	2	0
4	C	22	0	18	0	0
4	D	22	0	18	1	0
4	E	22	0	18	1	0
4	F	22	0	18	2	0
5	A	13	0	0	0	0
5	B	19	0	0	0	0
5	C	14	0	0	1	0
5	D	16	0	0	1	0
5	E	9	0	0	1	0
5	F	7	0	0	3	0
All	All	12512	0	12410	473	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 19.

The worst 5 of 473 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:F:498:TRP:CE3	1:F:502:VAL:HG11	1.66	1.28
1:F:506:ILE:HD13	1:F:506:ILE:C	1.67	1.11
1:E:727:ARG:HG2	1:E:727:ARG:HH11	1.18	1.09
1:F:498:TRP:HE3	1:F:502:VAL:HG11	0.99	1.08
1:F:447:LYS:HG2	1:F:460:GLU:HG3	1.40	1.01

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	257/312 (82%)	234 (91%)	21 (8%)	2 (1%)	24	64
1	B	257/312 (82%)	239 (93%)	17 (7%)	1 (0%)	39	77
1	C	253/312 (81%)	239 (94%)	13 (5%)	1 (0%)	39	77
1	D	255/312 (82%)	235 (92%)	17 (7%)	3 (1%)	16	53
1	E	253/312 (81%)	230 (91%)	21 (8%)	2 (1%)	24	64
1	F	214/312 (69%)	186 (87%)	21 (10%)	7 (3%)	5	26
All	All	1489/1872 (80%)	1363 (92%)	110 (7%)	16 (1%)	17	56

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	802	ASN
1	F	430	ASN
1	E	771	GLY
1	F	427	SER
1	F	429	SER

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	228/270 (84%)	197 (86%)	31 (14%)	5 20
1	B	229/270 (85%)	190 (83%)	39 (17%)	2 11
1	C	225/270 (83%)	193 (86%)	32 (14%)	4 18
1	D	227/270 (84%)	186 (82%)	41 (18%)	2 9
1	E	226/270 (84%)	185 (82%)	41 (18%)	2 9
1	F	199/270 (74%)	149 (75%)	50 (25%)	1 2
All	All	1334/1620 (82%)	1100 (82%)	234 (18%)	2 10

5 of 234 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	433	LEU
1	D	719	LYS
1	F	687	ASP
1	D	467	LEU
1	D	672	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	491	GLN
1	C	714	GLN
1	F	491	GLN
1	C	456	ASN
1	F	497	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)

2 carbohydrates 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	NAG	A	1	1,2	14,14,15	0.52	0	15,19,21	1.48	1 (6%)
2	FUC	A	2	2	10,10,11	0.79	0	14,14,16	2.13	6 (42%)

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	NAG	A	1	1,2	1/1/5/7	0/6/23/26	0/1/1/1
2	FUC	A	2	2	-	0/0/17/20	0/1/1/1

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	2	FUC	O5-C5-C6	2.26	109.87	106.13
2	A	2	FUC	C3-C4-C5	2.43	113.81	109.72
2	A	2	FUC	C2-C3-C4	2.67	115.58	111.04
2	A	2	FUC	O5-C5-C4	2.85	114.47	109.53
2	A	2	FUC	C1-O5-C5	3.83	118.30	112.38

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	1	NAG	C1

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1	NAG	1	0
2	A	2	FUC	1	0

5.6 Ligand geometry (i)

15 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	$\#\|Z\| > 2$
4	DOQ	A	1103	-	11,22,22	2.26	2 (18%)	6,30,30	2.31	3 (50%)
3	NAG	B	1201	1	14,14,15	0.80	1 (7%)	15,19,21	1.31	2 (13%)
3	NAG	B	1202	1	14,14,15	0.57	0	15,19,21	1.73	1 (6%)
4	DOQ	B	1203	-	11,22,22	2.28	1 (9%)	6,30,30	2.66	1 (16%)
3	NAG	C	1301	1	14,14,15	0.71	0	15,19,21	2.06	4 (26%)
3	NAG	C	1302	1	14,14,15	0.71	1 (7%)	15,19,21	1.25	3 (20%)
4	DOQ	C	1303	-	11,22,22	2.74	1 (9%)	6,30,30	1.30	1 (16%)
3	NAG	D	1401	1	14,14,15	0.52	0	15,19,21	1.64	4 (26%)
3	NAG	D	1402	1	14,14,15	0.68	0	15,19,21	1.33	1 (6%)
4	DOQ	D	1403	-	11,22,22	2.67	2 (18%)	6,30,30	1.97	1 (16%)
3	NAG	E	1501	1	14,14,15	0.69	0	15,19,21	1.54	3 (20%)
3	NAG	E	1502	1	14,14,15	0.85	0	15,19,21	1.53	3 (20%)
4	DOQ	E	1503	-	11,22,22	2.46	2 (18%)	6,30,30	2.80	1 (16%)
3	NAG	F	1601	1	14,14,15	0.53	0	15,19,21	1.23	1 (6%)
4	DOQ	F	1603	-	11,22,22	2.36	2 (18%)	6,30,30	2.27	3 (50%)

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
4	DOQ	A	1103	-	-	0/11/35/35	0/1/1/1
3	NAG	B	1201	1	-	0/6/23/26	0/1/1/1
3	NAG	B	1202	1	-	0/6/23/26	0/1/1/1
4	DOQ	B	1203	-	-	0/11/35/35	0/1/1/1
3	NAG	C	1301	1	-	0/6/23/26	0/1/1/1
3	NAG	C	1302	1	1/1/5/7	0/6/23/26	0/1/1/1
4	DOQ	C	1303	-	-	0/11/35/35	0/1/1/1
3	NAG	D	1401	1	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	D	1402	1	1/1/5/7	0/6/23/26	0/1/1/1
4	DOQ	D	1403	-	-	0/11/35/35	0/1/1/1
3	NAG	E	1501	1	-	0/6/23/26	0/1/1/1
3	NAG	E	1502	1	1/1/5/7	0/6/23/26	0/1/1/1
4	DOQ	E	1503	-	-	0/11/35/35	0/1/1/1
3	NAG	F	1601	1	1/1/5/7	0/6/23/26	0/1/1/1
4	DOQ	F	1603	-	-	0/11/35/35	0/1/1/1

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1103	DOQ	CAA-CAP	2.01	1.54	1.50
4	D	1403	DOQ	CAA-CAP	2.06	1.54	1.50
3	C	1302	NAG	C1-C2	2.15	1.55	1.52
4	F	1603	DOQ	CAA-CAP	2.22	1.54	1.50
4	E	1503	DOQ	CAS-CAK	2.29	1.54	1.51

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	E	1503	DOQ	CAI-CAJ-CAP	-6.29	120.77	127.39
4	B	1203	DOQ	CAI-CAJ-CAP	-5.87	121.21	127.39
4	F	1603	DOQ	CAI-CAJ-CAP	-4.20	122.97	127.39
4	A	1103	DOQ	CAI-CAJ-CAP	-3.74	123.46	127.39
4	D	1403	DOQ	CAI-CAJ-CAP	-3.61	123.59	127.39

All (5) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	D	1402	NAG	C1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atom
3	E	1502	NAG	C1
3	D	1401	NAG	C1
3	C	1302	NAG	C1
3	F	1601	NAG	C1

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1103	DOQ	2	0
3	B	1202	NAG	3	0
4	B	1203	DOQ	2	0
3	D	1402	NAG	2	0
4	D	1403	DOQ	1	0
4	E	1503	DOQ	1	0
3	F	1601	NAG	1	0
4	F	1603	DOQ	2	0

5.7 Other polymers

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	261/312 (83%)	-0.06	2 (0%) 87 75	24, 58, 73, 77	0
1	B	261/312 (83%)	-0.03	0 100 100	24, 57, 74, 79	0
1	C	257/312 (82%)	-0.02	3 (1%) 81 65	26, 58, 77, 86	0
1	D	259/312 (83%)	-0.03	3 (1%) 81 65	25, 58, 73, 79	0
1	E	257/312 (82%)	0.05	6 (2%) 64 41	26, 58, 81, 93	0
1	F	226/312 (72%)	0.09	6 (2%) 58 35	46, 58, 74, 79	0
All	All	1521/1872 (81%)	-0.00	20 (1%) 79 63	24, 58, 75, 93	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	677	GLN	3.2
1	E	678	THR	3.1
1	E	545	PRO	3.1
1	C	730	THR	2.8
1	E	679	LYS	2.7

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

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

6.3 Carbohydrates

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, 95^{th} 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	FUC	A	2	10/11	0.89	0.30	-	81,82,82,82	0
2	NAG	A	1	14/15	0.83	0.18	-	78,79,80,80	0

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, 95^{th} 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
3	NAG	B	1202	14/15	0.82	0.37	1.81	81,82,85,85	0
4	DOQ	E	1503	22/22	0.88	0.27	1.67	50,53,58,59	0
4	DOQ	F	1603	22/22	0.93	0.20	-0.23	59,60,62,64	0
4	DOQ	D	1403	22/22	0.92	0.20	-0.24	41,42,44,44	0
4	DOQ	A	1103	22/22	0.96	0.20	-0.80	35,36,41,42	0
4	DOQ	C	1303	22/22	0.95	0.19	-0.81	42,44,48,49	0
4	DOQ	B	1203	22/22	0.95	0.21	-0.91	33,35,37,38	0
3	NAG	D	1402	14/15	0.83	0.34	-	83,85,87,87	0
3	NAG	C	1301	14/15	0.77	0.22	-	82,84,85,85	0
3	NAG	D	1401	14/15	0.82	0.34	-	84,87,87,88	0
3	NAG	E	1502	14/15	0.72	0.43	-	84,86,88,88	0
3	NAG	B	1201	14/15	0.82	0.26	-	79,80,81,81	0
3	NAG	C	1302	14/15	0.81	0.36	-	85,88,89,90	0
3	NAG	E	1501	14/15	0.69	0.33	-	85,87,88,88	0
3	NAG	F	1601	14/15	0.81	0.31	-	88,90,91,92	0

6.5 Other polymers (i)

There are no such residues in this entry.