



# wwPDB X-ray Structure Validation Summary Report

Oct 21, 2021 – 04:22 pm BST

PDB ID : 7NWZ  
Title : ALK:ALKAL2 complex  
Authors : De Munck, S.; Savvides, S.N.  
Deposited on : 2021-03-17  
Resolution : 4.17 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 following versions of software and data (see [references](#) ) 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  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.23.2

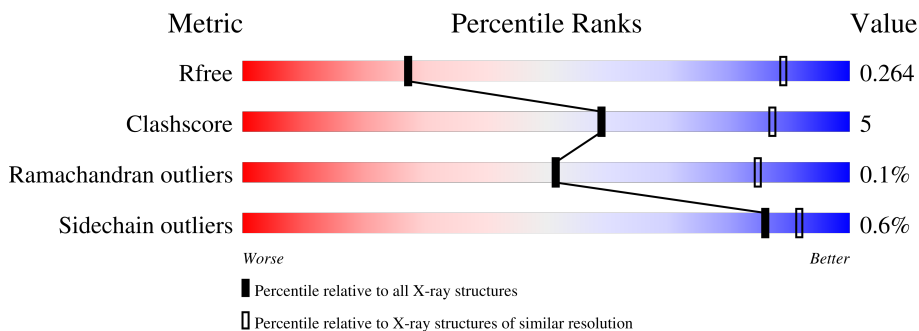
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.17 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))
$R_{free}$	130704	1034 (4.60-3.76)
Clashscore	141614	1030 (4.54-3.80)
Ramachandran outliers	138981	1006 (4.58-3.78)
Sidechain outliers	138945	1037 (4.60-3.76)

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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	C	81	
1	D	81	
2	A	344	
2	B	344	
2	E	344	
2	F	344	

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 9838 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 ALK and LTK ligand 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	D	58	485	312	91	76	6	0	0	0
1	C	58	485	312	91	76	6	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	153	GLY	-	expression tag	UNP Q6UX46
D	154	THR	-	expression tag	UNP Q6UX46
D	155	ASP	-	expression tag	UNP Q6UX46
D	156	GLU	-	expression tag	UNP Q6UX46
D	157	VAL	-	expression tag	UNP Q6UX46
D	158	ASP	-	expression tag	UNP Q6UX46
C	153	GLY	-	expression tag	UNP Q6UX46
C	154	THR	-	expression tag	UNP Q6UX46
C	155	ASP	-	expression tag	UNP Q6UX46
C	156	GLU	-	expression tag	UNP Q6UX46
C	157	VAL	-	expression tag	UNP Q6UX46
C	158	ASP	-	expression tag	UNP Q6UX46

- Molecule 2 is a protein called ALK tyrosine kinase receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	F	307	2190	1354	387	436	13	0	0	0
2	E	307	2190	1354	387	436	13	0	0	0
2	B	309	2202	1360	389	440	13	0	0	0
2	A	309	2202	1360	389	440	13	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	986	GLY	-	expression tag	UNP Q9UM73
F	987	THR	-	expression tag	UNP Q9UM73
F	988	ASP	-	expression tag	UNP Q9UM73
F	989	GLU	-	expression tag	UNP Q9UM73
F	990	VAL	-	expression tag	UNP Q9UM73
F	991	ASP	-	expression tag	UNP Q9UM73
E	986	GLY	-	expression tag	UNP Q9UM73
E	987	THR	-	expression tag	UNP Q9UM73
E	988	ASP	-	expression tag	UNP Q9UM73
E	989	GLU	-	expression tag	UNP Q9UM73
E	990	VAL	-	expression tag	UNP Q9UM73
E	991	ASP	-	expression tag	UNP Q9UM73
B	986	GLY	-	expression tag	UNP Q9UM73
B	987	THR	-	expression tag	UNP Q9UM73
B	988	ASP	-	expression tag	UNP Q9UM73
B	989	GLU	-	expression tag	UNP Q9UM73
B	990	VAL	-	expression tag	UNP Q9UM73
B	991	ASP	-	expression tag	UNP Q9UM73
A	986	GLY	-	expression tag	UNP Q9UM73
A	987	THR	-	expression tag	UNP Q9UM73
A	988	ASP	-	expression tag	UNP Q9UM73
A	989	GLU	-	expression tag	UNP Q9UM73
A	990	VAL	-	expression tag	UNP Q9UM73
A	991	ASP	-	expression tag	UNP Q9UM73

- Molecule 3 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
			Total	C	N	O		
3	F	1	14	8	1	5	0	0
3	E	1	14	8	1	5	0	0
3	B	1	14	8	1	5	0	0
3	B	1	14	8	1	5	0	0
3	B	1	14	8	1	5	0	0
3	A	1	14	8	1	5	0	0

### 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. 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.

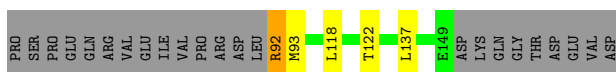
- Molecule 1: ALK and LTK ligand 2

Chain D: 




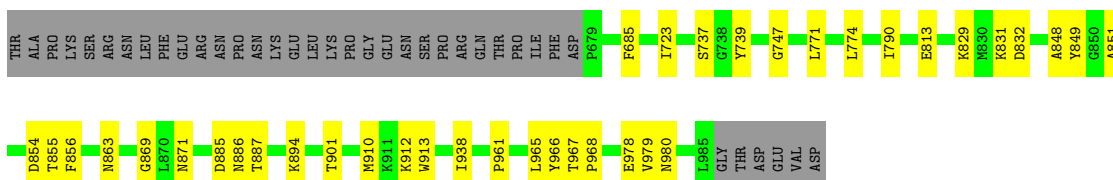
- Molecule 1: ALK and LTK ligand 2

Chain C: 




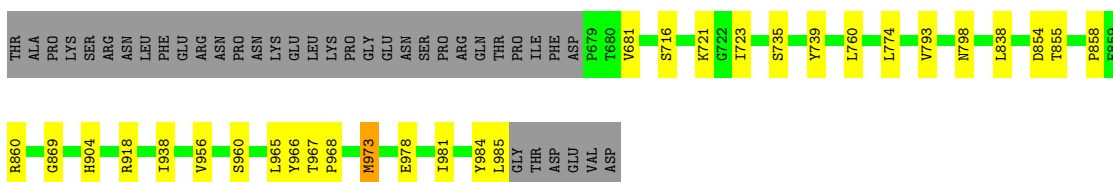
- Molecule 2: ALK tyrosine kinase receptor

Chain F: 




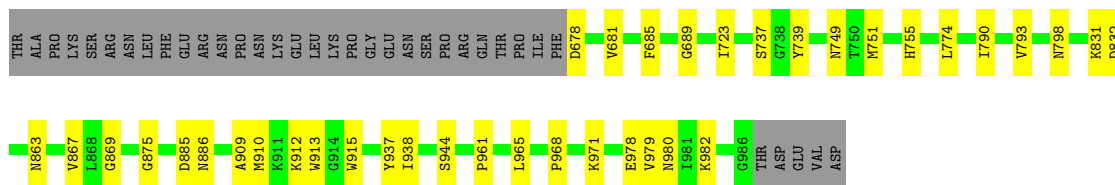
- Molecule 2: ALK tyrosine kinase receptor

Chain E: 



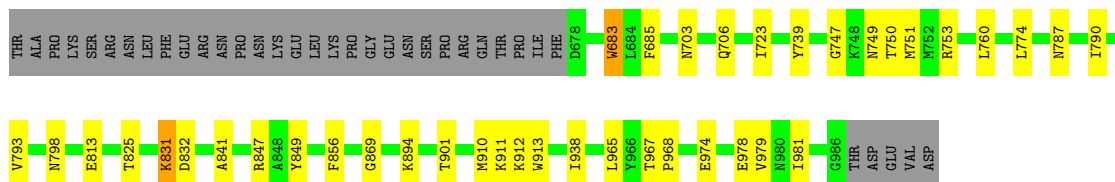
- Molecule 2: ALK tyrosine kinase receptor

Chain B: 



- Molecule 2: ALK tyrosine kinase receptor

Chain A: 78% 11% 10%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	97.91Å 97.91Å 356.17Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.96 – 4.17 48.95 – 4.17	Depositor EDS
% Data completeness (in resolution range)	93.9 (48.96-4.17) 94.0 (48.95-4.17)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.41 (at 4.14Å)	Xtrriage
Refinement program	BUSTER 2.11.2, PHENIX 1.17.1_3660	Depositor
R, $R_{free}$	0.255 , 0.268 0.253 , 0.264	Depositor DCC
$R_{free}$ test set	897 reflections (6.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	105.5	Xtrriage
Anisotropy	0.150	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.25$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.72	EDS
Total number of atoms	9838	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	163.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.45% 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

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	C	0.24	0/499	0.40	0/670
1	D	0.23	0/499	0.37	0/670
2	A	0.25	0/2251	0.44	0/3047
2	B	0.25	0/2251	0.45	0/3047
2	E	0.25	0/2239	0.46	0/3030
2	F	0.25	0/2239	0.45	0/3030
All	All	0.25	0/9978	0.44	0/13494

There are no bond length outliers.

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	C	485	0	495	6	0
1	D	485	0	495	10	0
2	A	2202	0	2059	22	0
2	B	2202	0	2057	22	0
2	E	2190	0	2053	16	0
2	F	2190	0	2053	22	0
3	A	14	0	13	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	42	0	39	0	0
3	E	14	0	13	0	0
3	F	14	0	13	0	0
All	All	9838	0	9290	90	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.

The worst 5 of 90 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:F:790:ILE:HD11	2:F:913:TRP:HB3	1.68	0.76
2:B:790:ILE:HD11	2:B:913:TRP:HB3	1.71	0.72
2:B:755:HIS:HB2	2:B:971:LYS:HG2	1.74	0.69
1:D:118:LEU:HD13	2:E:967:THR:HG21	1.80	0.63
2:F:965:LEU:HD21	2:F:968:PRO:HA	1.82	0.62

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	C	56/81 (69%)	56 (100%)	0	0	100	100
1	D	56/81 (69%)	55 (98%)	1 (2%)	0	100	100
2	A	307/344 (89%)	291 (95%)	15 (5%)	1 (0%)	41	75
2	B	307/344 (89%)	288 (94%)	19 (6%)	0	100	100
2	E	305/344 (89%)	292 (96%)	13 (4%)	0	100	100
2	F	305/344 (89%)	291 (95%)	13 (4%)	1 (0%)	41	75
All	All	1336/1538 (87%)	1273 (95%)	61 (5%)	2 (0%)	51	85

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	855	THR
2	A	683	TRP

### 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	C	54/76 (71%)	53 (98%)	1 (2%)	57	74
1	D	54/76 (71%)	54 (100%)	0	100	100
2	A	219/255 (86%)	218 (100%)	1 (0%)	88	93
2	B	219/255 (86%)	218 (100%)	1 (0%)	88	93
2	E	218/255 (86%)	215 (99%)	3 (1%)	67	80
2	F	218/255 (86%)	218 (100%)	0	100	100
All	All	982/1172 (84%)	976 (99%)	6 (1%)	86	92

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	681	VAL
2	A	831	LYS
1	C	92	ARG
2	E	973	MET
2	E	681	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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](#)

6 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
3	NAG	E	1001	2	14,14,15	1.07	1 (7%)	17,19,21	1.28	1 (5%)
3	NAG	F	1001	2	14,14,15	1.06	1 (7%)	17,19,21	1.25	1 (5%)
3	NAG	B	1002	2	14,14,15	0.28	0	17,19,21	0.48	0
3	NAG	B	1003	2	14,14,15	0.28	0	17,19,21	0.40	0
3	NAG	A	1001	2	14,14,15	1.07	1 (7%)	17,19,21	1.25	1 (5%)
3	NAG	B	1001	2	14,14,15	0.97	1 (7%)	17,19,21	1.15	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
3	NAG	E	1001	2	-	0/6/23/26	0/1/1/1
3	NAG	F	1001	2	-	2/6/23/26	0/1/1/1
3	NAG	B	1002	2	-	0/6/23/26	0/1/1/1
3	NAG	B	1003	2	-	2/6/23/26	0/1/1/1
3	NAG	A	1001	2	-	1/6/23/26	0/1/1/1
3	NAG	B	1001	2	-	2/6/23/26	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	1001	NAG	O5-C1	3.77	1.49	1.43
3	A	1001	NAG	O5-C1	3.75	1.49	1.43
3	E	1001	NAG	O5-C1	3.73	1.49	1.43
3	B	1001	NAG	O5-C1	3.38	1.49	1.43

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	1001	NAG	C1-O5-C5	5.03	119.01	112.19
3	F	1001	NAG	C1-O5-C5	4.96	118.91	112.19
3	A	1001	NAG	C1-O5-C5	4.90	118.83	112.19
3	B	1001	NAG	C1-O5-C5	4.51	118.31	112.19

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	1001	NAG	O5-C5-C6-O6
3	F	1001	NAG	C4-C5-C6-O6
3	F	1001	NAG	O5-C5-C6-O6
3	B	1001	NAG	C4-C5-C6-O6
3	B	1003	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

## 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

### 6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.