Biomol NMR Assign
DOI 10.1007/s12104-015-9626-y

ARTICLE

# $^{1}H$, $^{13}C$, and $^{15}N$ backbone resonance assignments of the full-length 40 kDa *S. acidocaldarius* Y-family DNA polymerase, *dinB* homolog

**Sean L. Moro[1] · Melanie J. Cocco[1,2]**

Received: 21 April 2015 / Accepted: 26 June 2015


**Abstract** The *dinB* homolog (Dbh) is a member of the Y-family of translesion DNA polymerases, which are specialized to accurately replicate DNA across from a wide variety of lesions in living cells. Lesioned bases block the progression of high-fidelity polymerases and cause detrimental replication fork stalling; Y-family polymerases can bypass these lesions. The active site of the translesion synthesis polymerase is more open than that of a replicative polymerase; consequently Dbh polymerizes with low fidelity. Bypass polymerases also have low processivity. Short extension past the lesion allows the high-fidelity polymerase to switch back onto the site of replication. Dbh and the other Y-family polymerases have been used as structural models to investigate the mechanisms of DNA polymerization and lesion bypass. Many high-resolution crystal structures of Y-family polymerases have been reported. NMR dynamics studies can complement these structures by providing a measure of protein motions. Here we report the $^{15}N$, $^{1}H$, and $^{13}C$ backbone resonance assignments at two temperatures (35 and 50 °C) for *Sulfolobus acidocaldarius* Dbh polymerase. Backbone resonance assignments have been obtained for 86 % of the residues. The polymerase active site is assigned as well as the majority of residues in each of the four domains.

✉ Melanie J. Cocco
mcocco@uci.edu

[1] Department of Molecular Biology and Biochemistry, University of California, Irvine, CA, USA

[2] Department of Pharmaceutical Sciences, University of California, 1218 Natural Sciences I, Irvine, CA 92697-3900, USA

**Keywords** Y-family polymerase · Bypass polymerase · Translesion synthesis · *Sulfolobus acidocaldarius* · Dbh (*dinB* homolog)

## Biological context

Dbh (*dinB* homolog) is a 354 amino acid DNA polymerase from the thermophilic archaeon *Sulfolobus acidocaldarius* (Boudsocq et al. 2004). Dbh belongs to the Y-family of translesion (TLS) DNA polymerases that can accurately bypass a variety of damaged DNA templates (Yang and Woodgate 2007). Members of the Y-family include Dpo4 from *Sulfolobus solfataricus,* pol IV [*E. coli*, (Wagner et al. 1999)] mammalian pol κ (Ogi et al. 1999; Ohashi et al. 2000; Gerlach et al. 2001), pol η (Johnson et al. 1999a, b), pol ι (McDonald et al. 1999; Tissier et al. 2000), and Rev1 (Lin et al. 1999; Wiltrout and Walker 2011). In particular, Dbh shares 54 % sequence identity to Dpo4. Members of the Y-family share canonical polymerase architecture consisting of a catalytic core composed of palm, fingers, and thumb domains, and have an additional unique C-terminal domain termed the "wrist", "polymerase-associated domain" (PAD), or "little finger" (LF) domain. Notably, the LF domain is tethered to the catalytic core by a flexible linker and has been found to occupy a variety of conformations (Pata 2010). Y-family polymerases also have a significantly larger active site than that of replicative polymerases, which allows the accommodation of bulky DNA adducts, but increases the error-rate extending primers across from undamaged DNA (Sale et al. 2012). Hence, Y-family polymerases are usually highly regulated in eukaryotes, most often through monoubiquitination of PCNA, to prevent error-prone replication of cellular DNA (Yang et al. 2013). TLS polymerases can allow cancer cells

Published online: 08 July 2015

to tolerate DNA damage caused by chemotherapeutics; for example, pol η can bypass cisplatin adducts (Alt et al. 2007). In this way, TLS polymerases that are upregulated can confer resistance to chemotherapy. Furthermore, loss of translesion synthesis can generate chromosomal abnormalities that can contribute to a cancerous phenotype (reviewed in Lange et al. 2011).

Different Y-family polymerases generate characteristic mutations; in particular, Dbh is highly prone to generating single-base deletions resulting in −1 frameshift mutations in vitro (Potapova et al. 2002). The first crystal structures of the apo form of Dbh were reported in 2001 (Silvian et al. 2001; Zhou et al. 2001), followed by preinsertion binary (Dbh, primer/template DNA), insertion ternary (Dbh, primer/template DNA, incoming dNTP) and postinsertion binary (Dbh, primer +1/template DNA) crystal structures (Wilson and Pata 2008). The binary and ternary structures indicated that Dbh generates single-base deletions by a template-slippage mechanism (Wilson and Pata 2008). Dpo4 also generates single-base deletions at high rates (Kokoska et al. 2002), albeit with a somewhat altered mutational spectrum from Dbh. Intriguingly, the mutational and lesion bypass properties of Dbh can be made to resemble those of Dpo4 by generating chimeras of Dbh/Dpo4 containing the catalytic core of Dbh with the LF domain of Dpo4 (Boudsocq et al. 2004), Dbh core/LF with the linker region of Dpo4 (residues 231–245) (Wilson et al. 2013), or even merely three residues of the Dpo4 linker (R242,K243,S244) substituted for residues (K243,I244,P245) of Dbh (Mukherjee et al. 2014). The crystal structures of the Dbh chimeras (Boudsocq et al. 2004; Wilson et al. 2013; Mukherjee et al. 2014) reveal that the orientation of the LF domain with respect to the core resembles that of the LF in the Dpo4 crystal structure (Ling et al. 2001). This observation raises the question of how the position, flexibility, and dynamics of the LF domain affects substrate alignment in the active site of Dbh. Here we report the $^{15}$N, $^{1}$H, and $^{13}$C backbone assignments of Dbh, which will be used to interpret residue specific $T_1$, $T_2$, and $^{15}$N–$^{1}$H NOE relaxation data, and CPMG relaxation dispersion data. The relaxation data will provide valuable information on the mechanism of polymerization of DNA and generation of single-base deletion errors by Dbh.

## Methods and experiments

### Protein expression and purification

The Dbh gene was incorporated into the vector pKKT7-H (a derivative of pKK233, Promega) containing an N-terminal $His_6$ tag (MHHHHHHLVPRGM). Quick-change mutagenesis (Stratagene) was used to change Cys31 to Ser (hereafter referred to as C31S-Dbh) to eliminate potential formation of disulfide bonds. Transfected *E. coli* BL21 cells were grown in 1L Neidhart's minimal media (Neidhardt et al. 1974) at 37 °C containing 1 g $^{15}$N ammonium chloride ($^{15}$N-labeled samples), or 1 g $^{15}$N ammonium chloride, 3 g $^{13}$C glucose, and 80 % $D_2O$ ($^{2}$H, $^{15}$N, $^{13}$C-labeled samples) to ~1.0 OD; expression was induced by the addition of 1 mM IPTG. Protein was expressed for 5 h; subsequently, the cells were harvested by centrifugation and frozen at −80 °C. Dbh or C31S-Dbh were purified from cell lysate by Ni–NTA affinity chromatography under native conditions, and then dialyzed into buffer (20 mM HEPES, 100 mM NaCl, 50 μM EDTA, 50 μM $NaN_3$, pH 7.5) at 4 °C, then one change of buffer without EDTA (20 mM HEPES, 100 mM NaCl, 50 μM $NaN_3$, pH 7.5). To prepare the NMR samples, Dbh or C31S-Dbh protein was concentrated to at least 0.5 mM, and transferred into a Shigemi NMR tube. $D_2O$ was added to the sample for a final concentration of 10 % v/v. Since polymerase enzymes use aspartic acid side chains to coordinate $Mg^{2+}$ at the active site, we were particularly interested in assigning the Asp groups. $^{15}$N-HSQC spectra of selectively $^{15}$N labeled (Asp) and un-labeled (Asn, Arg, Gly, Lys, His, Met, Ser) samples were used to confirm amino acid identity within the sequence. Selectively labeled $^{15}$N-Asp C31S-Dbh was prepared using the *E. coli* auxotroph strain EA1, which is unable to convert Asp to Asn (Muchmore et al. 1989), and by supplementing 1L of Neidhart's media with 100 mg $^{15}$N-Asp. Selectively unlabeled Asn, Arg, Gly, Lys, His, Met, Ser C31S-Dbh samples were prepared using 1 g $^{15}$N ammonium chloride per L of Neidhart's minimal media, BL21 cells, and by supplementing with 0.5 g of each $^{14}$N amino acid separately. Since the HNCO experiment on C31S-Dbh was by far the most sensitive, an HNCO spectrum of a selective $^{13}$C'-Leu, fully $^{15}$N enriched sample was used to confirm amide resonances preceded by leucine residues. $^{13}$C'-Leu was incorporated by expressing the protein in BL21 cells and adding 150 mg of $^{13}$C'-labeled Leu to 1 L of Neidhart's medium containing 1 g $^{15}$N ammonium chloride.

### NMR experiments

NMR data were acquired at 35 and 50 °C on a Varian INOVA 800 MHz NMR spectrometer equipped with a 5-mm triple resonance XYZ-gradient probe. The chemical shifts were referenced using 2,2-dimethyl-2-silapentane-sulfonic acid (DSS). All spectra were processed using NMRPipe/NMRDraw (Delaglio et al. 1995) and analyzed using CcpNmr Analysis (Vranken et al. 2005). A set of 3D triple resonance experiments, including HNCO, HN(CA)CO, HN(CO)CA, and HNCACB were carried out using TROSY (Pervushin et al. 1997) for the sequential backbone resonance assignment (Kay 1997). In addition,

$^{15}$N-edited NOESY-HSQC spectra were also used to confirm resonance assignments.

## Assignment and data deposition

Figure 1 displays the $^{15}$N-HSQC of the full length C31S-Dbh protein (354 amino acids) with a N-terminal hexahistidine tag. Complete or partial backbone resonance assignments have been obtained for 86 % (306 of 354) residues in C31S-Dbh, and 81 % of amide resonances (276 of 339 non-proline residues). Twelve additional peaks were visible in the $^{15}$N-HSQC spectrum and the 3D spectra; we were unable to find connectivities for these resonances. In addition, four resonances in the $^{15}$N-HSQC did not have visible corresponding resonances in the 3D data or in the $^{15}$N-edited NOESY spectrum. Figure 2 presents the C31S-Dbh amino acid sequence with assigned residues indicated. Residues 36–38 and the C-terminus (residues 345–354) were disordered in the crystal structures of Dbh [PDB entries 1K1S/1K1Q (Silvian et al. 2001) and 3BQ0 (Wilson and Pata 2008)]; unfortunately, we were unable to assign the backbone resonances of the regions that were not visible in the crystal structures. If we omit the 13 residues

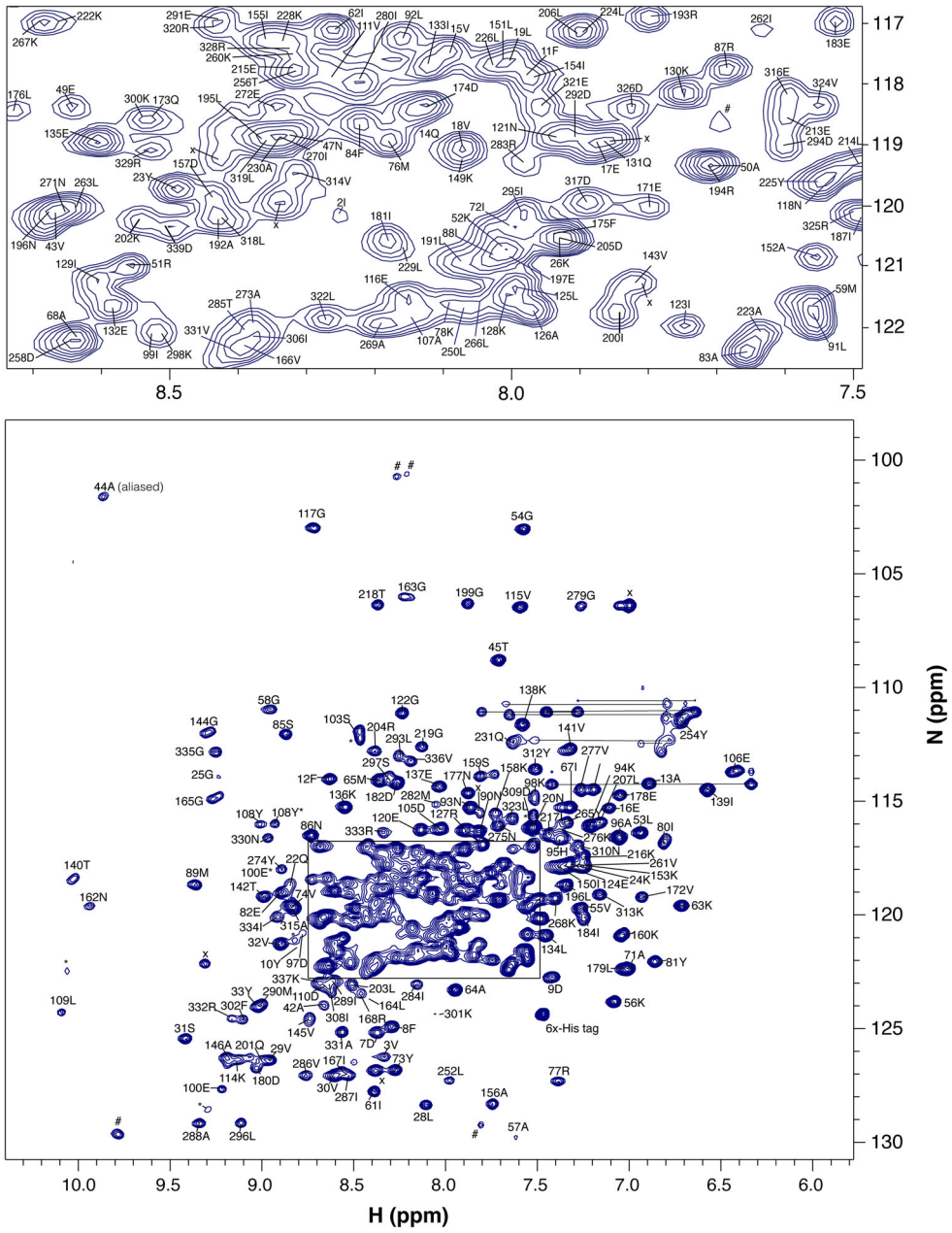


**Fig. 1** 2D $^{15}$N-HSQC spectra of C31S-Dbh, recorded at 800 MHz and 50 °C (323K). The one-letter amino acid code and the residue number indicate resonance assignments. Unassigned resonances are indicated by an "*x*" symbol. Additional conformations of a particular residue are indicated by an *asterisk* "*". Peaks in the HSQC which did not have visible corresponding peaks in the 3D spectra are marked by a *hash* "#". The crowded central region of the spectra is displayed in the *inset* for clarity. 44A is aliased in the $^{15}$N dimension; its true $^{15}$N shift is 133.701 ppm

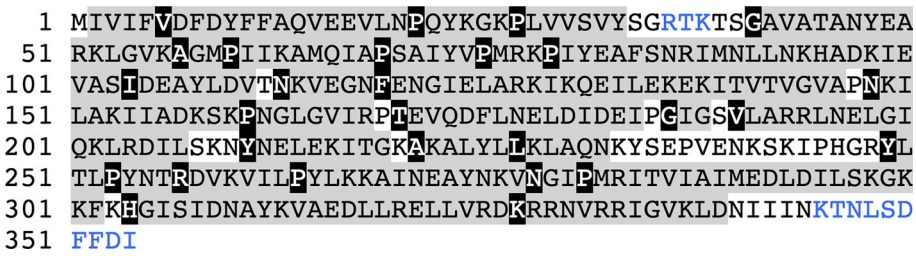


**Fig. 2** C31S-Dbh amino acid sequence with assignments indicated. *Grey shade* is used to indicate residues with at least one backbone atom assigned. *Black background* with *white lettering* indicates residues where only the $^{13}C'$ or $^{13}C^{\alpha}$ were found but not the NH. *White background* represents residues that have not been assigned. In *blue lettering*, amino acids 36–38 and 345–354 are too disordered in crystal structures to be detected [PDB entries 1K1S/1K1Q (Silvian et al. 2001) and 3BQ0 (Wilson and Pata 2008)]. These regions are also not detected by NMR

(36–38; 345–354) that are disordered in the crystal structures of Dbh, we can account for 89 % of residues with at least one assigned backbone resonance. Likewise, most of the linker region (residues 232–245) between the LF and thumb domains was not assigned. The linker and disordered regions from the crystal structure are flexible and solvent exposed. Hence, the signals may be missing due to intrinsic exchange with the solvent. In conclusion, for portions of C31S-Dbh where we would expect rigid structure to enable us to detect NMR signals, we have assigned 94 % of the protein.

A small part of the structure appears to have an alternate conformation; several residues have two corresponding peaks in the $^{15}N$-HSQC spectrum and 3D data (97D, 98K, 100E, 101V, 102A, 103S, 108Y, 109L, and 110D), consistent with slow chemical exchange. These residues are located in the β-sheet structure of the active site palm domain, surrounding the metal ion coordinating residues 105D and 106E.

The mutation of residue 31 from Cys to Ser does not appear to affect the structure. For example, the crystal structures of apo C31S-Dbh [PDB entries 1K1S/1KIQ (Silvian et al. 2001)] superimpose very well with ligand-bound forms of WT Dbh [PDB entries 3BQ0, 3BQ1, and 3BQ2 (Wilson and Pata 2008)]. In addition, this mutation did not significantly affect the NMR spectrum, as the $^{15}N$-HSQC of WT Dbh and C31S-Dbh overlay extremely well (data not shown). Only 11 peaks (9D, 10Y, 12F, 31S, 32V, 45T, 56K, 64A, 77R, 140T, and 301K) were observed to have shifted by any appreciable amount (>0.05 ppm for $^{1}H$ or >0.2 ppm for $^{15}N$) upon mutation. Unsurprisingly, the peaks for residue 31 and adjacent 32V are shifted, and all but two (140T, 301K) of the remaining shifted peaks are located in the same domain (finger) as residue 31.

The backbone resonance assignments of Dpo4 catalytic core (Ma et al. 2010) and LF domain (Ma et al. 2011) at 50 °C have been published. Given the homology between the two proteins (54 % sequence identity) and similar tertiary structure, some of the resonances of the two proteins would be expected to have similar chemical shifts. It should be noted that all of the assignments of C31S-Dbh were completed independently using only our own data; the Dpo4 assignments were compared to those of C31S-Dbh after we completed our assignments. The mutually assigned amide peak positions of identical residues do not correlate very well, with only 62 of 135 (46 %) available $^{1}H$ shifts within 0.2 ppm and 54 of 135 (40 %) available $^{15}N$ shifts within 0.8 ppm. However, the nearest neighbors of a particular residue can have a significant effect on the amide chemical shift, even if the residues are identical. For instance, a neighboring isoleucine residue on the C-terminal side of an amide could influence the $^{1}H$ shift downfield by ~0.2 ppm (Schwarzinger et al. 2001) and the $^{15}N$ shift downfield by almost 5 ppm (Braun et al. 1994; Schwarzinger et al. 2001). Eliminating identical residues from the comparison that do not also have identical neighbors, the correlation between the two sets of shifts is improved: 30 of 54 (56 %) $^{1}H$ shifts within 0.2 ppm and (57 %) 31 of 54 $^{15}N$ shifts within 0.8 ppm. All but one of the $^{1}H$ shifts and one of the $^{15}N$ shifts in the preceding comparison are found in the polymerase core: 29 of 45 (64 %) $^{1}H$ shifts within 0.2 ppm and 30 of 45 (67 %) $^{15}N$ shifts within 0.8 ppm. This is not surprising, since the polymerase core between the two proteins shares 59 % sequence identity, while the LF domains of the two proteins only have 41 % sequence identity. Even though the assignments of Dpo4 were completed on the polymerase core and the LF as separate constructs, the core and LF domains appear to fold independently into roughly the same native structure as found in the full-length protein based on chemical shifts.

The chemical shifts of C31S-Dbh polymerase at 308K and 323K (35 and 50 °C) have been deposited in the BioMagResBank database under accession number 26564 [http://www.bmrb.wisc.edu, (Ulrich et al. 2008)]. The backbone assignments will be used for relaxation dynamics studies of Dbh, and the published crystal structures of Dbh will be used to interpret NMR relaxation data through molecular modeling and molecular dynamics simulations.

**Acknowledgments** We thank Professor Tom Steitz (Yale University) for providing the plasmid containing the Dbh gene and F. W. (Rick) Dahlquist for the gift of *E. Coli* strain EA1. Egest Pone prepared several of the selectively labeled or unlabeled samples. This work was funded by the UC Cancer Research Coordinating Committee-35159.

## References

Alt A, Lammens K, Chiocchini C, Lammens A, Pieck JC, Kuch D, Hopfner KP, Carell T (2007) Bypass of DNA lesions generated during anticancer treatment with cisplatin by DNA polymerase eta. Science 318:967–970

Boudsocq F, Kokoska RJ, Plosky BS, Vaisman A, Ling H, Kunkel TA, Yang W, Woodgate R (2004) Investigating the role of the little finger domain of Y-family DNA polymerases in low fidelity synthesis and translesion replication. J Biol Chem 279:32932–32940

Braun D, Wider G, Wüthrich K (1994) Sequence-Corrected 15 N "Random Coil" Chemical Shifts. J Am Chem Soc 116: 8466–8469

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277–293

Gerlach VL, Feaver WJ, Fischhaber PL, Friedberg EC (2001) Purification and characterization of pol kappa, a DNA polymerase encoded by the human DINB1 gene. J Biol Chem 276:92–98

Johnson RE, Kondratick CM, Prakash S, Prakash L (1999a) hRAD30 mutations in the variant form of xeroderma pigmentosum. Science 285:263–265

Johnson RE, Prakash S, Prakash L (1999b) Efficient bypass of a thymine–thymine dimer by yeast DNA polymerase, Poleta. Science 283:1001–1004

Kay LE (1997) NMR methods for the study of protein structure and dynamics. Biochem Cell Biol 75:1–15

Kokoska RJ, Bebenek K, Boudsocq F, Woodgate R, Kunkel TA (2002) Low fidelity DNA synthesis by a Y family DNA polymerase due to misalignment in the active site. J Biol Chem 277:19633–19638

Lin W, Xin H, Zhang Y, Wu X, Yuan F, Wang Z (1999) The human REV1 gene codes for a DNA template-dependent dCMP transferase. Nucleic Acids Res 27:4468–4475

Ling H, Boudsocq F, Woodgate R, Yang W (2001) Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication. Cell 107:91–102

Ma D, Fowler JD, Yuan C, Suo Z (2010) Backbone assignment of the catalytic core of a Y-family DNA polymerase. Biomol NMR Assign 4:207–209

Ma D, Fowler JD, Suo Z (2011) Backbone assignment of the little finger domain of a Y-family DNA polymerase. Biomol NMR Assign 5:195–198

McDonald JP, Rapic-Otrin V, Epstein JA, Broughton BC, Wang X, Lehmann AR, Wolgemuth DJ, Woodgate R (1999) Novel human and mouse homologs of Saccharomyces cerevisiae DNA polymerase eta. Genomics 60:20–30

Muchmore DC, McIntosh LP, Russell CB, Anderson DE, Dahlquist FW (1989) Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. Methods Enzymol 177:44–73

Mukherjee P, Wilson RC, Lahiri I, Pata JD (2014) Three residues of the interdomain linker determine the conformation and single-base deletion fidelity of Y-family translesion polymerases. J Biol Chem 289:6323–6331

Neidhardt FC, Bloch PL, Smith DF (1974) Culture medium for enterobacteria. J Bacteriol 119:736–747

Ogi T, Kato T Jr, Kato T, Ohmori H (1999) Mutation enhancement by DINB1, a mammalian homologue of the Escherichia coli mutagenesis protein dinB. Genes Cells 4:607–618

Ohashi E, Bebenek K, Matsuda T, Feaver WJ, Gerlach VL, Friedberg EC, Ohmori H, Kunkel TA (2000) Fidelity and processivity of DNA synthesis by DNA polymerase kappa, the product of the human DINB1 gene. J Biol Chem 275:39678–39684

Pata JD (2010) Structural diversity of the Y-family DNA polymerases. Biochim Biophys Acta 1804:1124–1135

Pervushin K, Riek R, Wider G, Wuthrich K (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci USA 94:12366–12371

Potapova O, Grindley ND, Joyce CM (2002) The mutational specificity of the Dbh lesion bypass polymerase and its implications. J Biol Chem 277:28157–28166

Sale JE, Lehmann AR, Woodgate R (2012) Y-family DNA polymerases and their role in tolerance of cellular DNA damage. Nat Rev Mol Cell Biol 13:141–152

Schwarzinger S, Kroon GJ, Foss TR, Chung J, Wright PE, Dyson HJ (2001) Sequence-dependent correction of random coil NMR chemical shifts. J Am Chem Soc 123:2970–2978

Silvian LF, Toth EA, Pham P, Goodman MF, Ellenberger T (2001) Crystal structure of a DinB family error-prone DNA polymerase from Sulfolobus solfataricus. Nat Struct Biol 8:984–989

Tissier A, McDonald JP, Frank EG, Woodgate R (2000) poliota, a remarkably error-prone human DNA polymerase. Genes Dev 14:1642–1650

Ulrich EL, Akutsu H, Doreleijers JF, Harano Y, Ioannidis YE, Lin J, Livny M, Mading S, Maziuk D, Miller Z et al (2008) BioMagResBank. Nucleic Acids Res 36:D402–D408

Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins 59:687–696

Wagner J, Gruz P, Kim SR, Yamada M, Matsui K, Fuchs RP, Nohmi T (1999) The dinB gene encodes a novel *E. coli* DNA polymerase, DNA pol IV, involved in mutagenesis. Mol Cell 4:281–286

Wilson RC, Pata JD (2008) Structural insights into the generation of single-base deletions by the Y family DNA polymerase dbh. Mol Cell 29:767–779

Wilson RC, Jackson MA, Pata JD (2013) Y-family polymerase conformation is a major determinant of fidelity and translesion specificity. Structure 21:20–31

Wiltrout ME, Walker GC (2011) The DNA polymerase activity of Saccharomyces cerevisiae Rev1 is biologically significant. Genetics 187:21–35

Yang W, Woodgate R (2007) What a difference a decade makes: insights into translesion DNA synthesis. Proc Natl Acad Sci USA 104:15591–15598

Yang K, Weinacht CP, Zhuang Z (2013) Regulatory role of ubiquitin in eukaryotic DNA translesion synthesis. Biochemistry 52:3217–3228

Zhou BL, Pata JD, Steitz TA (2001) Crystal structure of a DinB lesion bypass DNA polymerase catalytic fragment reveals a classic polymerase catalytic domain. Mol Cell 8:427–437