



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 2, 2024 – 12:32 PM EDT

PDB ID : 5VTA  
Title : Co-Crystal Structure of DPPIV with a Chemibody Inhibitor  
Authors : Wang, Z.; Johnstone, S.; Cheng, A.  
Deposited on : 2017-05-16  
Resolution : 2.80 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 : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.20.1  
EDS : 3.0  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

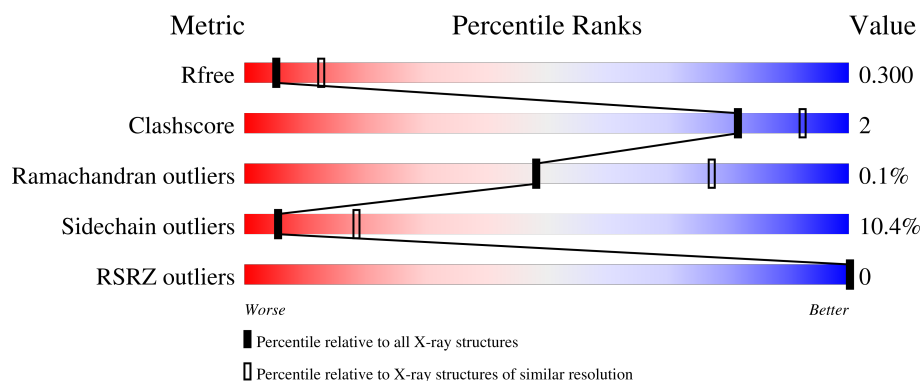
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*


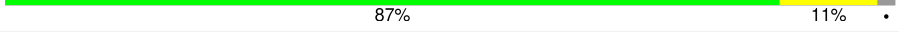
The reported resolution of this entry is 2.80 Å.

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))
$R_{free}$	164625	3657 (2.80-2.80)
Clashscore	180529	4123 (2.80-2.80)
Ramachandran outliers	177936	4071 (2.80-2.80)
Sidechain outliers	177891	4073 (2.80-2.80)
RSRZ outliers	164620	3659 (2.80-2.80)

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  $\leq 5\%$ . 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	739	
1	B	739	
1	C	739	
1	D	739	
2	E	213	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	213	<div><div></div><div>38%9%53%</div></div>
2	J	213	<div><div></div><div>30%10%58%</div></div>
2	L	213	<div><div></div><div>71%23%</div></div>
3	F	217	<div><div></div><div>88%12%</div></div>
3	H	217	<div><div></div><div>82%17%</div></div>
3	I	217	<div><div></div><div>42%8%47%</div></div>
3	K	217	<div><div></div><div>38%10%49%</div></div>

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 33875 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 Dipeptidyl peptidase 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	728	Total	C	N	O	S	0	0	0
			5929	3798	981	1124	26			
1	B	718	Total	C	N	O	S	0	1	0
			5852	3752	967	1107	26			
1	C	728	Total	C	N	O	S	0	0	0
			5929	3798	981	1124	26			
1	D	727	Total	C	N	O	S	0	0	0
			5921	3792	980	1123	26			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	768	HIS	-	expression tag	UNP P14740
A	769	HIS	-	expression tag	UNP P14740
A	770	HIS	-	expression tag	UNP P14740
A	771	HIS	-	expression tag	UNP P14740
A	772	HIS	-	expression tag	UNP P14740
A	773	HIS	-	expression tag	UNP P14740
A	774	HIS	-	expression tag	UNP P14740
A	775	HIS	-	expression tag	UNP P14740
B	768	HIS	-	expression tag	UNP P14740
B	769	HIS	-	expression tag	UNP P14740
B	770	HIS	-	expression tag	UNP P14740
B	771	HIS	-	expression tag	UNP P14740
B	772	HIS	-	expression tag	UNP P14740
B	773	HIS	-	expression tag	UNP P14740
B	774	HIS	-	expression tag	UNP P14740
B	775	HIS	-	expression tag	UNP P14740
C	768	HIS	-	expression tag	UNP P14740
C	769	HIS	-	expression tag	UNP P14740
C	770	HIS	-	expression tag	UNP P14740
C	771	HIS	-	expression tag	UNP P14740
C	772	HIS	-	expression tag	UNP P14740

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	773	HIS	-	expression tag	UNP P14740
C	774	HIS	-	expression tag	UNP P14740
C	775	HIS	-	expression tag	UNP P14740
D	768	HIS	-	expression tag	UNP P14740
D	769	HIS	-	expression tag	UNP P14740
D	770	HIS	-	expression tag	UNP P14740
D	771	HIS	-	expression tag	UNP P14740
D	772	HIS	-	expression tag	UNP P14740
D	773	HIS	-	expression tag	UNP P14740
D	774	HIS	-	expression tag	UNP P14740
D	775	HIS	-	expression tag	UNP P14740

- Molecule 2 is a protein called Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	208	Total	C	N	O	S	0	0	0
			1578	986	268	317	7			
2	E	204	Total	C	N	O	S	0	0	0
			1546	966	263	310	7			
2	G	100	Total	C	N	O	S	0	0	0
			747	467	131	144	5			
2	J	89	Total	C	N	O	S	0	0	0
			669	420	119	125	5			

- Molecule 3 is a protein called Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	H	217	Total	C	N	O	S	0	0	0
			1639	1040	267	327	5			
3	F	217	Total	C	N	O	S	0	0	0
			1639	1040	267	327	5			
3	I	115	Total	C	N	O	S	0	0	0
			899	573	146	177	3			
3	K	110	Total	C	N	O	S	0	0	0
			860	549	141	167	3			

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	C	1	Total	C	N	O	0	0
			14	8	1	5		
4	C	1	Total	C	N	O	0	0
			14	8	1	5		
4	C	1	Total	C	N	O	0	0
			14	8	1	5		
4	C	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	C	1	Total	C	N	O	0	0
			14	8	1	5		
4	D	1	Total	C	N	O	0	0
			14	8	1	5		
4	D	1	Total	C	N	O	0	0
			14	8	1	5		
4	D	1	Total	C	N	O	0	0
			14	8	1	5		
4	D	1	Total	C	N	O	0	0
			14	8	1	5		
4	D	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula:  $C_8H_{18}O_5$ ).



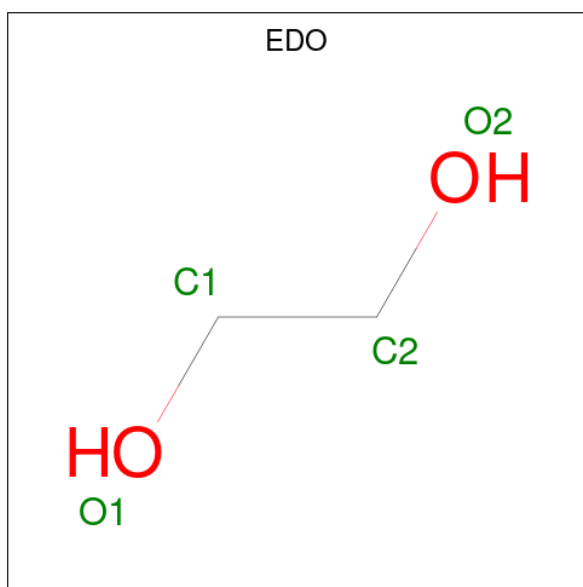
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			13	8	5		

- Molecule 6 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			7	4	3		
6	A	1	Total	C	O	0	0
			7	4	3		
6	D	1	Total	C	O	0	0
			7	4	3		

- Molecule 7 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			4	2	2		

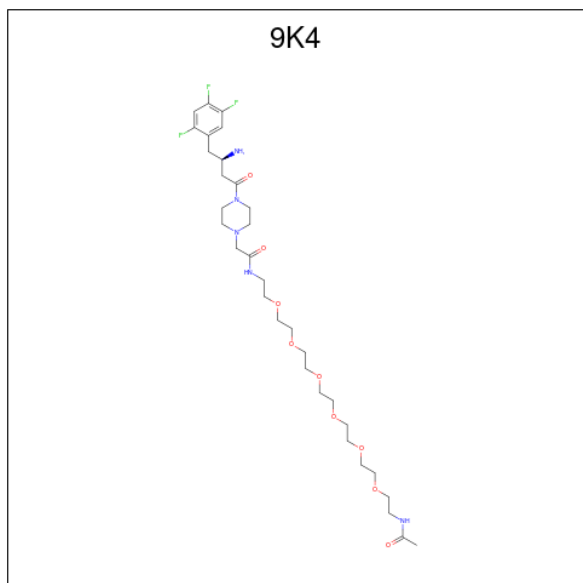
*Continued on next page...*



Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			4	2	2		
7	A	1	Total	C	O	0	0
			4	2	2		
7	A	1	Total	C	O	0	0
			4	2	2		
7	B	1	Total	C	O	0	0
			4	2	2		
7	C	1	Total	C	O	0	0
			4	2	2		
7	C	1	Total	C	O	0	0
			4	2	2		
7	C	1	Total	C	O	0	0
			4	2	2		
7	C	1	Total	C	O	0	0
			4	2	2		

- Molecule 8 is 2-{4-[(3R)-3-amino-4-(2,4,5-trifluorophenyl)butanoyl]piperazin-1-yl}-N-(22-oxo-3,6,9,12,15,18-hexaoxa-21-azatricosan-1-yl)acetamide (three-letter code: 9K4) (formula:  $C_{32}H_{52}F_3N_5O_9$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
8	A	1	Total	C	F	N	O	0	0
			43	28	3	4	8		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
8	B	1	Total	C	F	N	O	0	0
			49	32	3	5	9		
8	C	1	Total	C	F	N	O	0	0
			40	26	3	4	7		
8	D	1	Total	C	F	N	O	0	0
			49	32	3	5	9		

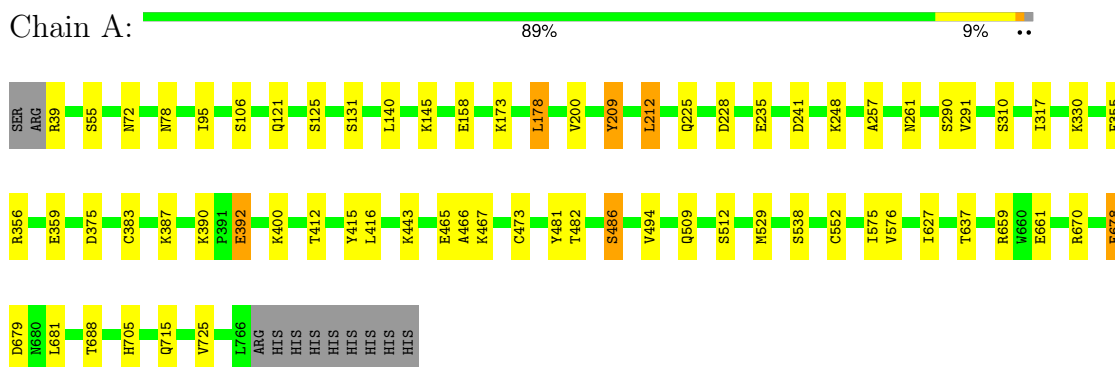
- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	36	Total	O	0	0
			36	36		
9	B	28	Total	O	0	0
			28	28		
9	C	24	Total	O	0	0
			24	24		
9	D	16	Total	O	0	0
			16	16		
9	L	2	Total	O	0	0
			2	2		
9	E	2	Total	O	0	0
			2	2		
9	H	5	Total	O	0	0
			5	5		
9	F	2	Total	O	0	0
			2	2		
9	G	1	Total	O	0	0
			1	1		
9	J	1	Total	O	0	0
			1	1		
9	K	1	Total	O	0	0
			1	1		

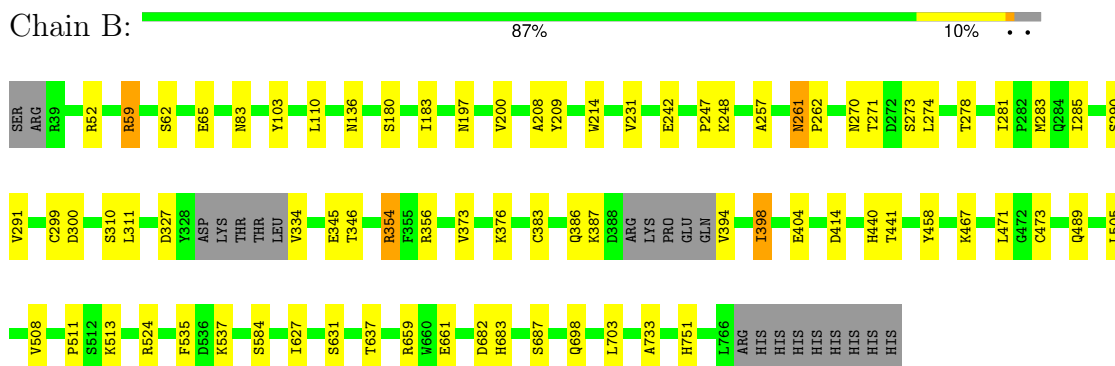
### 3 Residue-property plots

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 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.

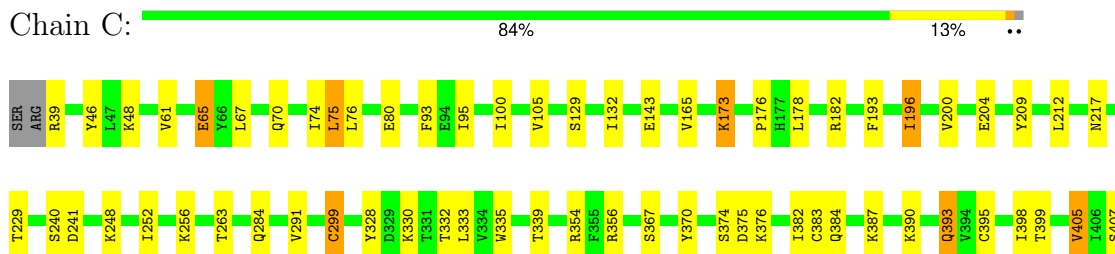
#### • Molecule 1: Dipeptidyl peptidase 4

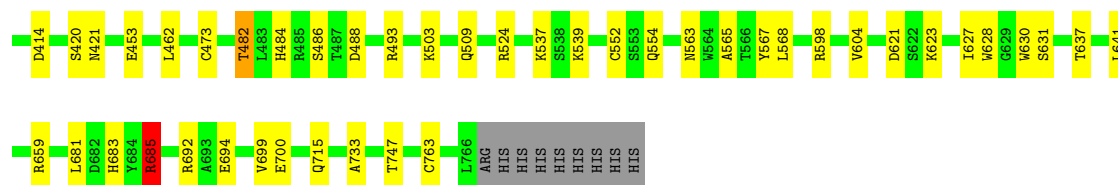


#### • Molecule 1: Dipeptidyl peptidase 4



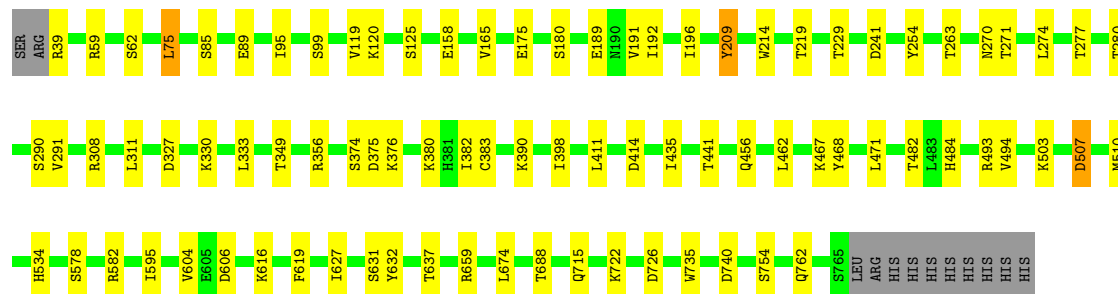
#### • Molecule 1: Dipeptidyl peptidase 4





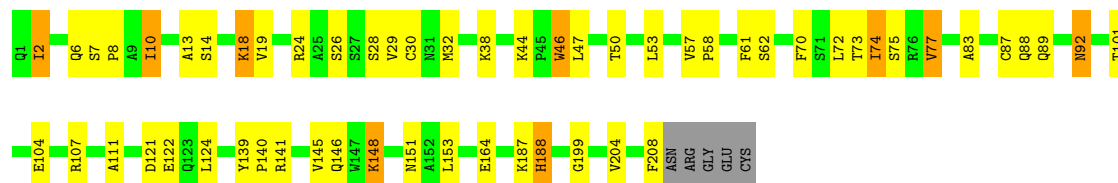
• Molecule 1: Dipeptidyl peptidase 4

Chain D: 87% 11%



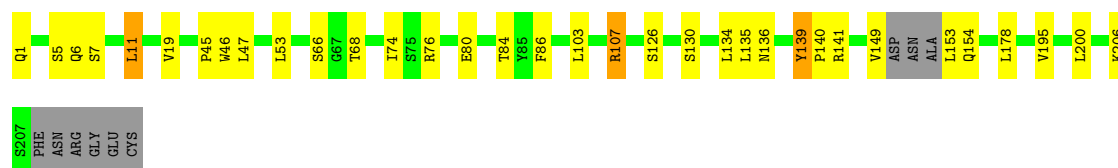
• Molecule 2: Fab light chain

Chain L: 71% 23%



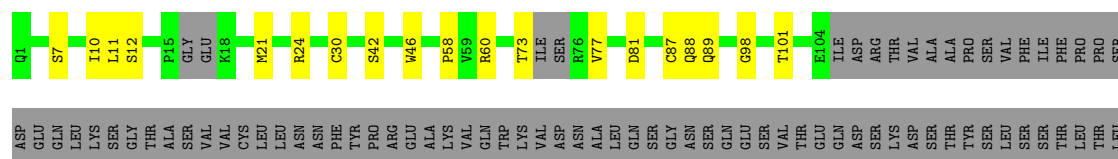
• Molecule 2: Fab light chain

Chain E: 80% 15%



• Molecule 2: Fab light chain

Chain G: 38% 9% 53%



GLU	F2	Q3	Q3	L4	Q5	Q6	S7	L11	K19	F29	T30	ASP	TYR	N33	L34	N35	N36	M37	K38	Q39	L45	I52	N61	Q62	K63	V72	D73	Q74	T78	A79	C96	R98	F99	R100	V101	F102	V105	T110	T113	VAL	SER	SER	ALA	SER	SER	THR	LYS	TRP
-----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----

PRO SER VAL PHE PRO LEU ALA PRO SER SER LYS SER THR SER GLY THR ALA ALA LEU GLY CYS LEU VAL LYS ASP TYR PHE PRO GLU PRO VAL THR VAL SER TRP ASN SER GLY ALA LEU THR SER GLY VAL HIS THR PHE PRO ALA VAL GLN SER SER GLY LEU TYR SER LEU

SER SER VAL PHE VAL PRO SER SER GLY THR GLN THR TYR ILE CYS ASN VAL ASN HIS LYS PRO SER ASN THR LYS VAL ASP LYS LYS VAL GLU PRO LYS

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	120.33Å 123.23Å 129.02Å 62.34° 77.21° 75.91°	Depositor
Resolution (Å)	30.00 – 2.80 30.00 – 2.80	Depositor EDS
% Data completeness (in resolution range)	88.5 (30.00-2.80) 88.6 (30.00-2.80)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.55 (at 2.80Å)	Xtriage
Refinement program	REFMAC 5.8.0155	Depositor
R, $R_{free}$	0.250 , 0.300 0.253 , 0.300	Depositor DCC
$R_{free}$ test set	6869 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	46.0	Xtriage
Anisotropy	0.582	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 26.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	33875	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.81% 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: NAG, 9K4, EDO, PEG, PG4

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	0.56	0/6098	0.73	0/8291
1	B	0.59	0/6022	0.76	0/8186
1	C	0.56	0/6098	0.73	1/8291 (0.0%)
1	D	0.55	0/6090	0.73	0/8280
2	E	0.54	0/1583	0.72	0/2152
2	G	0.57	0/767	0.73	0/1041
2	J	0.55	0/688	0.68	0/933
2	L	0.61	0/1617	0.75	0/2200
3	F	0.55	0/1679	0.70	0/2287
3	H	0.54	0/1679	0.73	1/2287 (0.0%)
3	I	0.57	0/920	0.75	1/1249 (0.1%)
3	K	0.61	0/879	0.82	2/1191 (0.2%)
All	All	0.56	0/34120	0.74	5/46388 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	685	ARG	NE-CZ-NH2	9.34	124.97	120.30
3	H	151	GLU	C-N-CD	-5.76	107.93	120.60
3	K	98	ARG	NE-CZ-NH1	5.52	123.06	120.30
3	K	98	ARG	NE-CZ-NH2	-5.39	117.61	120.30
3	I	100	ARG	NE-CZ-NH1	5.14	122.87	120.30

There are no chirality outliers.

There are no planarity outliers.



## 5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5929	0	5664	23	0
1	B	5852	0	5581	18	0
1	C	5929	0	5663	34	0
1	D	5921	0	5652	26	0
2	E	1546	0	1499	9	0
2	G	747	0	716	3	0
2	J	669	0	642	5	0
2	L	1578	0	1524	23	0
3	F	1639	0	1613	8	0
3	H	1639	0	1613	7	0
3	I	899	0	870	8	0
3	K	860	0	837	7	0
4	A	70	0	65	1	0
4	B	56	0	52	1	0
4	C	84	0	78	1	0
4	D	84	0	78	1	0
5	A	13	0	18	0	0
6	A	14	0	20	0	0
6	D	7	0	10	1	0
7	A	16	0	24	1	0
7	B	4	0	6	1	0
7	C	20	0	30	0	0
8	A	43	0	0	0	0
8	B	49	0	0	1	0
8	C	40	0	0	2	0
8	D	49	0	0	0	0
9	A	36	0	0	0	0
9	B	28	0	0	0	0
9	C	24	0	0	0	0
9	D	16	0	0	0	0
9	E	2	0	0	0	0
9	F	2	0	0	0	0
9	G	1	0	0	0	0
9	H	5	0	0	0	0
9	J	1	0	0	0	0
9	K	1	0	0	0	0
9	L	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	33875	0	32255	161	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 2.

The worst 5 of 161 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)
2:E:134:LEU:HD22	3:F:184:VAL:HG11	1.61	0.80
1:D:219:THR:HG22	1:D:270:ASN:HD21	1.57	0.69
2:J:60:PHE:CE1	2:J:73:ILE:HD13	2.31	0.66
2:L:47:LEU:HD11	2:L:72:LEU:HD13	1.77	0.65
3:K:39:GLN:HB2	3:K:45:LEU:HD23	1.79	0.64

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	726/739 (98%)	690 (95%)	36 (5%)	0	100	100
1	B	713/739 (96%)	682 (96%)	31 (4%)	0	100	100
1	C	726/739 (98%)	686 (94%)	40 (6%)	0	100	100
1	D	725/739 (98%)	676 (93%)	49 (7%)	0	100	100
2	E	200/213 (94%)	190 (95%)	10 (5%)	0	100	100
2	G	94/213 (44%)	89 (95%)	4 (4%)	1 (1%)	12	37
2	J	83/213 (39%)	80 (96%)	2 (2%)	1 (1%)	11	34
2	L	206/213 (97%)	196 (95%)	10 (5%)	0	100	100
3	F	215/217 (99%)	204 (95%)	10 (5%)	1 (0%)	25	56

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	H	215/217 (99%)	207 (96%)	7 (3%)	1 (0%)	25	56
3	I	113/217 (52%)	105 (93%)	8 (7%)	0	100	100
3	K	106/217 (49%)	100 (94%)	5 (5%)	1 (1%)	14	42
All	All	4122/4676 (88%)	3905 (95%)	212 (5%)	5 (0%)	48	77

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	H	152	PRO
3	F	101	VAL
3	K	101	VAL
2	J	76	VAL
2	G	58	PRO

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	649/660 (98%)	608 (94%)	41 (6%)	15	42
1	B	640/660 (97%)	589 (92%)	51 (8%)	10	30
1	C	649/660 (98%)	584 (90%)	65 (10%)	6	20
1	D	648/660 (98%)	595 (92%)	53 (8%)	9	29
2	E	178/185 (96%)	154 (86%)	24 (14%)	3	10
2	G	84/185 (45%)	70 (83%)	14 (17%)	2	6
2	J	76/185 (41%)	55 (72%)	21 (28%)	0	1
2	L	181/185 (98%)	149 (82%)	32 (18%)	1	5
3	F	189/189 (100%)	173 (92%)	16 (8%)	8	27
3	H	189/189 (100%)	160 (85%)	29 (15%)	2	7
3	I	101/189 (53%)	85 (84%)	16 (16%)	2	7
3	K	96/189 (51%)	77 (80%)	19 (20%)	1	3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	3680/4136 (89%)	3299 (90%)	381 (10%)	5 18

5 of 381 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	L	107	ARG
3	H	141	LEU
2	L	187	LYS
2	E	149	VAL
3	F	65	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
2	G	88	GLN
3	I	62	GLN
2	J	51	ASN
1	D	49	ASN
1	C	762	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

39 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	NAG	D	802	1	14,14,15	0.47	0	17,19,21	1.63	3 (17%)
4	NAG	C	803	1	14,14,15	0.47	0	17,19,21	2.14	1 (5%)
4	NAG	A	804	1	14,14,15	0.37	0	17,19,21	1.38	2 (11%)
4	NAG	C	801	1	14,14,15	0.58	0	17,19,21	1.96	2 (11%)
4	NAG	D	804	1	14,14,15	0.59	0	17,19,21	1.33	2 (11%)
4	NAG	B	803	1	14,14,15	0.43	0	17,19,21	1.78	1 (5%)
4	NAG	B	804	1	14,14,15	0.68	1 (7%)	17,19,21	1.50	1 (5%)
6	PEG	D	807	-	6,6,6	0.71	0	5,5,5	0.59	0
4	NAG	D	801	1	14,14,15	0.48	0	17,19,21	1.47	1 (5%)
4	NAG	A	805	1	14,14,15	0.66	0	17,19,21	1.33	2 (11%)
7	EDO	C	807	-	3,3,3	0.42	0	2,2,2	0.31	0
4	NAG	D	805	1	14,14,15	0.35	0	17,19,21	1.71	3 (17%)
7	EDO	A	811	-	3,3,3	0.40	0	2,2,2	0.01	0
4	NAG	C	805	1	14,14,15	0.52	0	17,19,21	2.26	5 (29%)
7	EDO	C	809	-	3,3,3	0.53	0	2,2,2	0.14	0
4	NAG	C	806	1	14,14,15	0.44	0	17,19,21	1.56	3 (17%)
4	NAG	B	802	1	14,14,15	0.74	0	17,19,21	1.68	4 (23%)
4	NAG	A	803	1	14,14,15	0.54	0	17,19,21	1.94	5 (29%)
7	EDO	C	808	-	3,3,3	0.44	0	2,2,2	0.12	0
8	9K4	D	808	-	50,50,50	2.56	12 (24%)	57,61,61	1.39	8 (14%)
8	9K4	C	812	-	41,41,50	2.51	7 (17%)	47,51,61	1.11	4 (8%)
4	NAG	A	801	1	14,14,15	0.79	1 (7%)	17,19,21	1.82	5 (29%)
7	EDO	C	810	-	3,3,3	0.54	0	2,2,2	0.09	0
4	NAG	B	801	1	14,14,15	0.65	0	17,19,21	2.11	5 (29%)
8	9K4	B	806	-	50,50,50	2.50	11 (22%)	57,61,61	1.29	6 (10%)
7	EDO	A	809	-	3,3,3	0.71	0	2,2,2	0.42	0
7	EDO	A	810	-	3,3,3	0.42	0	2,2,2	0.34	0
8	9K4	A	813	-	44,44,50	2.42	7 (15%)	50,54,61	1.24	4 (8%)
4	NAG	C	804	1	14,14,15	0.33	0	17,19,21	1.53	3 (17%)
4	NAG	D	803	1	14,14,15	0.56	0	17,19,21	1.33	3 (17%)
7	EDO	A	812	-	3,3,3	0.58	0	2,2,2	0.27	0
7	EDO	B	805	-	3,3,3	0.56	0	2,2,2	0.27	0
5	PG4	A	806	-	12,12,12	0.59	0	11,11,11	0.41	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
7	EDO	C	811	-	3,3,3	0.43	0	2,2,2	0.16	0
4	NAG	C	802	1	14,14,15	0.45	0	17,19,21	1.43	2 (11%)
4	NAG	A	802	1	14,14,15	0.50	0	17,19,21	1.26	2 (11%)
6	PEG	A	808	-	6,6,6	0.56	0	5,5,5	0.42	0
6	PEG	A	807	-	6,6,6	0.46	0	5,5,5	0.34	0
4	NAG	D	806	1	14,14,15	0.34	0	17,19,21	1.35	1 (5%)

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	NAG	D	802	1	-	2/6/23/26	0/1/1/1
4	NAG	C	803	1	-	2/6/23/26	0/1/1/1
4	NAG	A	804	1	-	1/6/23/26	0/1/1/1
4	NAG	C	801	1	-	2/6/23/26	0/1/1/1
4	NAG	D	804	1	-	1/6/23/26	0/1/1/1
4	NAG	B	803	1	-	2/6/23/26	0/1/1/1
4	NAG	B	804	1	-	2/6/23/26	0/1/1/1
6	PEG	D	807	-	-	1/4/4/4	-
4	NAG	D	801	1	-	2/6/23/26	0/1/1/1
4	NAG	A	805	1	-	2/6/23/26	0/1/1/1
7	EDO	C	807	-	-	1/1/1/1	-
4	NAG	D	805	1	-	2/6/23/26	0/1/1/1
7	EDO	A	811	-	-	0/1/1/1	-
4	NAG	C	805	1	-	5/6/23/26	0/1/1/1
7	EDO	C	809	-	-	1/1/1/1	-
4	NAG	C	806	1	-	2/6/23/26	0/1/1/1
4	NAG	B	802	1	-	1/6/23/26	0/1/1/1
4	NAG	A	803	1	-	2/6/23/26	0/1/1/1
7	EDO	C	808	-	-	0/1/1/1	-
8	9K4	D	808	-	-	18/41/51/51	0/2/2/2
8	9K4	C	812	-	-	11/32/42/51	0/2/2/2
4	NAG	A	801	1	-	4/6/23/26	0/1/1/1
7	EDO	C	810	-	-	0/1/1/1	-
4	NAG	B	801	1	-	3/6/23/26	0/1/1/1
8	9K4	B	806	-	-	10/41/51/51	0/2/2/2

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	EDO	A	809	-	-	1/1/1/1	-
7	EDO	A	810	-	-	1/1/1/1	-
8	9K4	A	813	-	-	13/35/45/51	0/2/2/2
4	NAG	C	804	1	-	2/6/23/26	0/1/1/1
4	NAG	D	803	1	-	3/6/23/26	0/1/1/1
7	EDO	A	812	-	-	1/1/1/1	-
7	EDO	B	805	-	-	1/1/1/1	-
5	PG4	A	806	-	-	8/10/10/10	-
7	EDO	C	811	-	-	1/1/1/1	-
4	NAG	C	802	1	-	2/6/23/26	0/1/1/1
4	NAG	A	802	1	-	1/6/23/26	0/1/1/1
6	PEG	A	808	-	-	4/4/4/4	-
6	PEG	A	807	-	-	3/4/4/4	-
4	NAG	D	806	1	-	2/6/23/26	0/1/1/1

The worst 5 of 39 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	D	808	9K4	C18-N17	-9.99	1.25	1.47
8	C	812	9K4	C18-N17	-9.46	1.26	1.47
8	A	813	9K4	C18-N17	-9.22	1.27	1.47
8	B	806	9K4	C18-N17	-9.10	1.27	1.47
8	A	813	9K4	C13-N14	7.80	1.50	1.35

The worst 5 of 78 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	803	NAG	C1-O5-C5	8.46	123.53	112.19
4	B	803	NAG	C1-O5-C5	6.42	120.79	112.19
4	C	801	NAG	C1-O5-C5	6.21	120.51	112.19
4	C	805	NAG	C8-C7-N2	5.73	125.62	116.12
4	D	805	NAG	O5-C1-C2	-5.42	102.91	111.29

There are no chirality outliers.

5 of 120 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
8	D	808	9K4	C08-C09-C10-C12
8	D	808	9K4	C09-C10-C12-C13
8	D	808	9K4	N11-C10-C12-C13

*Continued on next page...*

*Continued from previous page...*

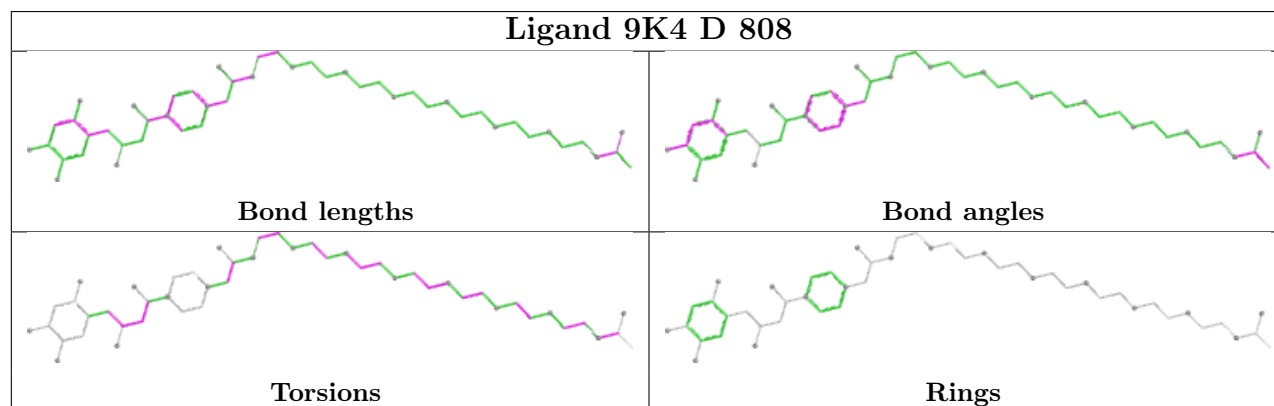
Mol	Chain	Res	Type	Atoms
4	B	803	NAG	O5-C5-C6-O6
4	A	805	NAG	O5-C5-C6-O6

There are no ring outliers.

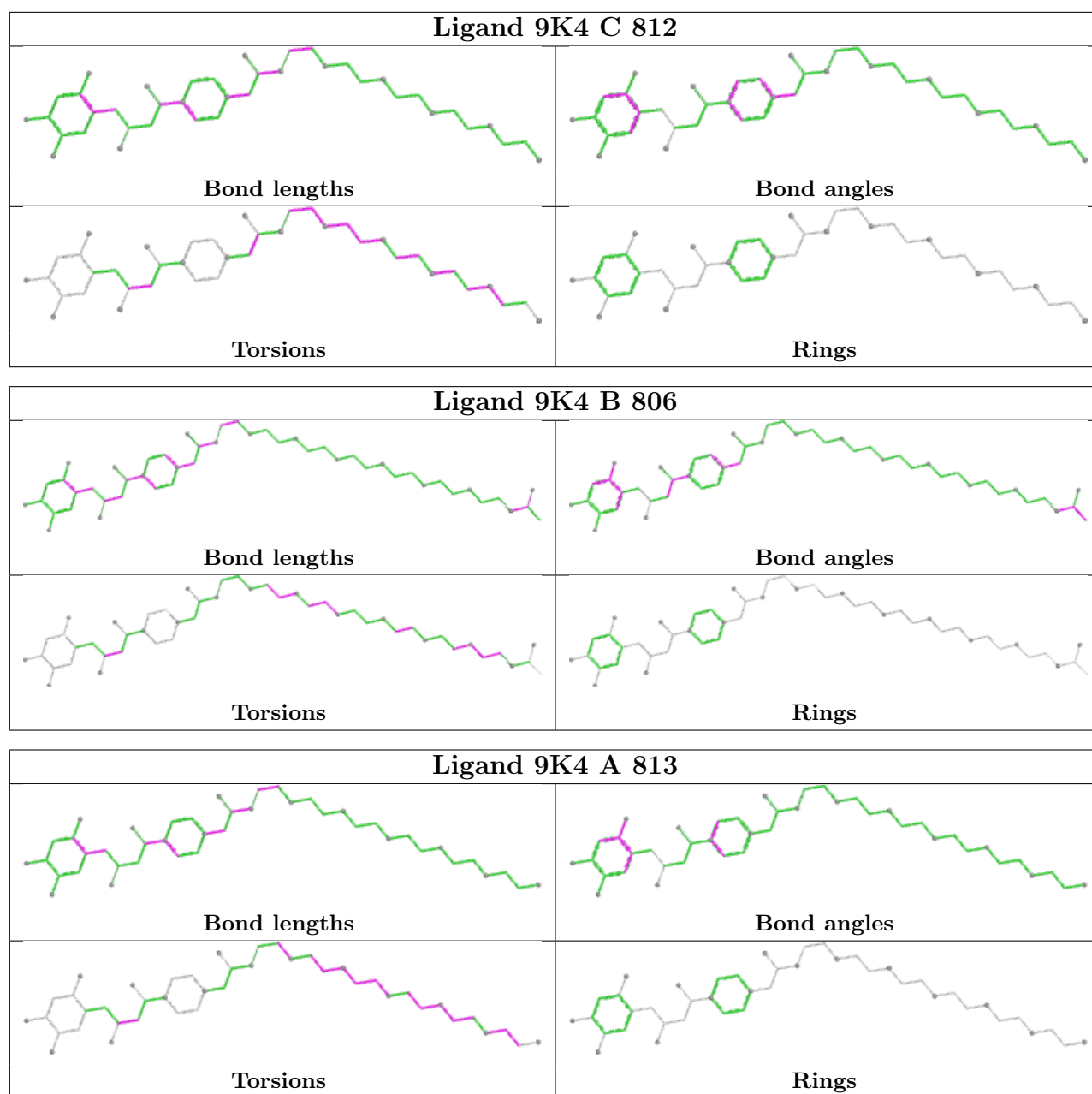
9 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	D	807	PEG	1	0
7	A	811	EDO	1	0
4	C	805	NAG	1	0
8	C	812	9K4	2	0
4	A	801	NAG	1	0
4	B	801	NAG	1	0
8	B	806	9K4	1	0
7	B	805	EDO	1	0
4	D	806	NAG	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.







## 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	728/739 (98%)	-1.88	0 100 100	19, 40, 62, 100	0
1	B	718/739 (97%)	-1.84	0 100 100	17, 43, 70, 86	1 (0%)
1	C	728/739 (98%)	-1.80	0 100 100	27, 51, 72, 98	0
1	D	727/739 (98%)	-1.81	0 100 100	28, 50, 74, 103	0
2	E	204/213 (95%)	-1.63	0 100 100	35, 65, 88, 102	0
2	G	100/213 (46%)	-1.64	0 100 100	32, 72, 102, 116	0
2	J	89/213 (41%)	-1.58	0 100 100	37, 70, 104, 126	0
2	L	208/213 (97%)	-1.65	0 100 100	35, 62, 85, 91	0
3	F	217/217 (100%)	-1.82	0 100 100	30, 47, 67, 78	0
3	H	217/217 (100%)	-1.85	0 100 100	25, 44, 62, 77	0
3	I	115/217 (52%)	-1.70	0 100 100	50, 70, 86, 98	0
3	K	110/217 (50%)	-1.56	0 100 100	47, 77, 101, 110	0
All	All	4161/4676 (88%)	-1.79	0 100 100	17, 49, 81, 126	1 (0%)

There are no RSRZ outliers to report.

### 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 monosaccharides in this entry.

## 6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	NAG	D	802	14/15	0.95	0.05	92,105,108,109	0
4	NAG	D	805	14/15	0.97	0.04	79,84,89,90	0
7	EDO	C	809	4/4	0.97	0.12	52,57,60,62	0
4	NAG	A	805	14/15	0.98	0.04	70,74,77,77	0
4	NAG	B	801	14/15	0.98	0.04	77,81,84,86	0
4	NAG	B	802	14/15	0.98	0.04	52,53,58,58	0
4	NAG	B	804	14/15	0.98	0.03	60,66,70,71	0
4	NAG	C	801	14/15	0.98	0.04	63,70,80,80	0
4	NAG	C	803	14/15	0.98	0.04	60,77,83,85	0
4	NAG	C	805	14/15	0.98	0.04	63,67,70,73	0
4	NAG	A	801	14/15	0.98	0.04	67,71,73,73	0
4	NAG	D	803	14/15	0.98	0.04	69,77,78,78	0
4	NAG	D	804	14/15	0.98	0.04	90,97,104,104	0
4	NAG	A	802	14/15	0.98	0.04	53,57,64,66	0
6	PEG	A	807	7/7	0.98	0.07	52,56,60,61	0
6	PEG	A	808	7/7	0.98	0.03	49,55,66,67	0
7	EDO	C	807	4/4	0.98	0.05	68,72,74,77	0
4	NAG	A	804	14/15	0.98	0.04	61,66,73,78	0
7	EDO	C	810	4/4	0.98	0.06	43,47,48,49	0
7	EDO	C	811	4/4	0.98	0.04	52,54,55,56	0
8	9K4	C	812	40/49	0.98	0.05	49,80,98,103	0
5	PG4	A	806	13/13	0.99	0.03	54,57,60,60	0
4	NAG	C	806	14/15	0.99	0.03	87,92,95,96	0
4	NAG	D	801	14/15	0.99	0.03	46,53,60,65	0
6	PEG	D	807	7/7	0.99	0.04	48,51,53,53	0
7	EDO	A	809	4/4	0.99	0.03	28,28,28,28	0
7	EDO	A	810	4/4	0.99	0.04	53,54,54,56	0
7	EDO	A	811	4/4	0.99	0.03	40,41,41,42	0
7	EDO	B	805	4/4	0.99	0.04	60,62,63,65	0
4	NAG	C	802	14/15	0.99	0.03	53,64,69,71	0
7	EDO	C	808	4/4	0.99	0.07	57,57,58,59	0
4	NAG	A	803	14/15	0.99	0.03	32,34,35,36	0
4	NAG	C	804	14/15	0.99	0.03	38,39,40,44	0
4	NAG	B	803	14/15	0.99	0.03	42,45,50,55	0
8	9K4	A	813	43/49	0.99	0.04	33,56,87,90	0
8	9K4	B	806	49/49	0.99	0.05	36,83,108,111	0
4	NAG	D	806	14/15	0.99	0.04	53,55,61,70	0

*Continued on next page...*

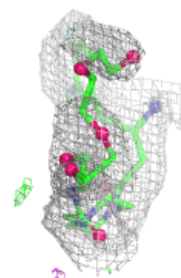
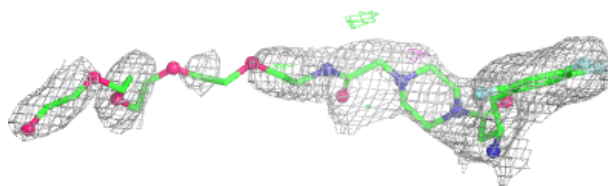
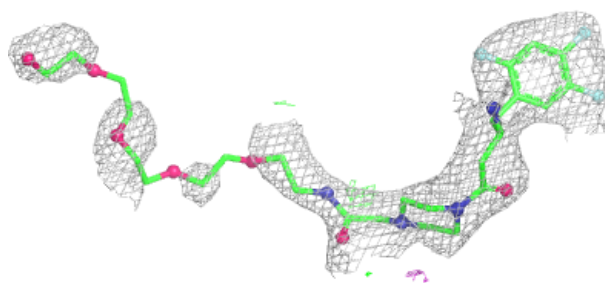
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
8	9K4	D	808	49/49	0.99	0.05	55,91,95,98	0
7	EDO	A	812	4/4	1.00	0.10	37,39,40,40	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

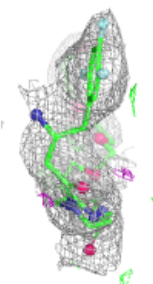
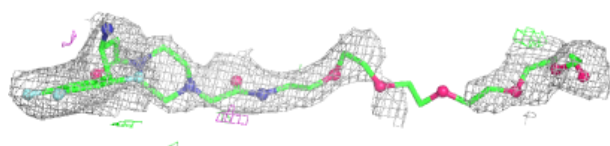
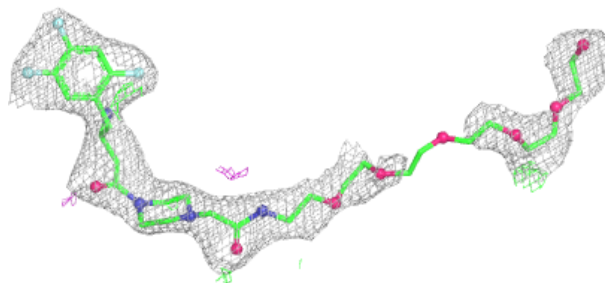
**Electron density around 9K4 C 812:**

2mF<sub>o</sub>-DF<sub>c</sub> (at 0.7 rmsd) in gray  
mF<sub>o</sub>-DF<sub>c</sub> (at 3 rmsd) in purple (negative)  
and green (positive)

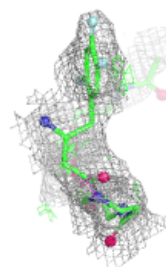
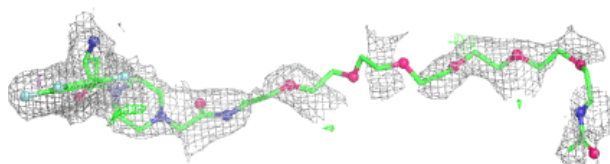
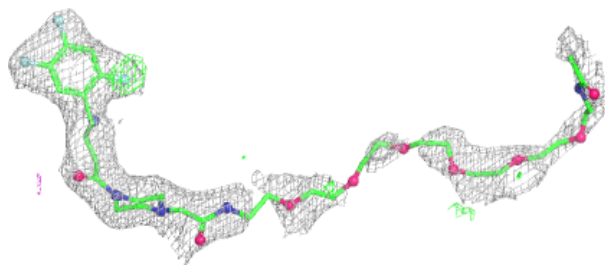


**Electron density around 9K4 A 813:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

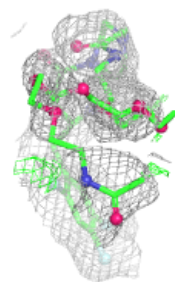
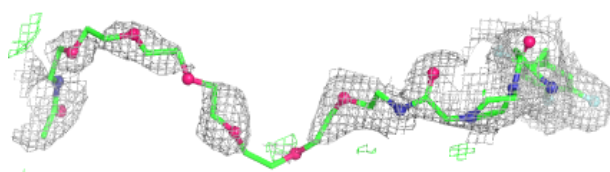
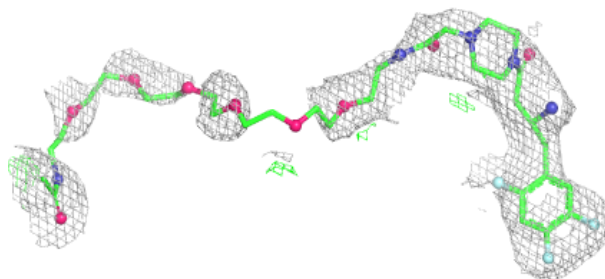
**Electron density around 9K4 B 806:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around 9K4 D 808:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



## 6.5 Other polymers [i](#)

There are no such residues in this entry.