



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 21, 2020 – 12:33 pm BST

PDB ID : 3CDJ  
Title : Crystal structure of the E. coli KH/S1 domain truncated PNPase  
Authors : Shi, Z.; Yang, W.Z.; Lin-Chao, S.; Chak, K.F.; Yuan, H.S.  
Deposited on : 2008-02-27  
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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

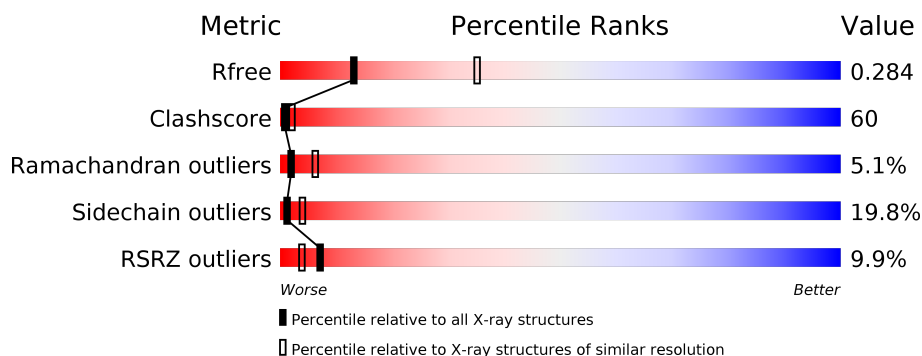
# 1 Overall quality at a glance

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))
$R_{free}$	130704	3140 (2.80-2.80)
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 on 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	559	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3418 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 Polynucleotide phosphorylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	445	Total	C	N	O	S	0	0	0
			3382	2132	585	647	18			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	36	Total	O	0	0
			36	36		

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 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.

Chain A:

8% 30% 34% 11% 20%

LYS ARG LYS Q37 E38 R39 T90 D21 T22 T23 T24 L25 L26 L27 L28 L29 L30 L31 L32 L33 L34 L35 L36 L37 L38 L39 L40 L41 L42 L43 L44 L45 L46 L47 L48 L49 L50 L51 L52 L53 L54 L55 L56 L57 L58 L59 L60 L61 L62 L63 L64 L65 L66 L67 L68 L69 L70 L71 L72 L73 L74 L75 L76 L77 L78 L79 L80 L81 L82 L83 L84 L85 L86 L87 L88 L89 L90 L91 L92 L93 L94 L95 L96 L97 L98 L99 L100 L101 L102 L103 L104 L105 L106 L107 L108 L109 L110 L111 L112 L113 L114 L115 L116 L117 L118 L119 L120 L121 L122 L123 L124 L125 L126 L127 L128 L129 L130 L131 L132 L133 L134 L135 L136 L137 L138 L139 L140 L141 L142 L143 L144 L145 L146 L147 L148 L149 L150 L151 L152 L153 L154 L155 L156 L157 L158 L159 L160 L161 L162 L163 L164 L165 L166 L167 L168 L169 L170 L171 L172 L173 L174 L175 L176 L177 L178 L179 L180 L181 L182 L183 L184 L185 L186 L187 L188 L189 L190 L191 L192 L193 L194 L195 L196 L197 L198 L199 L200 L201 L202 L203 L204 L205 L206 L207 L208 L209 L210 L211 L212 L213 L214 L215 L216 L217 L218 L219 L220 L221 L222 L223 L224 L225 L226 L227 L228 L229 L230 L231 L232 L233 L234 L235 L236 L237 L238 L239 L240 L241 L242 L243 L244 L245 L246 L247 L248 L249 L250 L251 L252 L253 L254 L255 L256 L257 L258 L259 L260 L261 L262 L263 L264 L265 L266 L267 L268 L269 L270 L271 L272 L273 L274 L275 L276 L277 L278 L279 L280 L281 L282 L283 L284 L285 L286 L287 L288 L289 L290 L291 L292 L293 L294 L295 L296 L297 L298 L299 L300 L301 L302 L303 L304 L305 L306 L307 L308 L309 L310 L311 L312 L313 L314 L315 L316 L317 L318 L319 L320 L321 L322 L323 L324 L325 L326 L327 L328 L329 L330 L331 L332 L333 L334 L335 L336 L337 L338 L339 L340 L341 L342 L343 L344 L345 L346 L347 L348 L349 L350 L351 L352 L353 L354 L355 L356 L357 L358 L359 L360 L361 L362 L363 L364 L365 L366 L367 L368 L369 L370 L371 L372 L373 L374 L375 L376 L377 L378 L379 L380 L381 L382 L383 L384 L385 L386 L387 L388 L389 L390 L391 L392 L393 L394 L395 L396 L397 L398 L399 L400 L401 L402 L403 L404 L405 L406 L407 L408 L409 L410 L411 L412 L413 L414 L415 L416 L417 L418 L419 L420 L421 L422 L423 L424 L425 L426 L427 L428 L429 L430 L431 L432 L433 L434 L435 L436 L437 L438 L439 L440 L441 L442 L443 L444 L445 L446 L447 L448 L449 L450 L451 L452 L453 L454 L455 L456 L457 L458 L459 L460 L461 L462 L463 L464 L465 L466 L467 L468 L469 L470 L471 L472 L473 L474 L475 L476 L477 L478 L479 L480 L481 L482 L483 L484 L485 L486 L487 L488 L489 L490 L491 L492 L493 L494 L495 L496 L497 L498 L499 L500 L501 L502 L503 L504 L505 L506 L507 L508 L509 L510 L511 L512 L513 L514 L515 L516 L517 L518 L519 L520 L521 L522 L523 L524 L525 L526 L527 L528 L529 L530 L531 L532 L533 L534 L535 L536 L537 L538 L539 L540 L541 L542 L543 L544 L545 L546 L547 L548 L549 L550 L551 L552 L553 L554 L555 L556 L557 L558 L559 L560 L561 L562 L563 L564 L565 L566 L567 L568 L569 L570 L571 L572 L573 L574 L575 L576 L577 L578 L579 L580 L581 L582 L583 L584 L585 L586 L587 L588 L589 L590 L591 L592 L593 L594 L595 L596 L597 L598 L599 L600 L601 L602 L603 L604 L605 L606 L607 L608 L609 L610 L611 L612 L613 L614 L615 L616 L617 L618 L619 L620 L621 L622 L623 L624 L625 L626 L627 L628 L629 L630 L631 L632 L633 L634 L635 L636 L637 L638 L639 L640 L641 L642 L643 L644 L645 L646 L647 L648 L649 L650 L651 L652 L653 L654 L655 L656 L657 L658 L659 L660 L661 L662 L663 L664 L665 L666 L667 L668 L669 L670 L671 L672 L673 L674 L675 L676 L677 L678 L679 L680 L681 L682 L683 L684 L685 L686 L687 L688 L689 L690 L691 L692 L693 L694 L695 L696 L697 L698 L699 L700 L701 L702 L703 L704 L705 L706 L707 L708 L709 L710 L711 L712 L713 L714 L715 L716 L717 L718 L719 L720 L721 L722 L723 L724 L725 L726 L727 L728 L729 L730 L731 L732 L733 L734 L735 L736 L737 L738 L739 L740 L741 L742 L743 L744 L745 L746 L747 L748 L749 L750 L751 L752 L753 L754 L755 L756 L757 L758 L759 L760 L761 L762 L763 L764 L765 L766 L767 L768 L769 L770 L771 L772 L773 L774 L775 L776 L777 L778 L779 L780 L781 L782 L783 L784 L785 L786 L787 L788 L789 L790 L791 L792 L793 L794 L795 L796 L797 L798 L799 L800 L801 L802 L803 L804 L805 L806 L807 L808 L809 L810 L811 L812 L813 L814 L815 L816 L817 L818 L819 L820 L821 L822 L823 L824 L825 L826 L827 L828 L829 L830 L831 L832 L833 L834 L835 L836 L837 L838 L839 L840 L841 L842 L843 L844 L845 L846 L847 L848 L849 L850 L851 L852 L853 L854 L855 L856 L857 L858 L859 L860 L861 L862 L863 L864 L865 L866 L867 L868 L869 L870 L871 L872 L873 L874 L875 L876 L877 L878 L879 L880 L881 L882 L883 L884 L885 L886 L887 L888 L889 L890 L891 L892 L893 L894 L895 L896 L897 L898 L899 L900 L901 L902 L903 L904 L905 L906 L907 L908 L909 L910 L911 L912 L913 L914 L915 L916 L917 L918 L919 L920 L921 L922 L923 L924 L925 L926 L927 L928 L929 L930 L931 L932 L933 L934 L935 L936 L937 L938 L939 L940 L941 L942 L943 L944 L945 L946 L947 L948 L949 L950 L951 L952 L953 L954 L955 L956 L957 L958 L959 L960 L961 L962 L963 L964 L965 L966 L967 L968 L969 L970 L971 L972 L973 L974 L975 L976 L977 L978 L979 L980 L981 L982 L983 L984 L985 L986 L987 L988 L989 L990 L991 L992 L993 L994 L995 L996 L997 L998 L999 L1000 L1001 L1002 L1003 L1004 L1005 L1006 L1007 L1008 L1009 L1010 L1011 L1012 L1013 L1014 L1015 L10

## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	160.09Å 160.09Å 153.15Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 2.80 49.58 – 2.80	Depositor EDS
% Data completeness (in resolution range)	92.3 (50.00-2.80) 95.2 (49.58-2.80)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	9.30 (at 2.81Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.270 , 0.294 0.280 , 0.284	Depositor DCC
$R_{free}$ test set	1817 reflections (10.22%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	60.1	Xtriage
Anisotropy	0.998	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 82.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	3418	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	91.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.45% 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](#)

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	1.18	17/3432 (0.5%)	1.29	49/4648 (1.1%)

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

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	91	TYR	CG-CD2	19.43	1.64	1.39
1	A	91	TYR	CE1-CZ	14.92	1.57	1.38
1	A	91	TYR	CA-CB	9.33	1.74	1.53
1	A	91	TYR	CE2-CZ	7.95	1.48	1.38
1	A	396	THR	CA-C	-7.09	1.34	1.52

The worst 5 of 49 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	395	ARG	N-CA-C	17.06	157.07	111.00
1	A	91	TYR	CB-CG-CD1	11.83	128.10	121.00
1	A	91	TYR	CB-CG-CD2	-11.70	113.98	121.00
1	A	338	ARG	N-CA-C	10.67	139.81	111.00
1	A	395	ARG	N-CA-CB	-10.61	91.50	110.60

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	394	GLU	Mainchain,Peptide

## 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	3382	0	3415	405	2
2	A	36	0	0	4	0
All	All	3418	0	3415	405	2

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 60.

The worst 5 of 405 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:91:TYR:CA	1:A:91:TYR:CB	1.74	1.58
1:A:395:ARG:HG2	1:A:396:THR:N	1.40	1.19
1:A:325:LYS:O	1:A:327:VAL:N	1.76	1.18
1:A:69:GLY:C	1:A:70:GLN:HG3	1.55	1.14
1:A:394:GLU:OE2	2:A:599:HOH:O	1.63	1.12

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:30:ARG:NH1	1:A:30:ARG:NH1[18_444]	2.09	0.11
1:A:30:ARG:NH2	1:A:30:ARG:NH2[18_444]	2.15	0.05

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	433/559 (78%)	381 (88%)	30 (7%)	22 (5%)	2 6

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	171	PRO
1	A	256	TRP
1	A	324	GLU
1	A	326	ASN
1	A	337	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	364/456 (80%)	292 (80%)	72 (20%)	1 4

5 of 72 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	222	ASP
1	A	329	ARG
1	A	533	LYS
1	A	225	LEU
1	A	259	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	181	ASN
1	A	183	GLN
1	A	497	ASN
1	A	148	ASN

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	372	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](#)

There are no ligands in this entry.

### 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	445/559 (79%)	0.72	44 (9%) <b>7</b> <b>4</b>	19, 84, 151, 166	0

The worst 5 of 44 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	334	ALA	5.2
1	A	336	GLU	4.4
1	A	78	ASP	4.3
1	A	260	PRO	4.2
1	A	257	GLN	4.1

### 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 carbohydrates 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.