Full wwPDB X-ray Structure Validation Report ⓘ

Apr 7, 2025 – 10:25 AM JST

PDB ID	:	8WS0 / pdb_00008ws0
Title	:	Crystal structure of human NEK7 S195D mutant
Authors	:	Bijpuria, S.; Athresh, S.; Subbiah, R.; Mollard, A.; Bearss, D.
Deposited on	:	2023-10-16
Resolution	:	2.12 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 ⓘ 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:

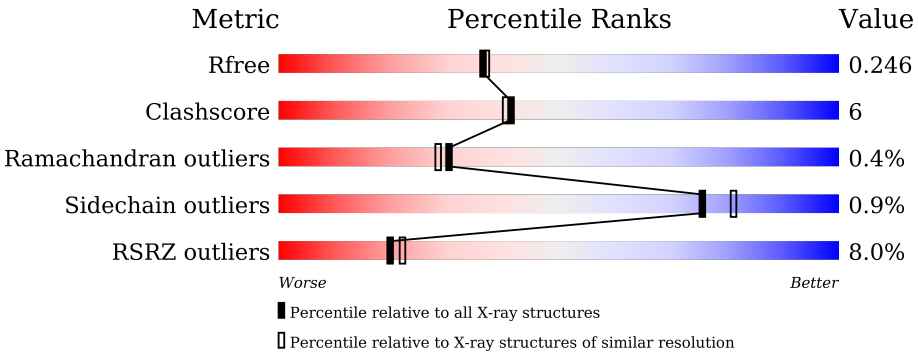
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.42

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 2.12 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	7689 (2.14-2.10)
Clashscore	180529	8431 (2.14-2.10)
Ramachandran outliers	177936	8366 (2.14-2.10)
Sidechain outliers	177891	8367 (2.14-2.10)
RSRZ outliers	164620	7689 (2.14-2.10)

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	310	5% 78% 8% 14%
1	B	310	9% 70% 16% • 13%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4540 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 Serine/threonine-protein kinase Nek7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	267	Total	C	N	O	S	5	0	0
			2158	1381	373	386	18			
1	B	270	Total	C	N	O	S	0	5	0
			2192	1405	374	396	17			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	195	ASP	SER	engineered mutation	UNP Q8TDX7
A	303	LEU	-	expression tag	UNP Q8TDX7
A	304	GLU	-	expression tag	UNP Q8TDX7
A	305	HIS	-	expression tag	UNP Q8TDX7
A	306	HIS	-	expression tag	UNP Q8TDX7
A	307	HIS	-	expression tag	UNP Q8TDX7
A	308	HIS	-	expression tag	UNP Q8TDX7
A	309	HIS	-	expression tag	UNP Q8TDX7
A	310	HIS	-	expression tag	UNP Q8TDX7
B	195	ASP	SER	engineered mutation	UNP Q8TDX7
B	303	LEU	-	expression tag	UNP Q8TDX7
B	304	GLU	-	expression tag	UNP Q8TDX7
B	305	HIS	-	expression tag	UNP Q8TDX7
B	306	HIS	-	expression tag	UNP Q8TDX7
B	307	HIS	-	expression tag	UNP Q8TDX7
B	308	HIS	-	expression tag	UNP Q8TDX7
B	309	HIS	-	expression tag	UNP Q8TDX7
B	310	HIS	-	expression tag	UNP Q8TDX7

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	99	Total	O	0	0
			99	99		

Continued on next page...

Continued from previous page...

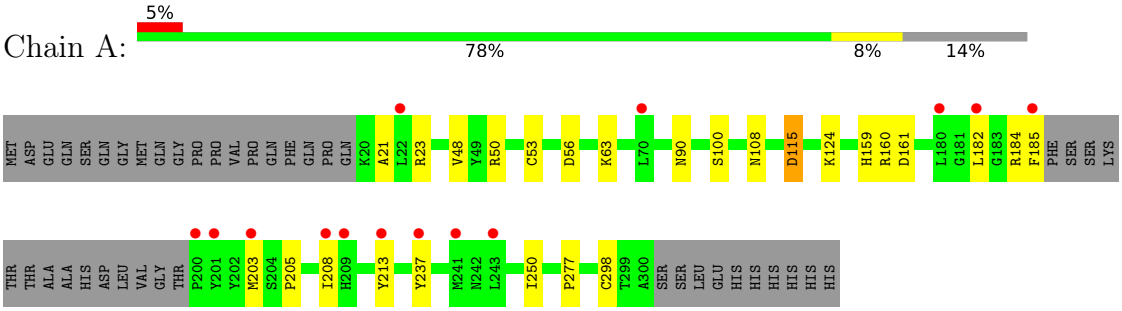
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	91	Total	O	0	0
			91	91		

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.

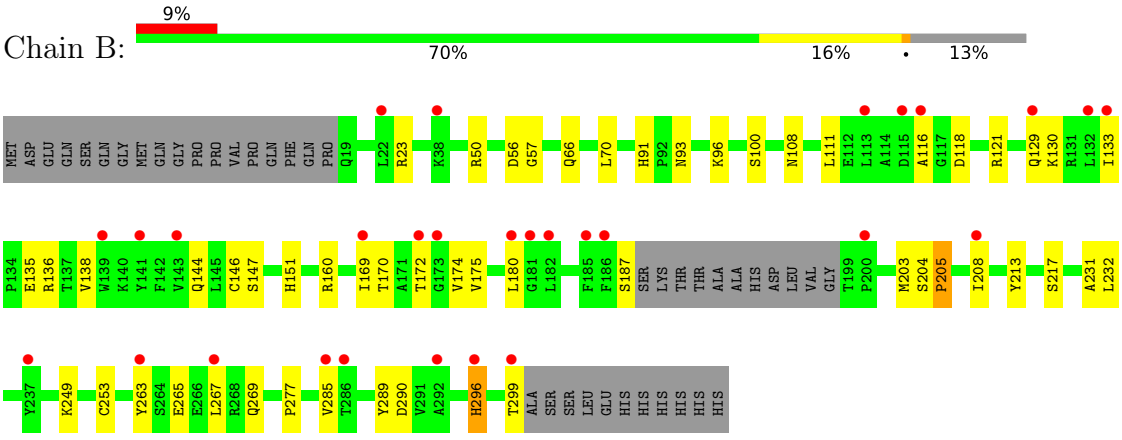
• Molecule 1: Serine/threonine-protein kinase Nek7

Chain A:



• Molecule 1: Serine/threonine-protein kinase Nek7

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	88.60Å 88.60Å 156.09Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.34 – 2.12 44.34 – 2.12	Depositor EDS
% Data completeness (in resolution range)	99.6 (44.34-2.12) 99.9 (44.34-2.12)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.76 (at 2.12Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.191 , 0.242 0.197 , 0.246	Depositor DCC
R_{free} test set	1948 reflections (4.75%)	wwPDB-VP
Wilson B-factor (Å^2)	50.7	Xtriage
Anisotropy	0.078	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 44.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.025 for -h,-k,l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	4540	wwPDB-VP
Average B, all atoms (Å^2)	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.39% 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

5.1 Standard geometry

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	0.91	2/2207 (0.1%)	1.04	3/2978 (0.1%)
1	B	1.13	10/2251 (0.4%)	1.05	4/3042 (0.1%)
All	All	1.03	12/4458 (0.3%)	1.05	7/6020 (0.1%)

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	136	ARG	C-O	12.71	1.47	1.23
1	A	50	ARG	NE-CZ	-10.68	1.19	1.33
1	A	124	LYS	CD-CE	9.51	1.75	1.51
1	B	263	TYR	CB-CG	7.76	1.63	1.51
1	B	146	CYS	C-O	7.08	1.36	1.23
1	B	135	GLU	CG-CD	6.55	1.61	1.51
1	B	285	VAL	C-O	5.55	1.33	1.23
1	B	263	TYR	C-O	5.53	1.33	1.23
1	B	296	HIS	ND1-CE1	-5.41	1.21	1.34
1	B	56	ASP	CG-OD2	5.20	1.37	1.25
1	B	263	TYR	CE2-CZ	5.17	1.45	1.38
1	B	57	GLY	N-CA	5.01	1.53	1.46

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	50	ARG	NE-CZ-NH1	-11.47	114.56	120.30
1	A	50	ARG	NE-CZ-NH2	11.26	125.93	120.30
1	A	50	ARG	CD-NE-CZ	10.23	137.92	123.60
1	B	56	ASP	CB-CG-OD1	-7.90	111.19	118.30
1	B	118	ASP	CB-CA-C	-5.90	98.60	110.40
1	B	263	TYR	CB-CA-C	5.67	121.73	110.40
1	B	50	ARG	NE-CZ-NH1	-5.60	117.50	120.30

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	2158	0	2159	18	0
1	B	2192	0	2177	38	0
2	A	99	0	0	2	0
2	B	91	0	0	3	0
All	All	4540	0	4336	55	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 6.

All (55) 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:B:91:HIS:HD2	1:B:93:ASN:H	1.24	0.81
1:B:91:HIS:CD2	1:B:93:ASN:H	2.05	0.74
1:B:144:GLN:HE22	1:B:174:VAL:HA	1.59	0.66
1:B:170:THR:CG2	1:B:172:THR:HG22	2.31	0.60
1:B:23:ARG:HD2	2:B:461:HOH:O	2.02	0.60
1:A:205:PRO:HG2	1:A:277:PRO:HA	1.84	0.59
1:A:160:ARG:HA	1:A:213:TYR:CZ	2.39	0.57
1:B:170:THR:HG22	1:B:172:THR:HG22	1.86	0.56
1:A:160:ARG:NH1	2:A:404:HOH:O	2.37	0.56
1:B:267:LEU:HD23	1:B:267:LEU:C	2.26	0.56
1:A:48:VAL:HG22	1:A:63:LYS:HG2	1.89	0.55
1:A:203:MET:HB3	1:A:250:ILE:HD11	1.90	0.54
1:A:63:LYS:HE2	1:A:182:LEU:HG	1.89	0.54
1:B:296:HIS:ND1	2:B:403:HOH:O	2.33	0.54
1:B:133[B]:ILE:HG23	1:B:138:VAL:CG2	2.39	0.52
1:B:231:ALA:O	1:B:232:LEU:HB2	2.10	0.52
1:A:159:HIS:O	1:A:160:ARG:HB2	2.10	0.51
1:B:203:MET:HG2	1:B:208:ILE:HG13	1.91	0.51
1:B:66:GLN:HG3	1:B:70:LEU:HD12	1.93	0.50
1:B:269:GLN:HA	1:B:269:GLN:OE1	2.12	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:63:LYS:HD3	1:A:182:LEU:HD11	1.95	0.48
1:A:90:ASN:OD1	1:B:96:LYS:NZ	2.45	0.47
1:A:21:ALA:O	1:A:23:ARG:HG2	2.14	0.47
1:A:161:ASP:OD1	2:A:401:HOH:O	2.20	0.47
1:B:91:HIS:HE1	1:B:147[B]:SER:OG	1.98	0.47
1:A:184:ARG:O	1:A:185:PHE:HB2	2.14	0.46
1:B:147[A]:SER:HB3	1:B:289:TYR:CZ	2.50	0.46
1:B:116:ALA:HB1	1:B:169:ILE:O	2.15	0.46
1:B:160:ARG:HA	1:B:213:TYR:CZ	2.51	0.46
1:B:205:PRO:HG2	1:B:277:PRO:HA	1.98	0.46
1:A:100:SER:HA	1:A:108:ASN:O	2.16	0.45
1:B:129[B]:GLN:O	1:B:129[B]:GLN:CG	2.65	0.45
1:B:133[B]:ILE:HG23	1:B:138:VAL:HG23	1.99	0.45
1:B:213:TYR:CE1	1:B:217:SER:HB2	2.51	0.45
1:A:160:ARG:HA	1:A:213:TYR:CE1	2.51	0.45
1:B:91:HIS:CE1	1:B:147[B]:SER:OG	2.70	0.45
1:A:53:CYS:HB3	1:A:56:ASP:OD1	2.18	0.44
1:A:208:ILE:HD12	1:A:250:ILE:HG21	2.00	0.44
1:B:253:CYS:SG	1:B:277:PRO:HD3	2.57	0.44
1:B:296:HIS:O	1:B:296:HIS:CG	2.71	0.44
1:B:147[A]:SER:HB3	1:B:289:TYR:OH	2.17	0.44
1:B:204[B]:SER:OG	1:B:205:PRO:HD2	2.18	0.43
1:B:299:THR:HB	2:B:445:HOH:O	2.17	0.43
1:B:111:LEU:HD13	1:B:180:LEU:HD11	1.99	0.43
1:B:91:HIS:HB2	1:B:151:HIS:CD2	2.53	0.42
1:B:172:THR:HG23	1:B:174:VAL:HG23	2.00	0.42
1:B:249:LYS:HB3	1:B:249:LYS:HE3	1.59	0.42
1:B:121:ARG:HH12	1:B:187:SER:C	2.23	0.42
1:A:237:TYR:CD1	1:A:237:TYR:C	2.94	0.41
1:B:100:SER:HA	1:B:108:ASN:O	2.21	0.41
1:B:265:GLU:O	1:B:269:GLN:HB2	2.20	0.41
1:B:133[B]:ILE:HG23	1:B:133[B]:ILE:O	2.21	0.41
1:A:63:LYS:HD3	1:A:182:LEU:CD1	2.51	0.41
1:B:144:GLN:NE2	1:B:175:VAL:H	2.18	0.41
1:B:203:MET:HG2	1:B:208:ILE:CG1	2.51	0.40

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	263/310 (85%)	250 (95%)	12 (5%)	1 (0%)	30 28
1	B	271/310 (87%)	260 (96%)	10 (4%)	1 (0%)	30 28
All	All	534/620 (86%)	510 (96%)	22 (4%)	2 (0%)	30 28

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	115	ASP
1	B	130	LYS

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	234/274 (85%)	232 (99%)	2 (1%)	75 81
1	B	238/274 (87%)	236 (99%)	2 (1%)	79 84
All	All	472/548 (86%)	468 (99%)	4 (1%)	75 84

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

Mol	Chain	Res	Type
1	A	115	ASP
1	A	298	CYS
1	B	205	PRO
1	B	290	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	91	HIS
1	B	93	ASN
1	B	144	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no oligosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	267/310 (86%)	0.22	14 (5%)	34	37	35, 56, 90, 112	4 (1%)
1	B	270/310 (87%)	0.67	29 (10%)	12	14	34, 61, 96, 123	6 (2%)
All	All	537/620 (86%)	0.45	43 (8%)	20	22	34, 58, 92, 123	10 (1%)

All (43) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	180	LEU	4.0
1	A	182	LEU	3.8
1	B	185	PHE	3.6
1	A	203	MET	3.6
1	A	70	LEU	3.3
1	B	182	LEU	3.3
1	B	186	PHE	3.1
1	B	22	LEU	3.1
1	B	292	ALA	3.1
1	A	180	LEU	3.0
1	B	133[A]	ILE	2.9
1	B	299	THR	2.9
1	A	185	PHE	2.7
1	B	169	ILE	2.7
1	B	129[A]	GLN	2.7
1	A	209	HIS	2.6
1	B	132	LEU	2.6
1	B	237	TYR	2.6
1	B	115	ASP	2.5
1	B	263	TYR	2.5
1	B	200	PRO	2.5
1	A	243	LEU	2.4
1	B	113	LEU	2.4
1	B	172	THR	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	181	GLY	2.3
1	B	208	ILE	2.3
1	B	173	GLY	2.3
1	B	38	LYS	2.2
1	A	208	ILE	2.2
1	B	141	TYR	2.2
1	B	286	THR	2.2
1	A	241	MET	2.2
1	B	139	TRP	2.2
1	B	116	ALA	2.2
1	B	296	HIS	2.2
1	A	237	TYR	2.1
1	B	143	VAL	2.1
1	B	285	VAL	2.1
1	B	267	LEU	2.1
1	A	200	PRO	2.1
1	A	201	TYR	2.1
1	A	22	LEU	2.1
1	A	213	TYR	2.0

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.