Full wwPDB X-ray Structure Validation Report (i)

May 14, 2020 – 02:36 am BST

PDB ID	:	2VH6
Title	:	Structure and property based design of factor Xa inhibitors: pyrrolidin-2-ones with biaryl P4 motifs
Authors	:	Young, R.J.; Borthwick, A.D.; Brown, D.; Burns-Kurtis, C.L.; Campbell, M.; Chan, C.; Charbaut, M.; Chung, C.W.; Convery, M.A.; Kelly, H.A.; King, N.P.; Kleanthous, S.; Mason, A.M.; Pateman, A.J.; Patikis, A.N.; Pinto, I.L.; Pollard, D.R.; Senger, S.; Shah, G.P.; Toomey, J.R.; Watson, N.S.; Weston, H.E.
Deposited on	:	2007-11-19
Resolution	:	1.95 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 (1)) were used in the production of this report:

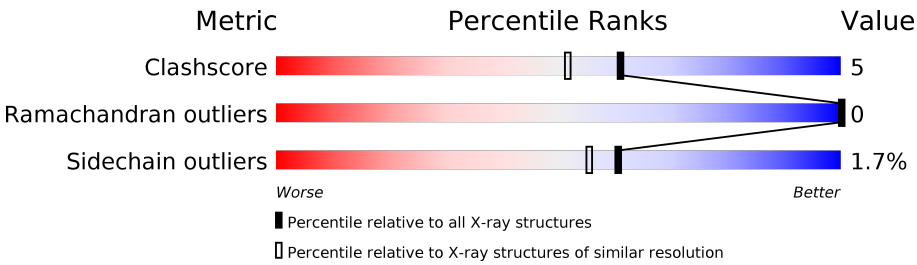
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 1.95 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	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)

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 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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	254	80% 11% 9%
2	B	134	40% • 60%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2433 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.

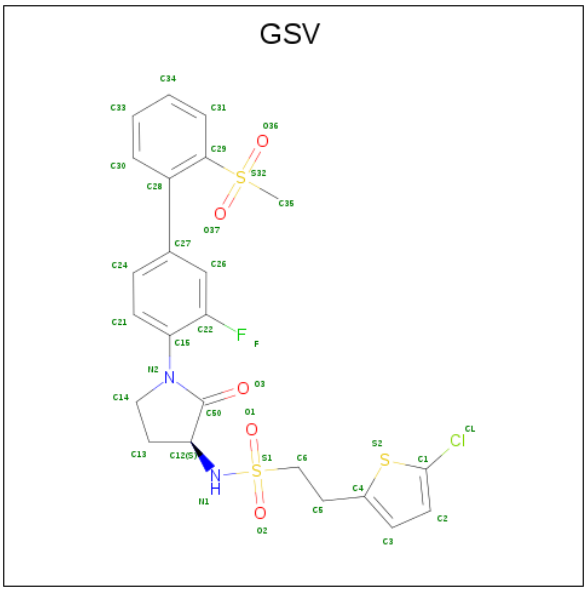
- Molecule 1 is a protein called ACTIVATED FACTOR XA HEAVY CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	232	1837	1158	321	344	14	0	0	0

- Molecule 2 is a protein called ACTIVATED FACTOR XA LIGHT CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	54	390	234	67	82	7	0	0	0

- Molecule 3 is 2-(5-chlorothiophen-2-yl)-N-{(3S)-1-[3-fluoro-2'-(methylsulfonyl)biphenyl-4-yl]-2-oxopyrrolidin-3-yl}ethanesulfonamide (three-letter code: GSV) (formula: $C_{23}H_{22}ClFN_2O_5S_3$).



Mol	Chain	Residues	Atoms							ZeroOcc	AltConf
			Total	C	Cl	F	N	O	S		
3	A	1	35	23	1	1	2	5	3	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	125	125	125	0	0
			Total	O		
4	B	46	46	46	0	0

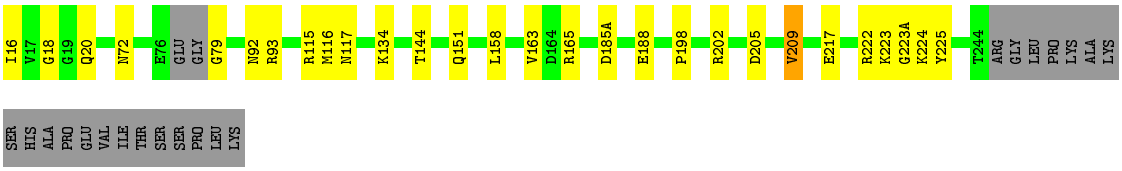
3 Residue-property plots (i)

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

Note EDS was not executed.

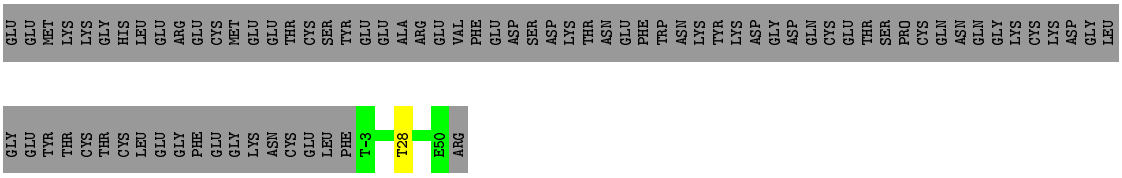
- Molecule 1: ACTIVATED FACTOR XA HEAVY CHAIN

Chain A:



- Molecule 2: ACTIVATED FACTOR XA LIGHT CHAIN

Chain B:



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	55.96Å 72.31Å 78.98Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.95	Depositor
% Data completeness (in resolution range)	100.0 (20.00-1.95)	Depositor
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC 5.3.0006	Depositor
R, R_{free}	0.197 , 0.242	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2433	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

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: GSV

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	0.77	1/1874 (0.1%)	0.77	0/2523
2	B	0.85	0/396	0.75	0/537
All	All	0.79	1/2270 (0.0%)	0.77	0/3060

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	209	VAL	CB-CG2	-5.08	1.42	1.52

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	198	PRO	Peptide
1	A	223(A)	GLY	Mainchain

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	1837	0	1793	19	0
2	B	390	0	345	0	0
3	A	35	0	22	1	0
4	A	125	0	0	6	0
4	B	46	0	0	0	0
All	All	2433	0	2160	20	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 5.

All (20) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:217:GLU:OE1	1:A:224:LYS:HE3	1.61	1.01
1:A:79:GLY:HA2	1:A:115:ARG:NE	2.07	0.69
1:A:217:GLU:CD	1:A:224:LYS:HE3	2.15	0.67
1:A:165:ARG:HG2	4:A:2081:HOH:O	2.00	0.60
1:A:18:GLY:HA3	1:A:188:GLU:HG2	1.85	0.58
1:A:202:ARG:NH1	1:A:205:ASP:OD1	2.22	0.56
1:A:93:ARG:NH2	4:A:2037:HOH:O	2.40	0.54
1:A:72:ASN:O	4:A:2023:HOH:O	2.18	0.53
1:A:217:GLU:OE1	1:A:224:LYS:CE	2.48	0.48
1:A:158:LEU:HD11	1:A:188:GLU:HB3	1.96	0.48
1:A:134:LYS:HG3	4:A:2063:HOH:O	2.13	0.47
1:A:16:ILE:O	1:A:144:THR:HA	2.16	0.46
1:A:134:LYS:HB3	1:A:134:LYS:NZ	2.31	0.45
1:A:185(A):ASP:O	1:A:223:LYS:HA	2.18	0.44
1:A:151:GLN:HB3	1:A:151:GLN:HE21	1.59	0.44
1:A:222:ARG:O	4:A:2113:HOH:O	2.21	0.43
1:A:79:GLY:N	4:A:2028:HOH:O	2.51	0.42
1:A:163:VAL:HG21	1:A:225:TYR:CD1	2.54	0.42
1:A:116:MET:O	1:A:117:ASN:HB2	2.19	0.42
3:A:1245:GSV:H353	3:A:1245:GSV:C24	2.51	0.41

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	228/254 (90%)	222 (97%)	6 (3%)	0	100	100
2	B	52/134 (39%)	48 (92%)	4 (8%)	0	100	100
All	All	280/388 (72%)	270 (96%)	10 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	197/216 (91%)	194 (98%)	3 (2%)	65	60
2	B	43/118 (36%)	42 (98%)	1 (2%)	50	42
All	All	240/334 (72%)	236 (98%)	4 (2%)	60	55

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	20	GLN
1	A	92	ASN
1	A	209	VAL
2	B	28	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	20	GLN
1	A	30	GLN
1	A	92	ASN
1	A	151	GLN
1	A	192	GLN
2	B	16	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 carbohydrates in this entry.

5.6 Ligand geometry (i)

1 ligand is 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	$\#\|Z\| > 2$
3	GSV	A	1245	-	37,38,38	1.39	2 (5%)	40,57,57	1.95	8 (20%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GSV	A	1245	-	-	10/24/38/38	0/4/4/4

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1245	GSV	C5-C4	-5.13	1.48	1.50
3	A	1245	GSV	C31-C29	2.50	1.42	1.39

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1245	GSV	C6-C5-C4	5.57	124.50	111.90
3	A	1245	GSV	O36-S32-C29	-5.19	102.22	108.27
3	A	1245	GSV	O2-S1-N1	-4.38	100.83	107.12
3	A	1245	GSV	C35-S32-C29	3.46	110.24	105.11
3	A	1245	GSV	O37-S32-C29	2.95	111.71	108.27
3	A	1245	GSV	O2-S1-C6	2.47	111.80	107.86
3	A	1245	GSV	C31-C29-C28	-2.39	119.13	121.89
3	A	1245	GSV	C27-C26-C22	-2.20	117.72	119.59

There are no chirality outliers.

All (10) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1245	GSV	C5-C6-S1-N1
3	A	1245	GSV	C5-C6-S1-O2
3	A	1245	GSV	C5-C6-S1-O1
3	A	1245	GSV	C3-C4-C5-C6
3	A	1245	GSV	C28-C29-S32-O37
3	A	1245	GSV	C28-C29-S32-C35
3	A	1245	GSV	C31-C29-S32-O37
3	A	1245	GSV	C24-C27-C28-C30
3	A	1245	GSV	C26-C27-C28-C30
3	A	1245	GSV	C24-C27-C28-C29

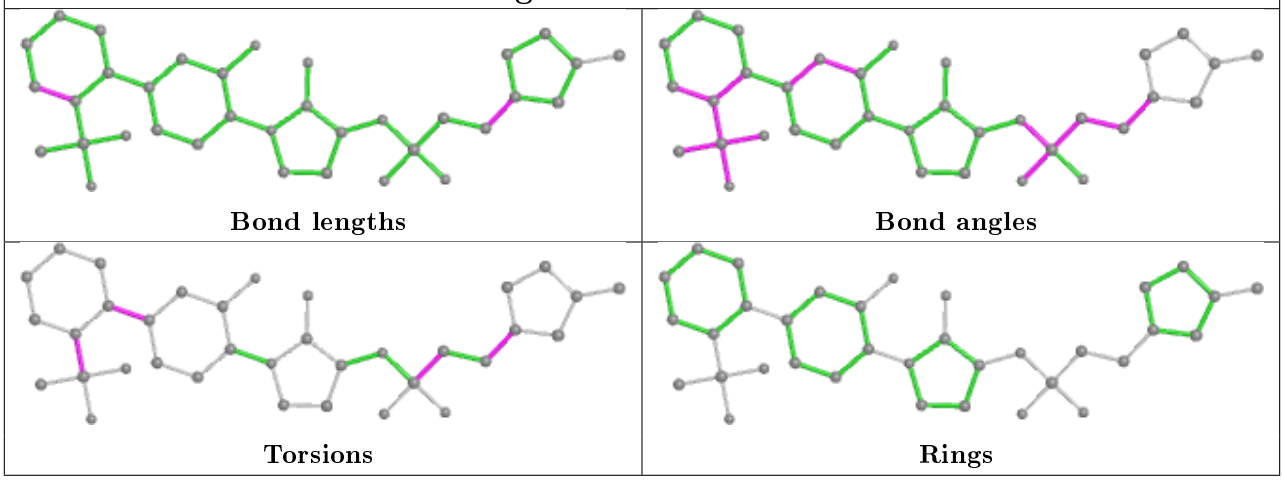
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1245	GSV	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 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 GSV A 1245



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)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.