WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 10:52 PM GMT

PDB ID	:	1VKX
Title	:	CRYSTAL STRUCTURE OF THE NFKB P50/P65 HETERODIMER COMPLEXED TO THE IMMUNOGLOBULIN KB DNA
Authors	:	Chen, F.; Huang, D.B.; Ghosh, G.
Deposited on	:	1997-09-17
Resolution	:	2.90 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.90 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	Percentile Ranks	Value
Clashscore		56
Ramachandran outliers		7.4%
Sidechain outliers		16.7%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
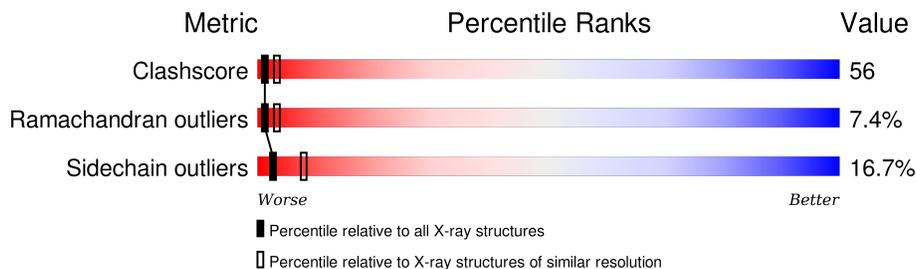
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)

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. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	C	12	8% 67% 17% 8%
2	D	12	17% 83%
3	A	273	30% 55% 13% •
4	B	312	32% 52% 14% •

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5116 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 DNA chain called DNA (5'-D(*TP*GP*GP*GP*GP*AP*CP*TP*TP*TP*CP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	12	Total 243	C 117	N 42	O 73	P 11	0	0	0

- Molecule 2 is a DNA chain called DNA (5'-D(*AP*GP*GP*AP*AP*AP*GP*TP*CP*CP*CP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	12	Total 243	C 116	N 49	O 67	P 11	0	0	0

- Molecule 3 is a protein called PROTEIN (NF-KAPPA B P65 SUBUNIT).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	273	Total 2176	C 1356	N 401	O 408	S 11	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	19	ALA	PRO	CONFLICT	UNP Q04207

- Molecule 4 is a protein called PROTEIN (NF-KAPPA B P50 SUBUNIT).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	312	Total 2454	C 1554	N 428	O 460	S 12	0	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

Note EDS was not executed.

- Molecule 1: DNA (5'-D(*TP*GP*GP*GP*GP*AP*CP*TP*TP*TP*CP*C)-3')

Chain C:
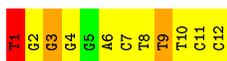

- Molecule 2: DNA (5'-D(*AP*GP*GP*AP*AP*AP*GP*TP*CP*CP*CP*C)-3')

Chain D:
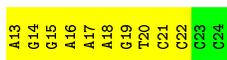
- Molecule 3: PROTEIN (NF-KAPPA B P65 SUBUNIT)

Chain A:
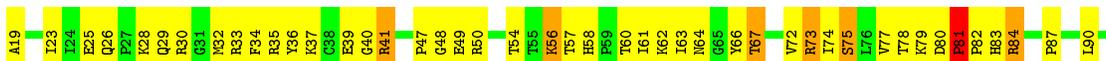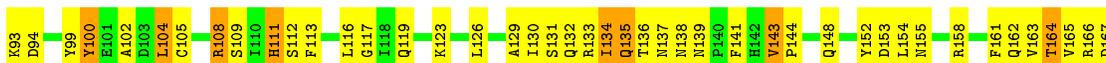
- Molecule 4: PROTEIN (NF-KAPPA B P50 SUBUNIT)

Chain B:

H407 A408 H409 S410 K414 H415 C416 E417 D418 G419 V420 C421 T422 A425 G426 P427 M430 V431 V432 G433 F434 A435 N436 L437 G438 I439 K444 V447 F448 L451 R454 M455 T456 E457 A458 C459 I460 R461 G462 Y463 N464 P465 G466 L467 L468 V469 H470 S471 D472 L473 A474 Y475 L476

Q477 A478 E479 G480 G481 R484 Q485 L486 T487 D488 R489 E490 K491 E492 I493 I494 R495 Q496 A497 A498 V499 Q500 Q501 T502 K503 E504 M505 D506 L507 R511 L512 M513 F514 T515 A516 L517 L518 P519 D520 S524 F525 T526 R527 R528 L529 E530 P531 V532 V533 S534 D535 A536 I537 S540 K541 A542

P543 N544 A545 S546 N547 L548 K549 I550 V551 R552 M553 D554 R555 C559 V560 T561 E564 E565 I566 Y567 L568 L569 C570 D571 K572 V573 Q574 K575 D576 D577 I578 Q579 I580 R581 F582 Y583 E584 E585 E586 E587 N588 G589 G590 V591 W592 E593 G594 D597 F598 S599 P600 T601 D602 V603 H604 R605 Q606

F607 V610 F611 K612 T613 F614 K615 Y616 K617 D618 V619 N620 I621 T622 K623 P624 A625 S626 V627 F628 V629 Q630 L631 R632 R633 L637 E638 T639 L646 Y647 Y648 P649 E650

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	106.61Å 106.61Å 206.56Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 2.90	Depositor
% Data completeness (in resolution range)	(Not available) (8.00-2.90)	Depositor
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, R_{free}	0.208 , 0.320	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5116	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

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	C	0.95	0/271	1.05	3/417 (0.7%)
2	D	0.84	0/273	0.86	0/419
3	A	0.69	0/2228	0.94	4/3021 (0.1%)
4	B	0.66	1/2506 (0.0%)	0.90	1/3384 (0.0%)
All	All	0.70	1/5278 (0.0%)	0.93	8/7241 (0.1%)

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	C	0	3
3	A	0	1
4	B	0	2
All	All	0	6

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	385	CYS	CB-SG	-5.05	1.73	1.81

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	1	DT	C1'-O4'-C4'	-7.32	102.78	110.10
1	C	1	DT	C3'-C2'-C1'	-7.22	93.83	102.50
1	C	1	DT	O4'-C1'-N1	6.69	112.68	108.00
3	A	262	LEU	CA-CB-CG	5.80	128.64	115.30
3	A	135	GLN	N-CA-C	-5.71	95.59	111.00

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	227	TYR	Sidechain
4	B	387	TYR	Sidechain
1	C	1	DT	Sidechain
1	C	3	DG	Sidechain
1	C	9	DT	Sidechain

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	243	0	138	39	0
2	D	243	0	135	42	0
3	A	2176	0	2137	224	0
4	B	2454	0	2449	276	0
All	All	5116	0	4859	554	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 56.

The worst 5 of 554 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)
4:B:487:THR:HB	4:B:490:GLU:HB2	1.26	1.17
2:D:13:DA:H2"	2:D:14:DG:H5"	1.22	1.16
1:C:2:DG:H2"	1:C:3:DG:C5'	1.82	1.10
4:B:465:PRO:HB2	4:B:467:LEU:HD13	1.32	1.08
1:C:2:DG:H2"	1:C:3:DG:H5'	1.35	1.06

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	
3	A	271/273 (99%)	208 (77%)	44 (16%)	19 (7%)	1	4
4	B	310/312 (99%)	241 (78%)	45 (14%)	24 (8%)	1	3
All	All	581/585 (99%)	449 (77%)	89 (15%)	43 (7%)	1	3

5 of 43 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	29	GLN
3	A	40	GLY
3	A	81	PRO
3	A	209	GLY
3	A	231	PRO

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	
3	A	242/242 (100%)	196 (81%)	46 (19%)	2	6
4	B	268/268 (100%)	229 (85%)	39 (15%)	4	12
All	All	510/510 (100%)	425 (83%)	85 (17%)	3	8

5 of 85 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	267	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	B	351	ARG
4	B	581	ARG
3	A	272	LEU
3	A	284	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
4	B	343	GLN
4	B	400	ASN
4	B	544	ASN
3	A	241	GLN
3	A	271	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)

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates (i)

EDS was not executed - this section will therefore be empty.

6.4 Ligands (i)

EDS was not executed - this section will therefore be empty.

6.5 Other polymers (i)

EDS was not executed - this section will therefore be empty.