wwPDB X-ray Structure Validation Summary Report (i)

Oct 10, 2021 – 10:42 AM EDT

PDB ID	:	3GQL
Title	:	Crystal Structure of activated receptor tyrosine kinase in complex with substrates
Authors	:	Bae, J.H.; Lew, E.D.; Yuzawa, S.; Tome, F.; Lax, I.; Schlessinger, J.
Deposited on	:	2009-03-24
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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

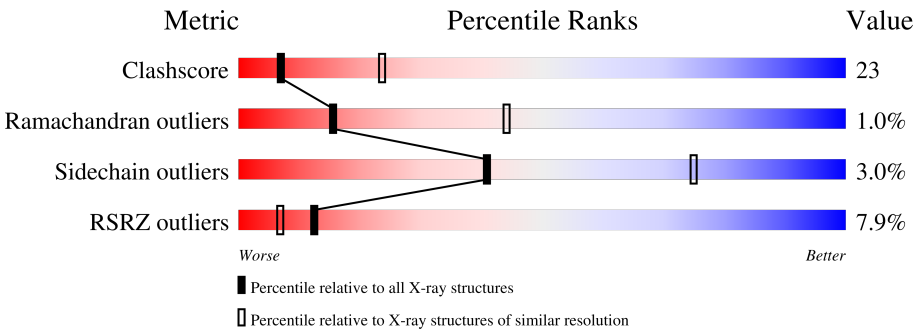
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 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))
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (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 >=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 <=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	326	5% 54% 31% • 12%
1	B	326	6% 54% 28% • 15%
1	C	326	9% 51% 32% • 16%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6786 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 Basic fibroblast growth factor receptor 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	287	2259	1433	390	418	18	0	0	0
1	B	276	2143	1360	365	401	17	0	0	0
1	C	275	2116	1343	364	393	16	0	0	0

There are 30 discrepancies between the modelled and reference sequences:

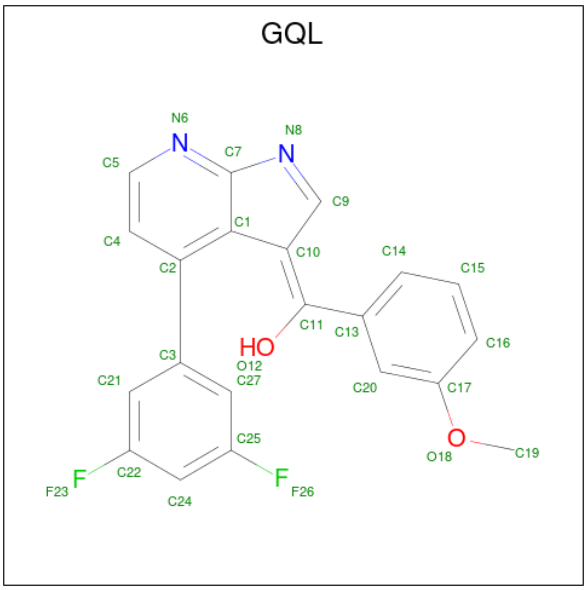
Chain	Residue	Modelled	Actual	Comment	Reference
A	449	MET	-	expression tag	UNP P11362
A	450	GLY	-	expression tag	UNP P11362
A	451	HIS	-	expression tag	UNP P11362
A	452	HIS	-	expression tag	UNP P11362
A	453	HIS	-	expression tag	UNP P11362
A	454	HIS	-	expression tag	UNP P11362
A	455	HIS	-	expression tag	UNP P11362
A	456	HIS	-	expression tag	UNP P11362
A	457	MET	-	expression tag	UNP P11362
A	488	ALA	CYS	engineered mutation	UNP P11362
B	449	MET	-	expression tag	UNP P11362
B	450	GLY	-	expression tag	UNP P11362
B	451	HIS	-	expression tag	UNP P11362
B	452	HIS	-	expression tag	UNP P11362
B	453	HIS	-	expression tag	UNP P11362
B	454	HIS	-	expression tag	UNP P11362
B	455	HIS	-	expression tag	UNP P11362
B	456	HIS	-	expression tag	UNP P11362
B	457	MET	-	expression tag	UNP P11362
B	488	ALA	CYS	engineered mutation	UNP P11362
C	449	MET	-	expression tag	UNP P11362
C	450	GLY	-	expression tag	UNP P11362
C	451	HIS	-	expression tag	UNP P11362

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	452	HIS	-	expression tag	UNP P11362
C	453	HIS	-	expression tag	UNP P11362
C	454	HIS	-	expression tag	UNP P11362
C	455	HIS	-	expression tag	UNP P11362
C	456	HIS	-	expression tag	UNP P11362
C	457	MET	-	expression tag	UNP P11362
C	488	ALA	CYS	engineered mutation	UNP P11362

- Molecule 2 is (E)-[4-(3,5-difluorophenyl)-3H-pyrrolo[2,3-b]pyridin-3-ylidene](3-methoxyphenyl)methanol (three-letter code: GQL) (formula: $C_{21}H_{14}F_2N_2O_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	F	N	O		
2	A	1	27	21	2	2	2	0	0
			Total	C	F	N	O		
2	B	1	27	21	2	2	2	0	0
			Total	C	F	N	O		
2	C	1	27	21	2	2	2	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	92	92	92	0	0
			Total	O		
3	B	45	45	45	0	0

Continued on next page...

Continued from previous page...

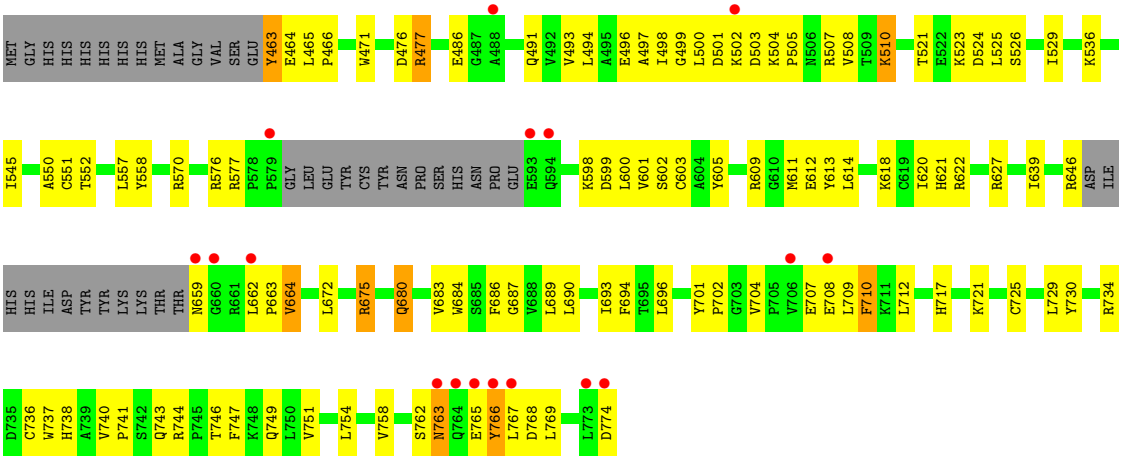
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	C	50	50	50	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

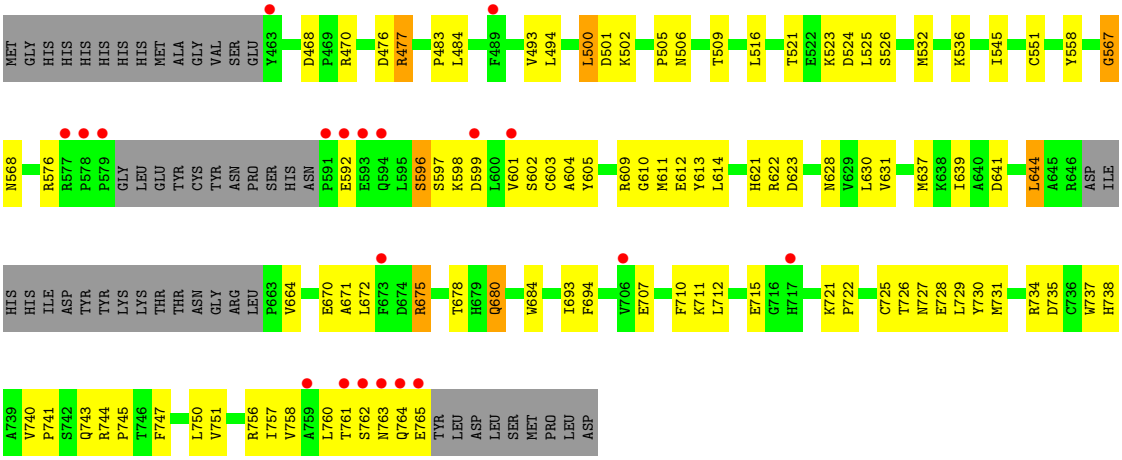
• Molecule 1: Basic fibroblast growth factor receptor 1

Chain A:



• Molecule 1: Basic fibroblast growth factor receptor 1

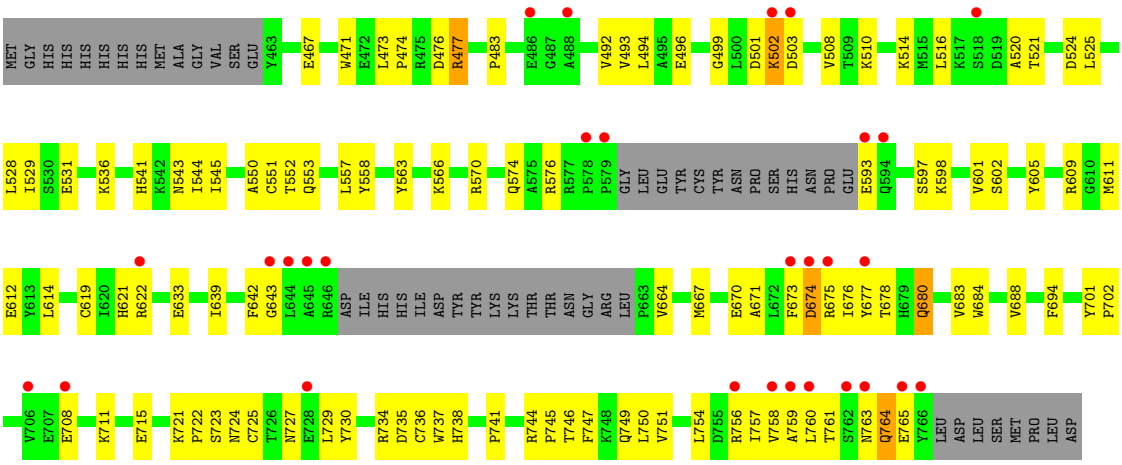
Chain B:



• Molecule 1: Basic fibroblast growth factor receptor 1

Chain C:

9%
51%
32%
16%


MET GLY HIS HIS HIS HIS HIS MET ALA GLY VAL SER GLU Y463 E467 W471 E472 L473 P474 R475 D476 R477 P483 E486 G487 A488 V492 V493 L494 A495 E496 G499 L500 D501 K502 D503 V508 I509 K510 K514 M515 L516 K517 S518 D519 A520 T521 D524 L525
L528 I529 S530 E531 K536 H541 K542 N543 I544 I545 A550 C551 T552 Q553 L557 Y558 Y563 K566 R570 Q574 A575 R576 R577 P578 P579 GLY LEU GLU TYR CYS TYR ASN PRO SER HIS ASN PRO GLU E593 Q594 S597 K598 V601 S602 Y605 R609 G610 M611
E612 Y613 L614 C619 I620 H621 R622 E633 I639 F642 G643 L644 A645 R646 ASP ILE HIS HIS ILE ASP TYR TYR LYS LYS THR THR ASN GLY ARG LEU P663 V664 M667 E670 A671 L672 F673 D674 R675 I676 Y677 T678 H679 Q680 V683 W684 V688 F694 Y701 P702
V706 E707 E708 K711 E715 K721 P722 S723 N724 C725 T726 N727 E728 L729 Y730 R734 D735 C736 W737 H738 P741 R744 P745 T746 F747 K748 Q749 L750 V751 L754 D755 R756 I757 V758 A759 L760 T761 S762 N763 Q764 E765 Y766 LEU ASP LEU SER MET PRO LEU ASP

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	194.17Å 78.27Å 98.46Å 90.00° 110.55° 90.00°	Depositor
Resolution (Å)	36.01 – 2.80 36.02 – 2.80	Depositor EDS
% Data completeness (in resolution range)	89.8 (36.01-2.80) 92.8 (36.02-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.22 (at 2.81Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.252 , 0.289 0.253 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	56.7	Xtriage
Anisotropy	0.831	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 51.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	6786	wwPDB-VP
Average B, all atoms (Å^2)	76.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.54% 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, 0.333 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: GQL

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	0.61	1/2303 (0.0%)	0.67	0/3113
1	B	0.47	0/2186	0.60	0/2961
1	C	0.44	0/2158	0.59	0/2925
All	All	0.51	1/6647 (0.0%)	0.62	0/8999

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	463	TYR	CB-CG	7.02	1.62	1.51

There are no bond angle outliers.

There are no chirality outliers.

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	2259	0	2263	104	0
1	B	2143	0	2107	98	0
1	C	2116	0	2068	96	0
2	A	27	0	13	2	0
2	B	27	0	13	6	0
2	C	27	0	13	6	0
3	A	92	0	0	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	45	0	0	6	0
3	C	50	0	0	6	0
All	All	6786	0	6477	301	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 23.

The worst 5 of 301 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:C:598:LYS:HB2	1:C:761:THR:HG21	1.30	1.10
1:A:504:LYS:HB3	1:A:507:ARG:HD3	1.36	1.05
1:A:763:ASN:N	1:A:763:ASN:HD22	1.63	0.96
1:C:676:ILE:HG13	3:C:62:HOH:O	1.67	0.93
1:A:477:ARG:HA	1:A:477:ARG:HH11	1.33	0.91

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	281/326 (86%)	253 (90%)	26 (9%)	2 (1%)	22	53
1	B	270/326 (83%)	248 (92%)	19 (7%)	3 (1%)	14	41
1	C	269/326 (82%)	246 (91%)	20 (7%)	3 (1%)	14	41
All	All	820/978 (84%)	747 (91%)	65 (8%)	8 (1%)	15	44

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	765	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	596	SER
1	C	764	GLN
1	B	710	PHE
1	C	499	GLY

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	245/286 (86%)	236 (96%)	9 (4%)	34	68
1	B	228/286 (80%)	221 (97%)	7 (3%)	40	74
1	C	221/286 (77%)	216 (98%)	5 (2%)	50	82
All	All	694/858 (81%)	673 (97%)	21 (3%)	41	75

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	680	GLN
1	C	502	LYS
1	C	680	GLN
1	C	673	PHE
1	C	477	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	738	HIS
1	C	724	ASN
1	B	749	GLN
1	C	680	GLN
1	B	738	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

3 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	GQL	A	1	-	30,30,30	2.43	7 (23%)	37,43,43	2.45	14 (37%)
2	GQL	C	1	-	30,30,30	2.43	7 (23%)	37,43,43	2.43	12 (32%)
2	GQL	B	1	-	30,30,30	2.43	7 (23%)	37,43,43	2.39	12 (32%)

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	GQL	A	1	-	-	6/14/23/23	0/4/4/4
2	GQL	C	1	-	-	6/14/23/23	0/4/4/4
2	GQL	B	1	-	-	10/14/23/23	0/4/4/4

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1	GQL	C10-C11	6.55	1.54	1.39
2	B	1	GQL	C10-C11	6.35	1.53	1.39
2	A	1	GQL	C10-C11	6.34	1.53	1.39
2	B	1	GQL	C9-C10	-5.93	1.33	1.40
2	A	1	GQL	C9-C10	-5.86	1.33	1.40

The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	1	GQL	C7-N8-C9	10.57	109.86	103.76
2	A	1	GQL	C7-N8-C9	10.43	109.78	103.76
2	B	1	GQL	C7-N8-C9	10.08	109.58	103.76
2	A	1	GQL	O12-C11-C10	-3.97	120.85	123.40
2	C	1	GQL	O12-C11-C10	-3.86	120.92	123.40

There are no chirality outliers.

5 of 22 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1	GQL	C10-C11-C13-C14
2	C	1	GQL	C10-C11-C13-C14
2	A	1	GQL	C20-C17-O18-C19
2	A	1	GQL	C16-C17-O18-C19
2	B	1	GQL	C20-C17-O18-C19

There are no ring outliers.

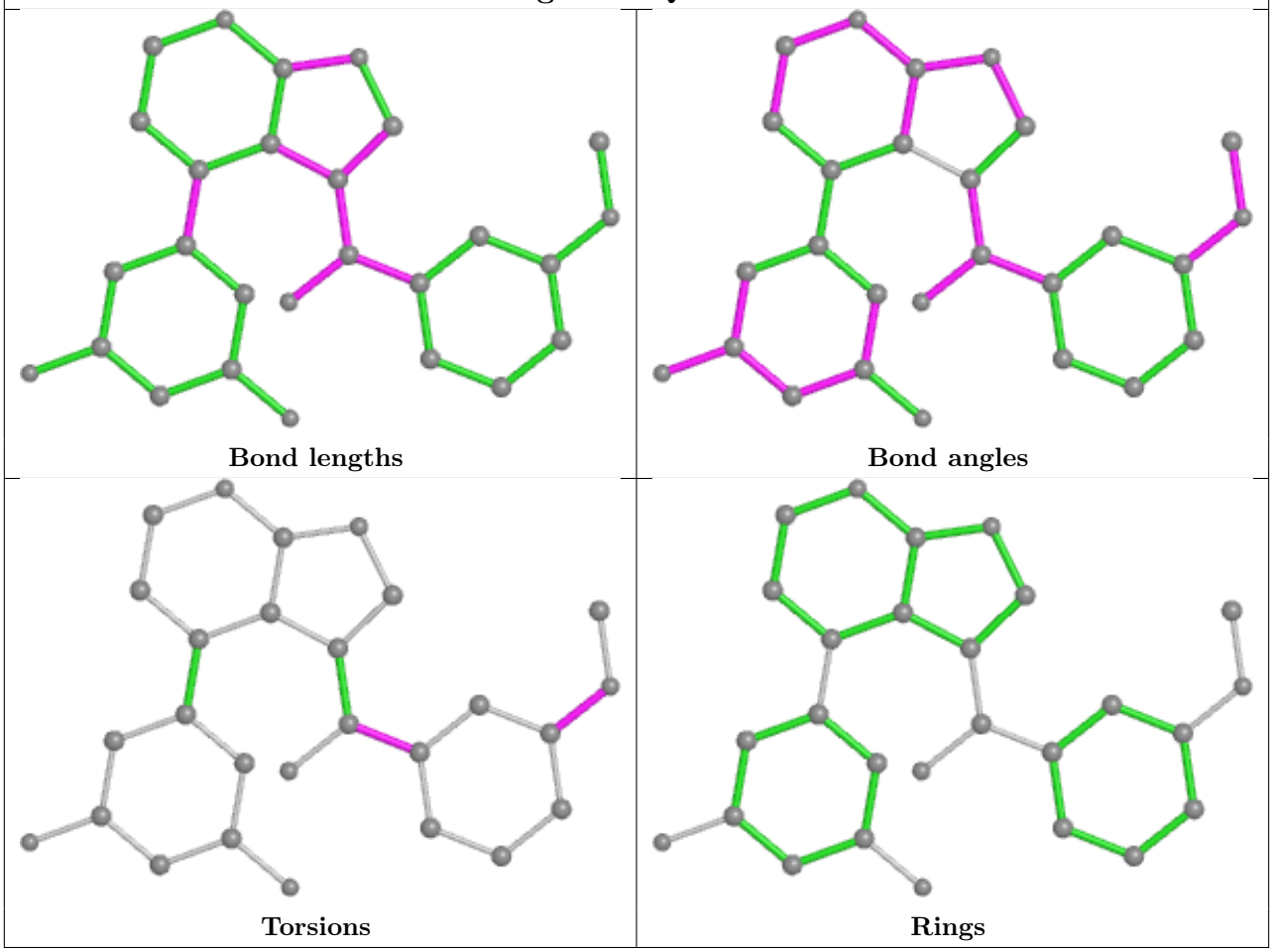
3 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1	GQL	2	0
2	C	1	GQL	6	0
2	B	1	GQL	6	0

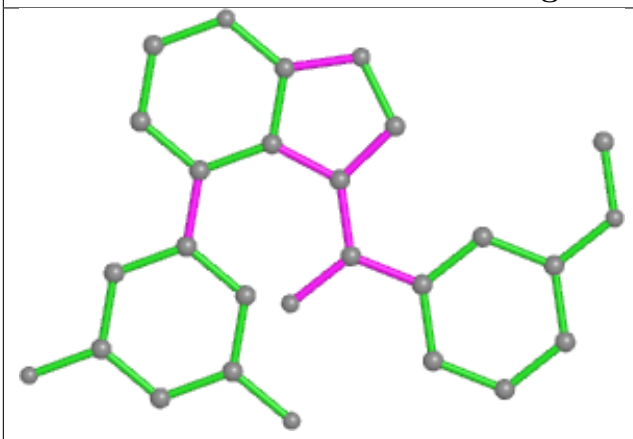
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 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.

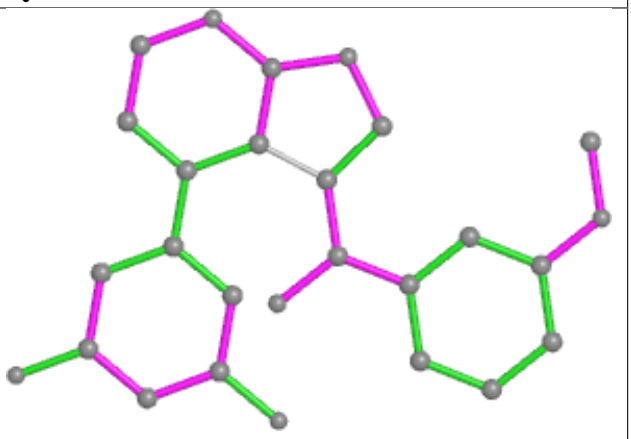
Ligand GQL A 1



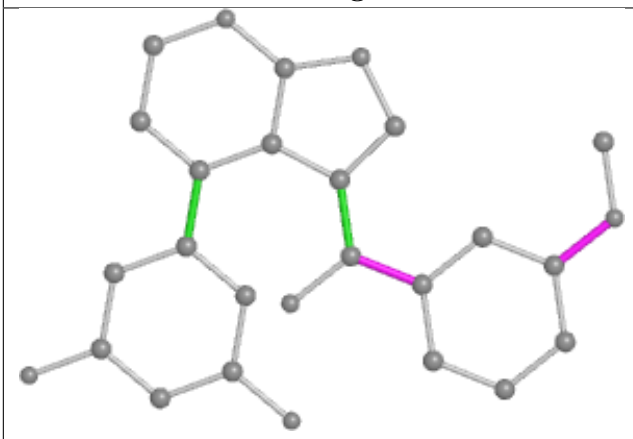
Ligand GQL C 1



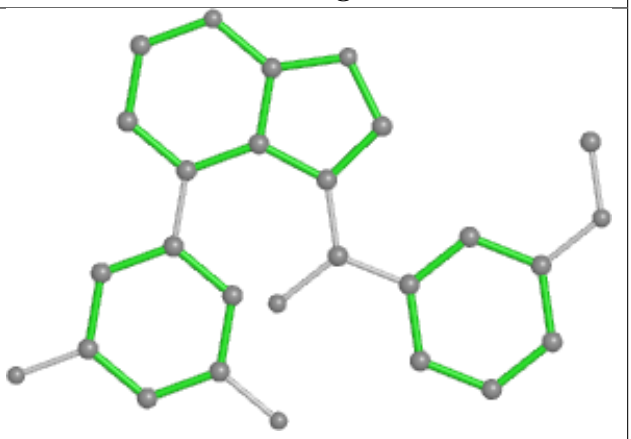
Bond lengths



Bond angles

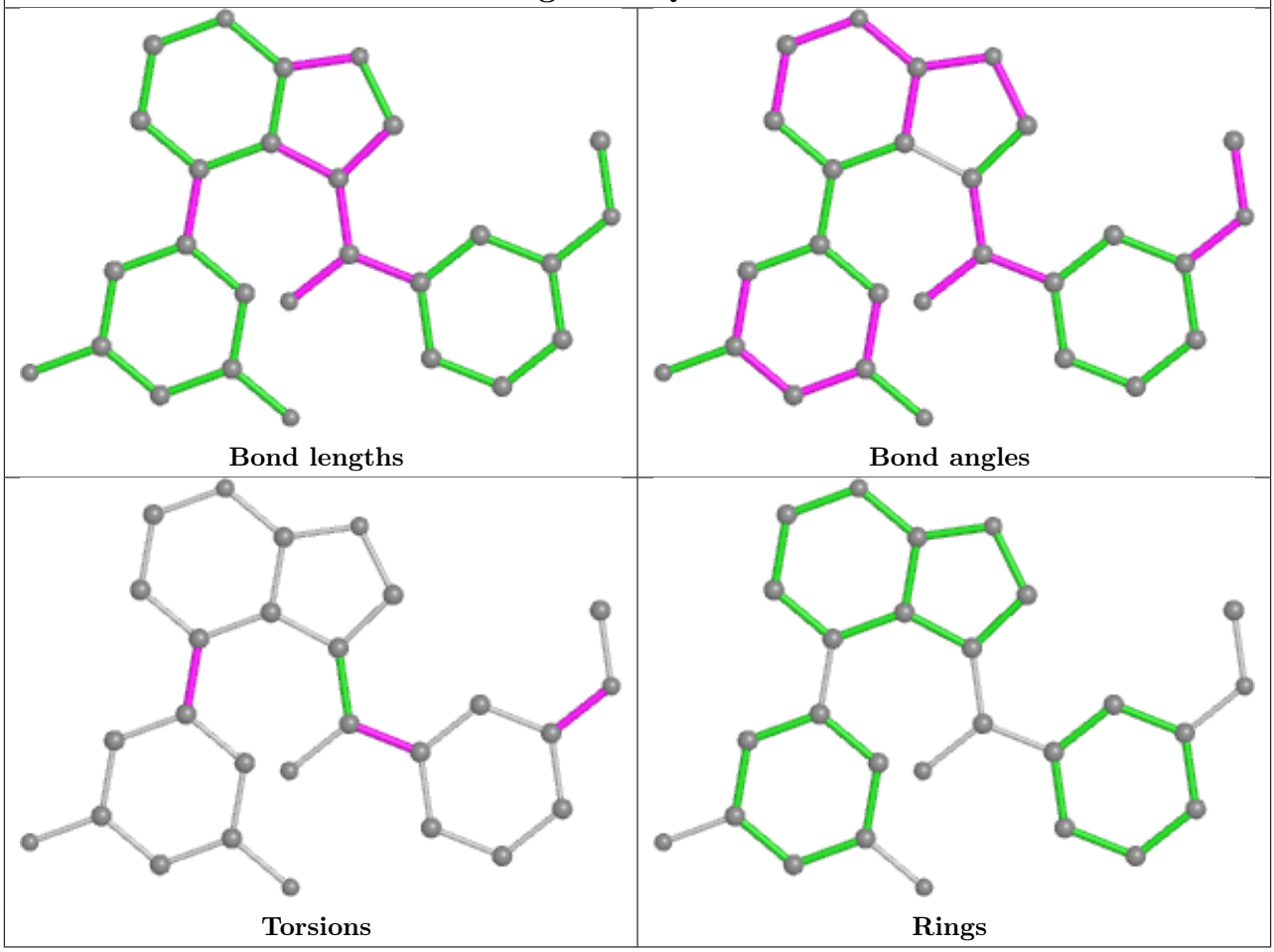


Torsions



Rings

Ligand GQL B 1



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, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	287/326 (88%)	0.18	17 (5%)	22	14	36, 62, 101, 116	0
1	B	276/326 (84%)	0.37	20 (7%)	15	8	45, 76, 111, 130	0
1	C	275/326 (84%)	0.41	29 (10%)	6	3	54, 85, 114, 119	0
All	All	838/978 (85%)	0.32	66 (7%)	12	7	36, 75, 110, 130	0

The worst 5 of 66 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	762	SER	8.7
1	C	766	TYR	7.9
1	B	764	GLN	7.2
1	B	591	PRO	6.9
1	B	763	ASN	6.8

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 (i)

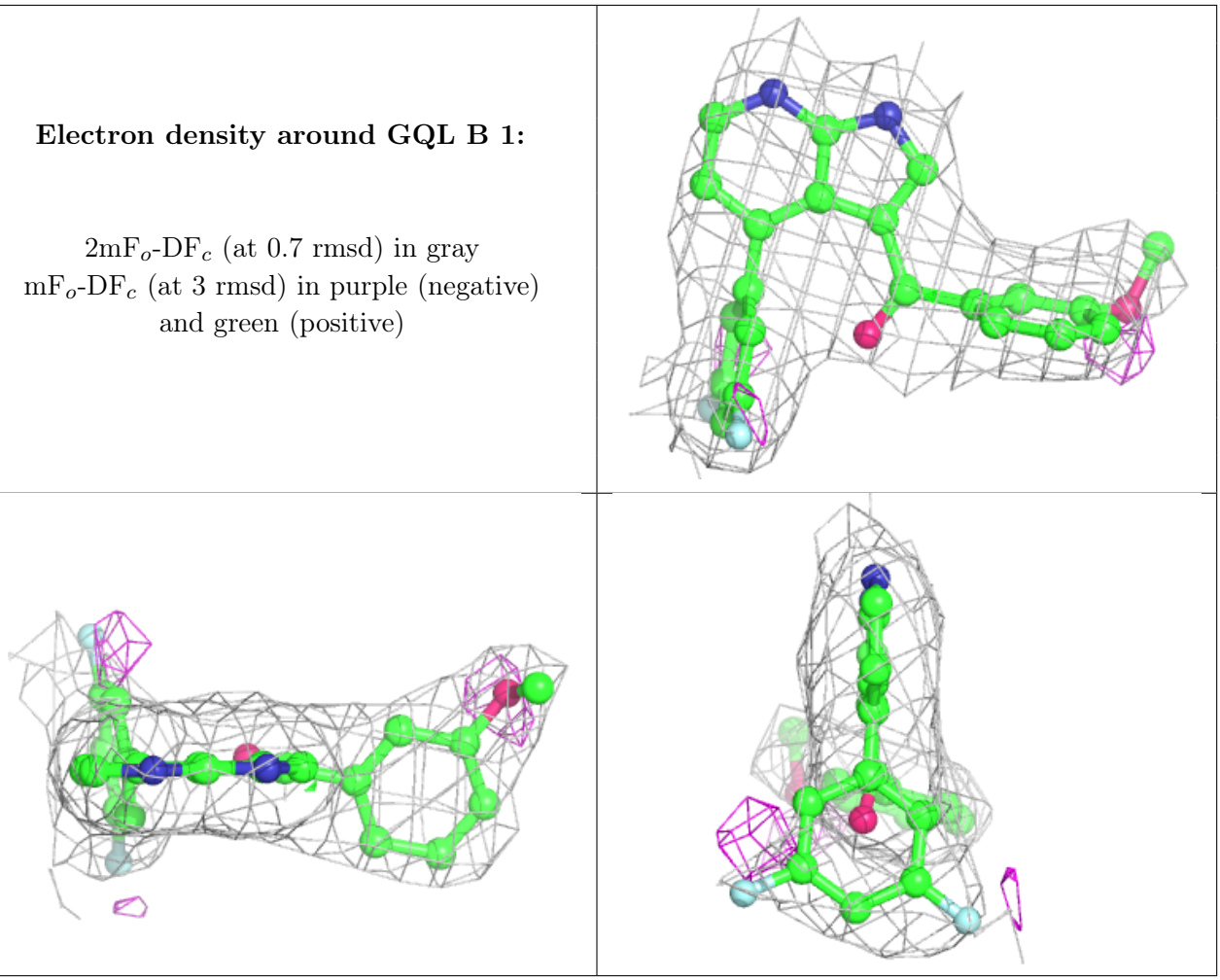
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	GQL	B	1	27/27	0.89	0.24	77,79,82,84	0
2	GQL	A	1	27/27	0.90	0.26	68,74,78,79	0
2	GQL	C	1	27/27	0.90	0.24	86,88,93,94	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.

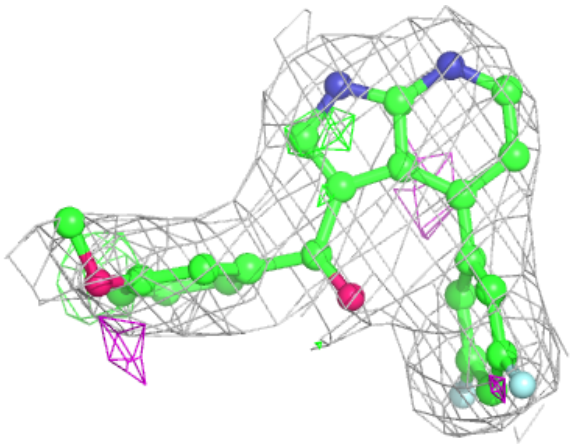
Electron density around GQL B 1:

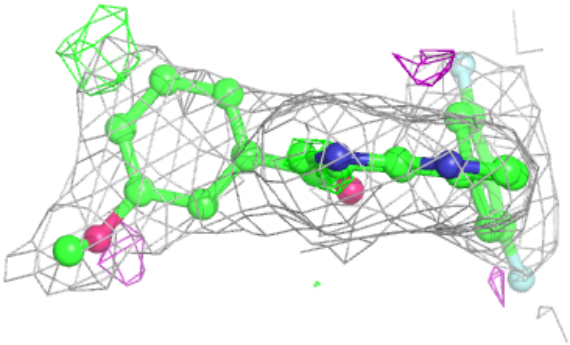
2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative) and green (positive)

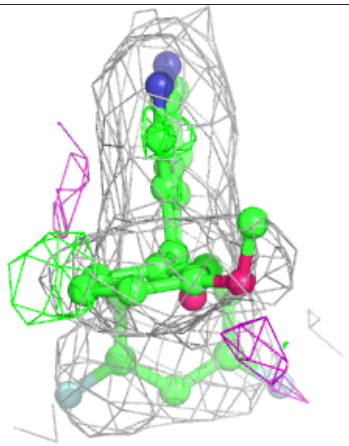


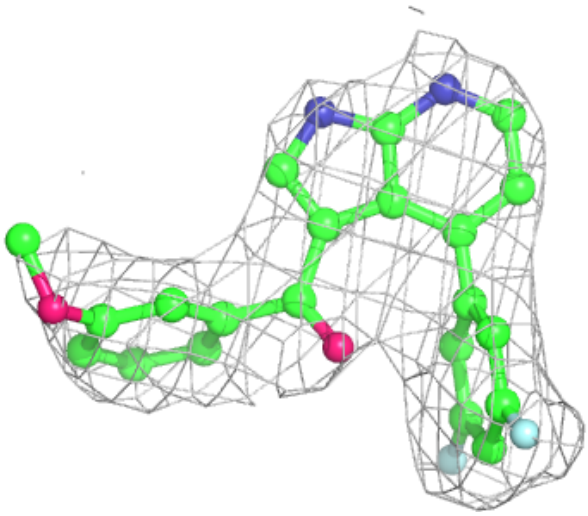
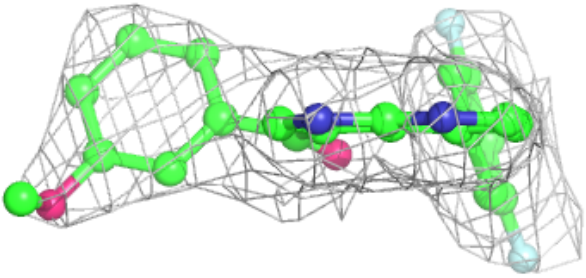
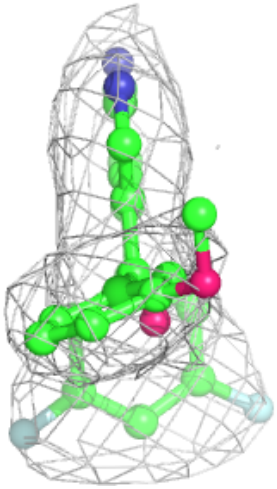
Electron density around GQL A 1:

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







Electron density around GQL C 1: 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.