



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 21, 2020 – 11:54 PM BST

PDB ID : 5W2A  
Title : Structure of human DNA polymerase kappa in complex with Lucidin-derived DNA adduct and incoming dCMPNPP  
Authors : Jha, V.K.; Ling, H.  
Deposited on : 2017-06-06  
Resolution : 2.90 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 : 2.13.1  
buster-report : 1.1.7 (2018)  
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) : 2.13.1

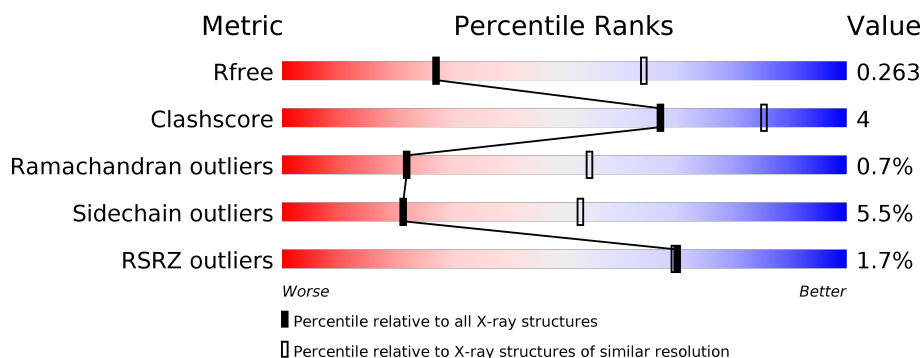
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (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  $\geq 3$ , 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	551	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 8%, green 70%, grey 22%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>70%</span> <span>8%</span> <span>22%</span> </div> </div>
1	B	551	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 2%, orange 1%, yellow 7%, green 70%, grey 22%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>2%</span> <span>70%</span> <span>7%</span> <span>22%</span> </div> </div>
2	C	9	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green 67%, yellow 33%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>67%</span> <span>33%</span> </div> </div>
2	P	9	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green 67%, yellow 33%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>67%</span> <span>33%</span> </div> </div>
3	D	13	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green 77%, yellow 23%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>77%</span> <span>23%</span> </div> </div>
3	T	13	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green 77%, yellow 23%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>77%</span> <span>23%</span> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	GOL	B	603	-	-	-	X
7	PGE	A	604	-	-	-	X
7	PGE	B	605	-	-	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 7735 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 DNA polymerase kappa.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	431	Total	C	N	O	S	0	0	0
			3349	2117	593	620	19			
1	B	431	Total	C	N	O	S	0	0	0
			3314	2101	585	608	20			

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-24	MET	-	initiating methionine	UNP Q9UBT6
A	-23	SER	-	expression tag	UNP Q9UBT6
A	-22	TYR	-	expression tag	UNP Q9UBT6
A	-21	TYR	-	expression tag	UNP Q9UBT6
A	-20	HIS	-	expression tag	UNP Q9UBT6
A	-19	HIS	-	expression tag	UNP Q9UBT6
A	-18	HIS	-	expression tag	UNP Q9UBT6
A	-17	HIS	-	expression tag	UNP Q9UBT6
A	-16	HIS	-	expression tag	UNP Q9UBT6
A	-15	HIS	-	expression tag	UNP Q9UBT6
A	-14	ASP	-	expression tag	UNP Q9UBT6
A	-13	TYR	-	expression tag	UNP Q9UBT6
A	-12	ASP	-	expression tag	UNP Q9UBT6
A	-11	ILE	-	expression tag	UNP Q9UBT6
A	-10	PRO	-	expression tag	UNP Q9UBT6
A	-9	THR	-	expression tag	UNP Q9UBT6
A	-8	THR	-	expression tag	UNP Q9UBT6
A	-7	GLU	-	expression tag	UNP Q9UBT6
A	-6	ASN	-	expression tag	UNP Q9UBT6
A	-5	LEU	-	expression tag	UNP Q9UBT6
A	-4	TYR	-	expression tag	UNP Q9UBT6
A	-3	PHE	-	expression tag	UNP Q9UBT6
A	-2	GLN	-	expression tag	UNP Q9UBT6
A	-1	GLY	-	expression tag	UNP Q9UBT6
A	0	ALA	-	expression tag	UNP Q9UBT6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-24	MET	-	initiating methionine	UNP Q9UBT6
B	-23	SER	-	expression tag	UNP Q9UBT6
B	-22	TYR	-	expression tag	UNP Q9UBT6
B	-21	TYR	-	expression tag	UNP Q9UBT6
B	-20	HIS	-	expression tag	UNP Q9UBT6
B	-19	HIS	-	expression tag	UNP Q9UBT6
B	-18	HIS	-	expression tag	UNP Q9UBT6
B	-17	HIS	-	expression tag	UNP Q9UBT6
B	-16	HIS	-	expression tag	UNP Q9UBT6
B	-15	HIS	-	expression tag	UNP Q9UBT6
B	-14	ASP	-	expression tag	UNP Q9UBT6
B	-13	TYR	-	expression tag	UNP Q9UBT6
B	-12	ASP	-	expression tag	UNP Q9UBT6
B	-11	ILE	-	expression tag	UNP Q9UBT6
B	-10	PRO	-	expression tag	UNP Q9UBT6
B	-9	THR	-	expression tag	UNP Q9UBT6
B	-8	THR	-	expression tag	UNP Q9UBT6
B	-7	GLU	-	expression tag	UNP Q9UBT6
B	-6	ASN	-	expression tag	UNP Q9UBT6
B	-5	LEU	-	expression tag	UNP Q9UBT6
B	-4	TYR	-	expression tag	UNP Q9UBT6
B	-3	PHE	-	expression tag	UNP Q9UBT6
B	-2	GLN	-	expression tag	UNP Q9UBT6
B	-1	GLY	-	expression tag	UNP Q9UBT6
B	0	ALA	-	expression tag	UNP Q9UBT6

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	9	Total	C	N	O	P	0	0	0
			188	88	38	53	9			
2	P	9	Total	C	N	O	P	0	0	0
			188	88	38	53	9			

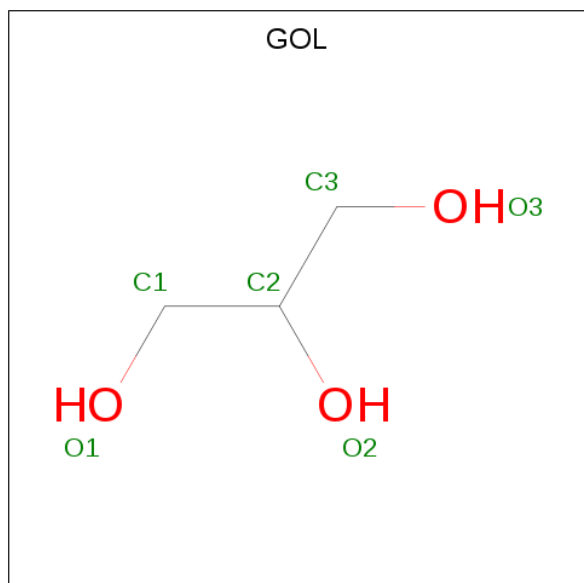
- Molecule 3 is a DNA chain called DNA (5'-D(\*AP\*TP\*GP\*(LDG)P\*CP\*TP\*GP\*AP\*TP\*CP\*CP\*GP\*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	13	Total	C	N	O	P	0	0	0
			283	143	48	80	12			
3	T	13	Total	C	N	O	P	0	0	0
			283	143	48	80	12			

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

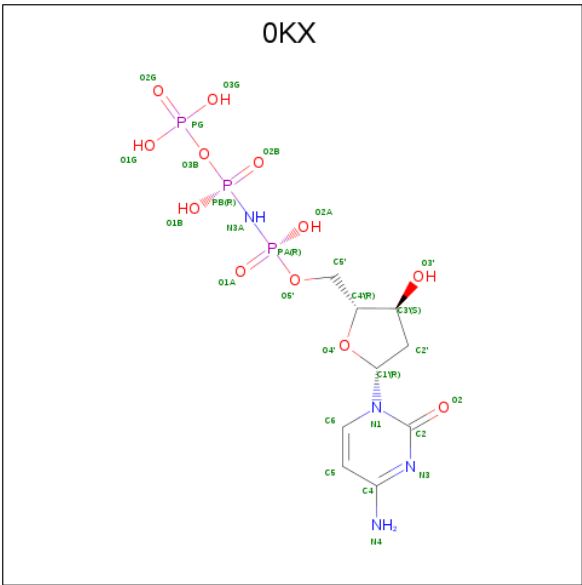
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Mg	0	0
			1	1		
4	A	1	Total	Mg	0	0
			1	1		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



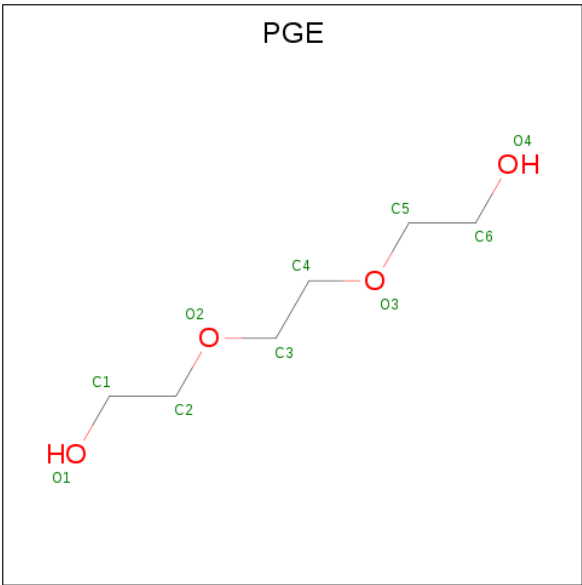
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			6	3	3		
5	B	1	Total	C	O	0	0
			6	3	3		
5	B	1	Total	C	O	0	0
			6	3	3		
5	P	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is 2'-deoxy-5'-O-[(R)-hydroxy{[(R)-hydroxy(phosphonooxy)phosphoryl]amino}phosphoryl]cytidine (three-letter code: 0KX) (formula: C<sub>9</sub>H<sub>17</sub>N<sub>4</sub>O<sub>12</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	A	1	Total	C	N	O	P	0	0
			28	9	4	12	3		
6	B	1	Total	C	N	O	P	0	0
			28	9	4	12	3		

- Molecule 7 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			10	6	4		
7	B	1	Total	C	O	0	0
			10	6	4		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	8	Total 8	O 8	0	0
8	B	14	Total 14	O 14	0	0
8	C	3	Total 3	O 3	0	0
8	D	2	Total 2	O 2	0	0
8	T	1	Total 1	O 1	0	0



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.

- [illegible]

- Chain B:
- 
- 2% 70% 7% 22%
- 0.20  
0.10  
0.00
- 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

- Chain C:  67% 33%




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

Chain P:  67% 33%




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

Chain D:  77% 23%



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

Chain T:  77% 23%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.89Å 129.01Å 167.35Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.25 – 2.90 47.25 – 2.90	Depositor EDS
% Data completeness (in resolution range)	98.9 (47.25-2.90) 99.0 (47.25-2.90)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.68 (at 2.91Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, $R_{free}$	0.231 , 0.270 0.231 , 0.263	Depositor DCC
$R_{free}$ test set	1228 reflections (3.94%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	76.8	Xtriage
Anisotropy	1.008	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 70.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7735	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	112.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.17% 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: GOL, MG, PGE, LDG, 0KX

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	0.45	0/3401	0.71	0/4590
1	B	0.45	0/3367	0.70	1/4548 (0.0%)
2	C	0.45	0/211	0.79	0/324
2	P	0.41	0/211	0.83	0/324
3	D	0.45	0/267	0.80	0/408
3	T	0.48	0/267	0.77	0/408
All	All	0.45	0/7724	0.72	1/10602 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	155	PHE	CB-CG-CD1	5.46	124.62	120.80

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	3349	0	3342	29	0
1	B	3314	0	3275	28	0
2	C	188	0	101	1	2
2	P	188	0	101	1	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	283	0	134	6	1
3	T	283	0	134	10	2
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	6	0	8	0	0
5	B	12	0	16	0	0
5	P	6	0	8	0	0
6	A	28	0	16	0	0
6	B	28	0	14	0	0
7	A	10	0	14	0	0
7	B	10	0	14	0	0
8	A	8	0	0	0	0
8	B	14	0	0	0	0
8	C	3	0	0	0	0
8	D	2	0	0	0	0
8	T	1	0	0	0	0
All	All	7735	0	7177	58	3

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 4.

The worst 5 of 58 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:500:PHE:HD1	1:A:501:PRO:HA	1.32	0.95
1:B:500:PHE:HD1	1:B:501:PRO:HA	1.34	0.92
1:B:155:PHE:CE1	1:B:156:ILE:HG13	2.21	0.75
1:B:155:PHE:HD1	1:B:156:ILE:N	1.87	0.73
1:A:500:PHE:CD1	1:A:501:PRO:HA	2.22	0.70

All (3) 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)
2:C:5:DG:OP2	3:T:14:DC:O3'[3_555]	1.78	0.42
2:C:5:DG:P	3:T:14:DC:O3'[3_555]	2.07	0.13
3:D:14:DC:O3'	2:P:5:DG:OP1[3_555]	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	427/551 (78%)	402 (94%)	22 (5%)	3 (1%)	22	54
1	B	427/551 (78%)	401 (94%)	23 (5%)	3 (1%)	22	54
All	All	854/1102 (78%)	803 (94%)	45 (5%)	6 (1%)	22	54

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	517	PRO
1	A	145	ARG
1	A	347	LYS
1	B	347	LYS
1	A	479	SER

### 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	353/495 (71%)	334 (95%)	19 (5%)	22	54
1	B	342/495 (69%)	323 (94%)	19 (6%)	21	52
All	All	695/990 (70%)	657 (94%)	38 (6%)	21	53

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	507	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	155	PHE
1	B	500	PHE
1	B	148	VAL
1	B	159	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	212	GLN
1	B	212	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 10 ligands modelled in this entry, 2 are monoatomic - leaving 8 for Mogul analysis.

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	$\# Z  > 2$
7	PGE	A	604	-	9,9,9	0.52	0	8,8,8	0.36	0
5	GOL	P	101	-	5,5,5	0.48	0	5,5,5	0.33	0
5	GOL	B	602	-	5,5,5	0.43	0	5,5,5	0.40	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
7	PGE	B	605	-	9,9,9	0.51	0	8,8,8	0.33	0
6	OKX	B	604	4	26,29,29	1.60	4 (15%)	33,45,45	1.28	4 (12%)
6	OKX	A	603	4	26,29,29	0.85	0	33,45,45	1.23	2 (6%)
5	GOL	B	603	-	5,5,5	0.47	0	5,5,5	0.23	0
5	GOL	A	602	-	5,5,5	0.45	0	5,5,5	0.38	0

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
7	PGE	A	604	-	-	4/7/7/7	-
5	GOL	P	101	-	-	0/4/4/4	-
5	GOL	B	602	-	-	1/4/4/4	-
7	PGE	B	605	-	-	4/7/7/7	-
6	OKX	B	604	4	-	2/16/34/34	0/2/2/2
6	OKX	A	603	4	-	2/16/34/34	0/2/2/2
5	GOL	B	603	-	-	3/4/4/4	-
5	GOL	A	602	-	-	2/4/4/4	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	604	OKX	PB-O1B	-4.04	1.45	1.56
6	B	604	OKX	PA-O2A	-3.83	1.46	1.56
6	B	604	OKX	PB-O2B	3.71	1.52	1.46
6	B	604	OKX	PA-O1A	3.56	1.51	1.46

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	603	OKX	C2-N3-C4	4.26	120.66	116.34
6	B	604	OKX	C2-N3-C4	4.05	120.45	116.34
6	A	603	OKX	N4-C4-N3	2.81	120.93	116.49
6	B	604	OKX	N4-C4-N3	2.66	120.69	116.49
6	B	604	OKX	C2'-C1'-N1	-2.14	109.33	114.27

There are no chirality outliers.

5 of 18 torsion outliers are listed below:

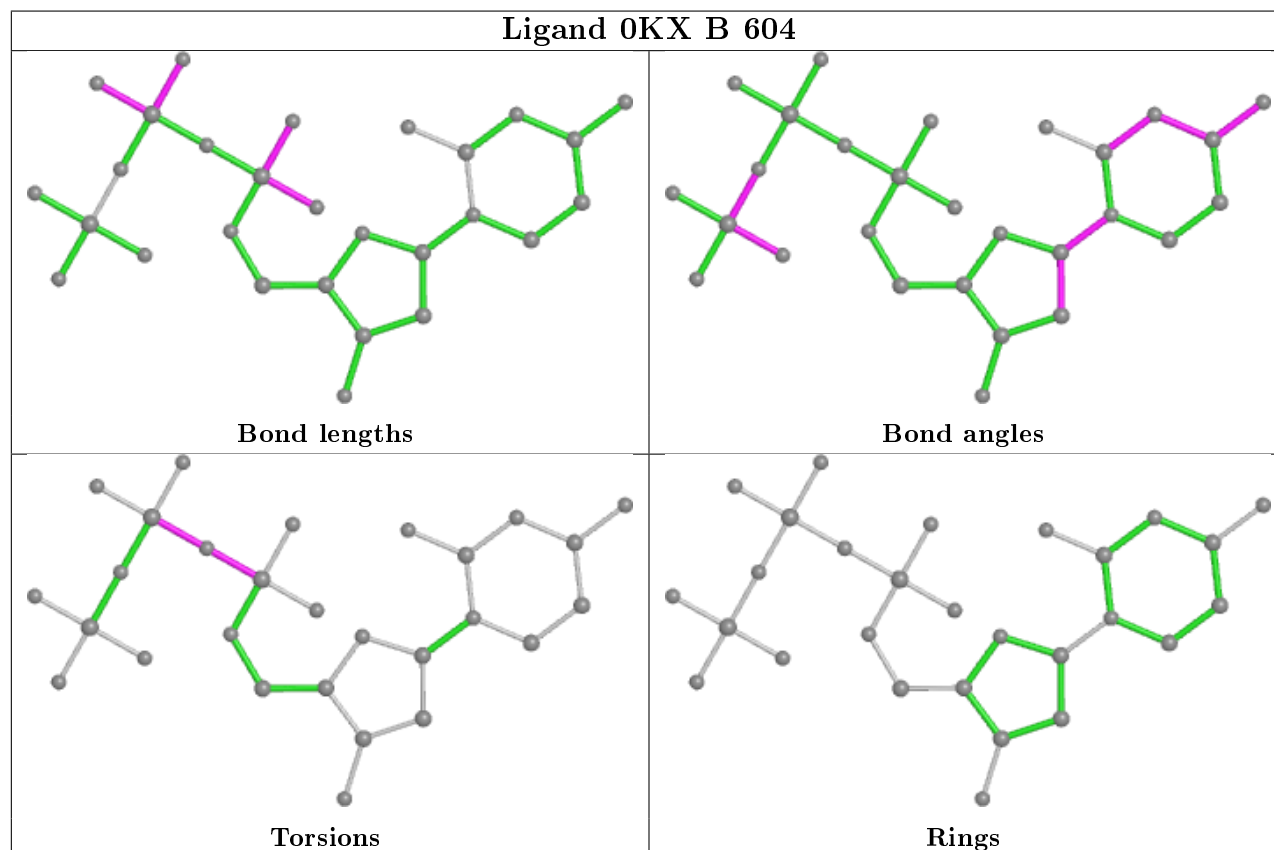


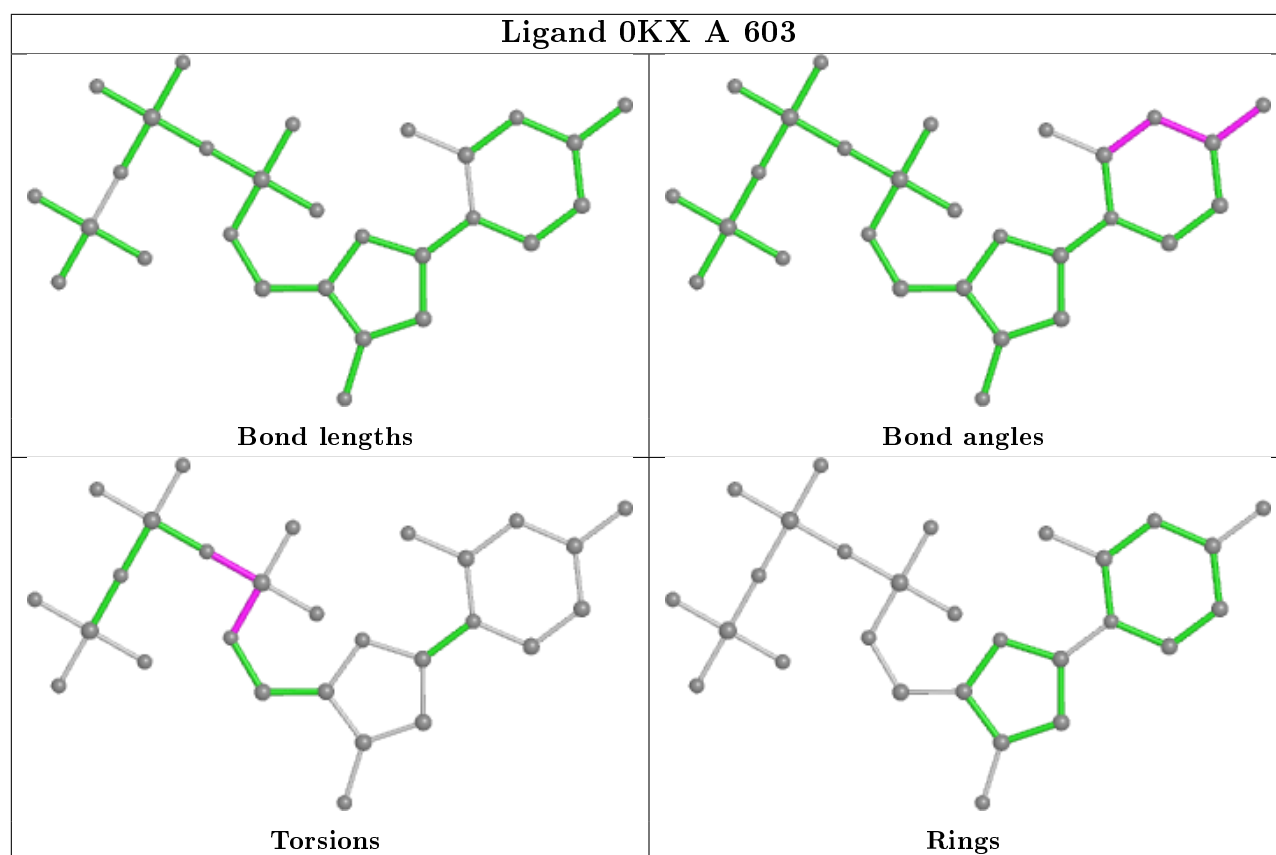
Mol	Chain	Res	Type	Atoms
6	B	604	0KX	PA-N3A-PB-O2B
6	B	604	0KX	PB-N3A-PA-O5'
6	A	603	0KX	PB-N3A-PA-O1A
6	A	603	0KX	C5'-O5'-PA-O2A
5	B	603	GOL	O1-C1-C2-C3

There are no ring outliers.

No monomer is involved in short contacts.

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 than 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.





## 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	431/551 (78%)	0.12	4 (0%) 84 84	75, 110, 142, 188	0
1	B	431/551 (78%)	0.12	11 (2%) 56 52	70, 112, 148, 191	0
2	C	9/9 (100%)	-0.53	0 100 100	92, 99, 118, 124	0
2	P	9/9 (100%)	-0.37	0 100 100	90, 100, 110, 113	0
3	D	12/13 (92%)	-0.20	0 100 100	86, 111, 133, 138	0
3	T	12/13 (92%)	-0.27	0 100 100	76, 108, 124, 127	0
All	All	904/1146 (78%)	0.10	15 (1%) 70 69	70, 111, 145, 191	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	160	LEU	3.9
1	B	400	LEU	3.6
1	B	146	PHE	3.6
1	B	222	PHE	3.3
1	B	31	LEU	2.8

### 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 ⓘ

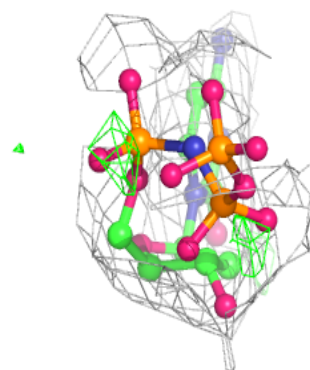
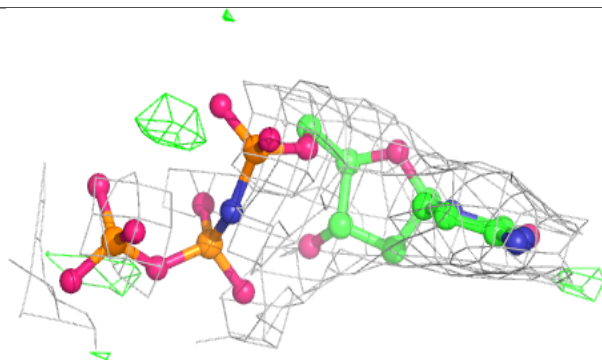
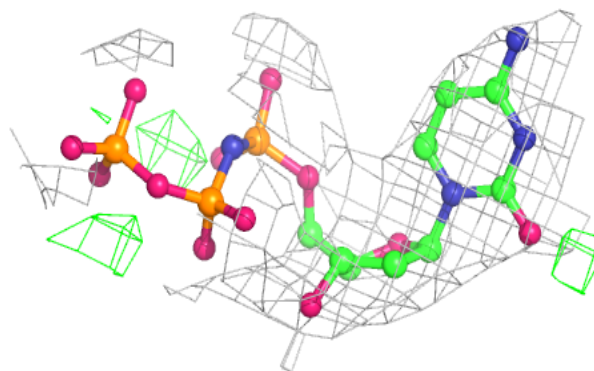
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	GOL	B	602	6/6	0.62	0.23	119,133,137,141	0
7	PGE	A	604	10/10	0.63	0.44	113,134,145,149	0
5	GOL	A	602	6/6	0.67	0.22	113,120,120,121	0
5	GOL	B	603	6/6	0.68	0.47	120,128,133,135	0
7	PGE	B	605	10/10	0.73	0.54	128,135,139,143	0
5	GOL	P	101	6/6	0.78	0.40	91,104,109,110	0
4	MG	A	601	1/1	0.94	0.26	79,79,79,79	0
4	MG	B	601	1/1	0.94	0.23	91,91,91,91	0
6	0KX	A	603	28/28	0.96	0.26	85,93,99,101	0
6	0KX	B	604	28/28	0.97	0.23	83,96,102,105	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

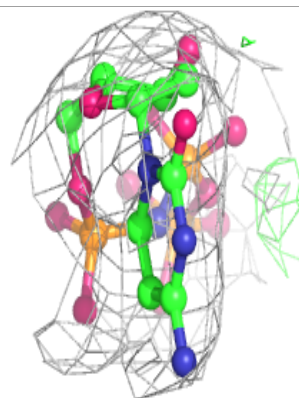
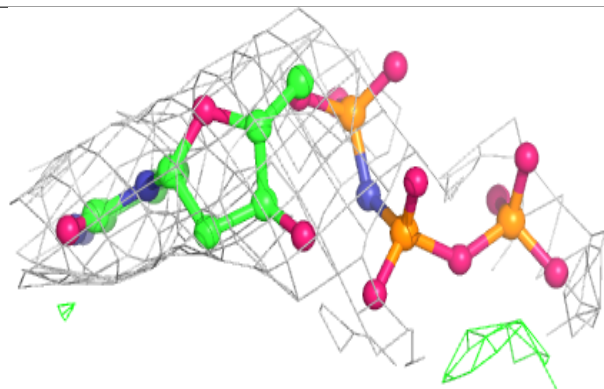
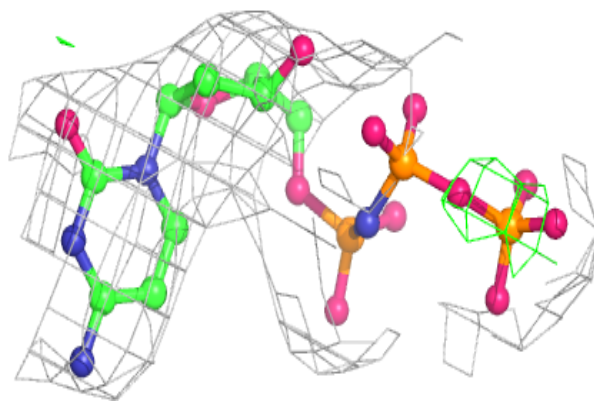
### Electron density around 0KX A 603:

2mF<sub>o</sub>-DF<sub>c</sub> (at 0.7 rmsd) in gray  
mF<sub>o</sub>-DF<sub>c</sub> (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around 0KX B 604:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



## 6.5 Other polymers [i](#)

There are no such residues in this entry.