Supplementary Material for:

Solution Structure and DNA-binding Properties of the Phosphoesterase Domain of DNA Ligase D

Aswin Natarajan, Kaushik Dutta, Deniz B. Temel, Pravin A. Nair, Stewart Shuman and Ranajeet Ghose

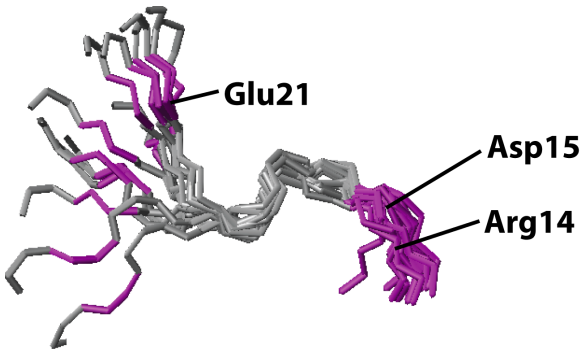

Figure S1. The N-terminus of *Pae*PE is highly disordered. However a small segment between Asp15 and Thr19 is more ordered with slightly higher ^{15}N-{^{1}H} NOE values (0.18 ± 0.01). This segment forms an extended conformation as shown above. The 15 structures of the NMR ensemble were overlaid by fitting the backbone atoms of residues Asp15 through Thr19. The residues (1) implicated in the 3'-phosphatase activity of LigD (Arg14, Asp15 and Glu21) are colored magenta and labeled.

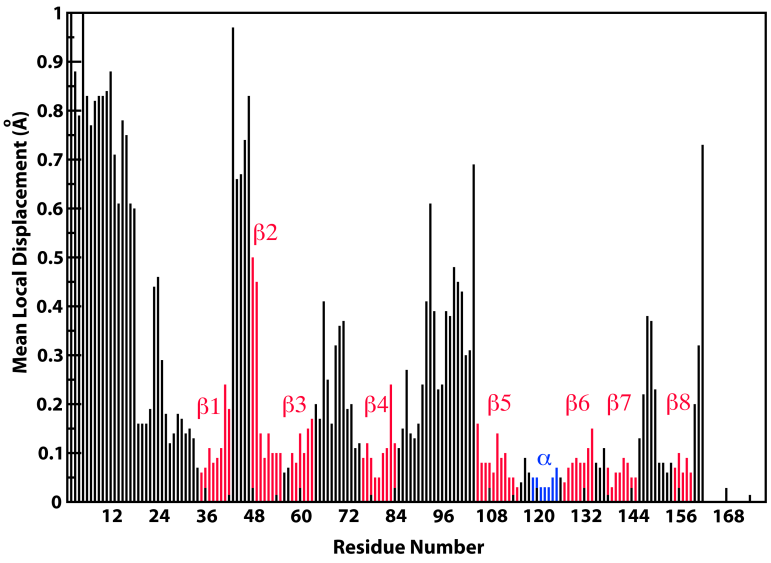

Figure S2. Mean local displacement for backbone residues, calculated using the program MOLMOL (2), in the final NMR ensemble of 15 structures is shown. The β-strands and the α-helix are colored red and blue respectively, and labeled.

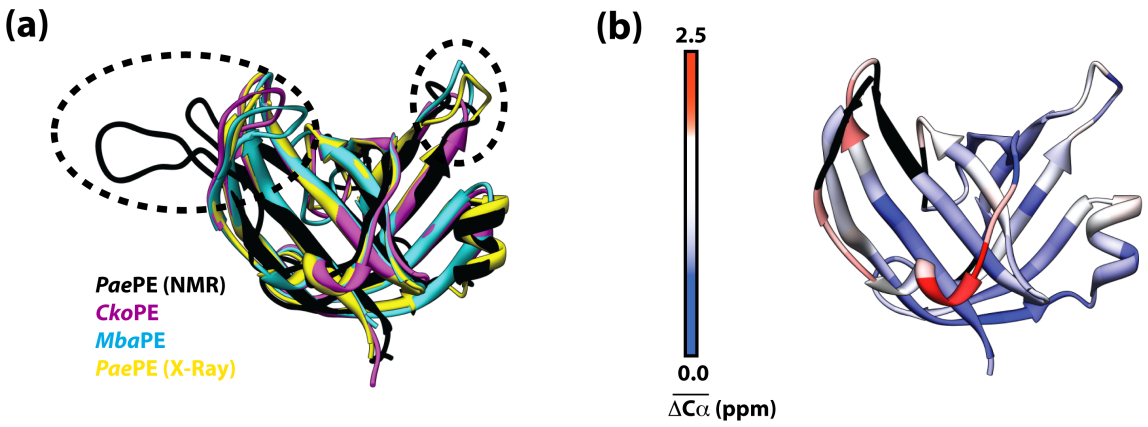

Figure S3. (a) Structures of the PE domains (3) from *Candidatus Korarchaeum cryptofilm* (*Cko*PE, magenta) and *Methanosarcina barkeri* (*Mba*PE, yellow) overlaid on the NMR structure (a representative member of the ensemble of 15 structures closest to the mean, black) and the crystal structure (yellow) of *Pae*PE. The two regions including the β1, β2, β5 face and the β7-β8 loop, discussed at length in the main text are indicated by the dashed ovals. Only the core residues are shown in all cases. **(b)** The absolute differences between the ^{13}Cα shifts measured in solution and those calculated from the crystal structures (PDB: 3N9B, 3N9D) of *Pae*PE (4) using the program SHIFTX2 (5), are depicted on a ribbon representation of the crystal structure (PDB: 3N9B, Chain A) of *Pae*PE using a blue-to-red gradient. Deviations in ^{13}Cα shifts averaged over three residues $\overline{\Delta C\alpha} = (1/3)\left[|\Delta C\alpha(i-1)| + |\Delta C\alpha(i)| + |\Delta C\alpha(i+1)|\right]$ are plotted. Residues for which ^{13}Cα shifts were excluded due to missing assignments or spectral overlap are colored back.

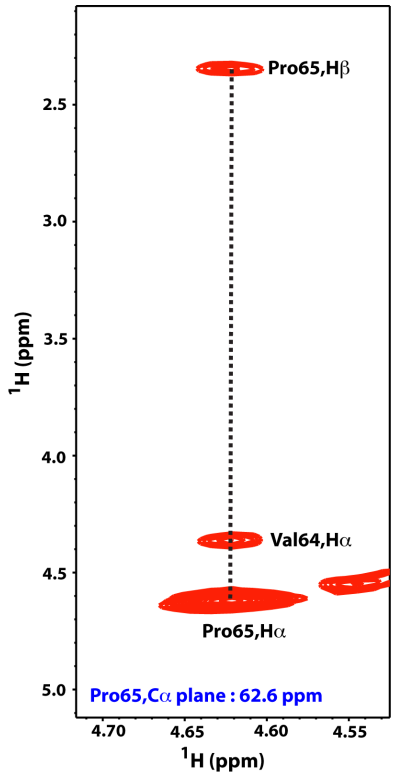

Figure S4. The ^{13}Cα plane for the Pro65 resonance in a ^{13}C-edited NOESY-HSQC (800 MHz) showing a strong Val64, Hα – Pro65, Hα cross-peak indicative of a *cis* Xaa-Pro bond. The difference (corrected for isotope shifts in an uniformly deuterated sample) between the ^{13}Cβ and ^{13}Cγ shifts for this residue (and for Pro72) was 8.4 ppm suggesting a *cis* conformation (6).

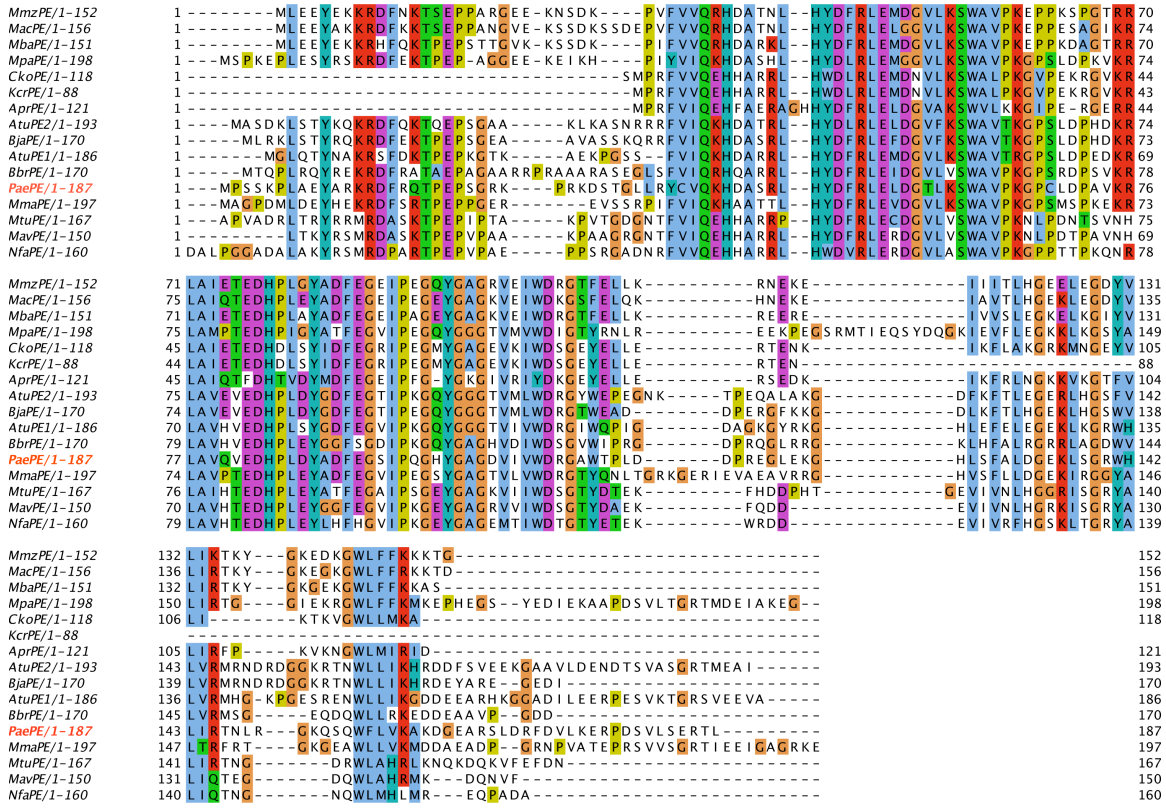

Figure S5. Sequence alignment of 16 bacterial and archaeal PE domains. *Pae*PE is labeled in red font. A Clustal (7) coloring scheme has been used.

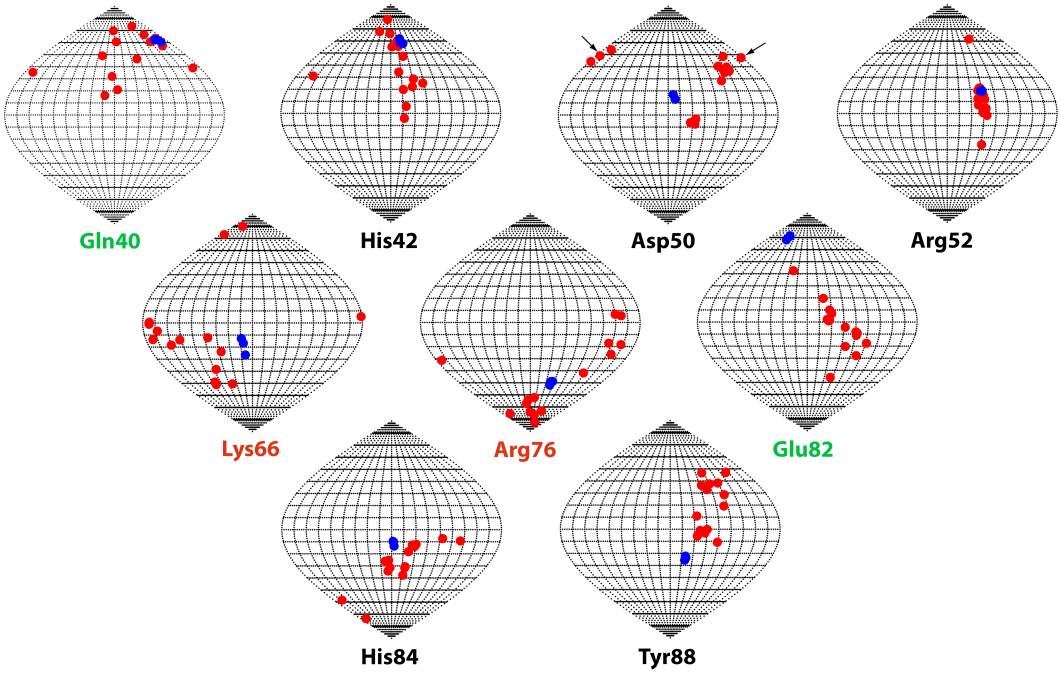

Figure S6. Sanson-Flamsteed (8) projection showing the overall sidechain orientations of key residues implicated in catalysis (1). The orientations have been calculated using the Cβ-Cδ vector for Gln40, the Cβ-Cε1 vector for His42, the Cβ-Cγ vector for Asp50, the Cβ-Cδ vector for Arg52, the Cβ-Cε vector for Lys66, Cβ-Cδ vector for Arg76, the Cβ-Cε1 vector for His84 and the Cγ-Cδ vector for Tyr88. The orientations of the sidechain vectors for each of the 15 structures in the NMR ensemble are shown as red circles. The corresponding orientations in the crystal structures (PDB: 3N9B, two molecules in the asymmetric unit; PDB: 3N9D) (4) are depicted using blue circles. Gln40 and Glu82 affect the 3'-phosphatase activity of LigD (along with Arg14, Asp15 and Glu21 from the disordered N-terminus); Lys66 and Arg76 affect the 3'-ribonucleoside activity; His42, Asp50, Arg52, His84 and Tyr88 play a role in both activities (1). No assignments are available for His48 that also plays a major role in catalysis (affects both 3'-ribonucleoside and 3'-phosphatase activities). Data for this residue are therefore not shown. All orientations have been defined relative to a common arbitrary frame of reference. Note that points at the opposite edges of the projection maps have very similar orientations (see for example, the points for Asp50 indicated by the arrows for which the orientations differ by only 6.7º).

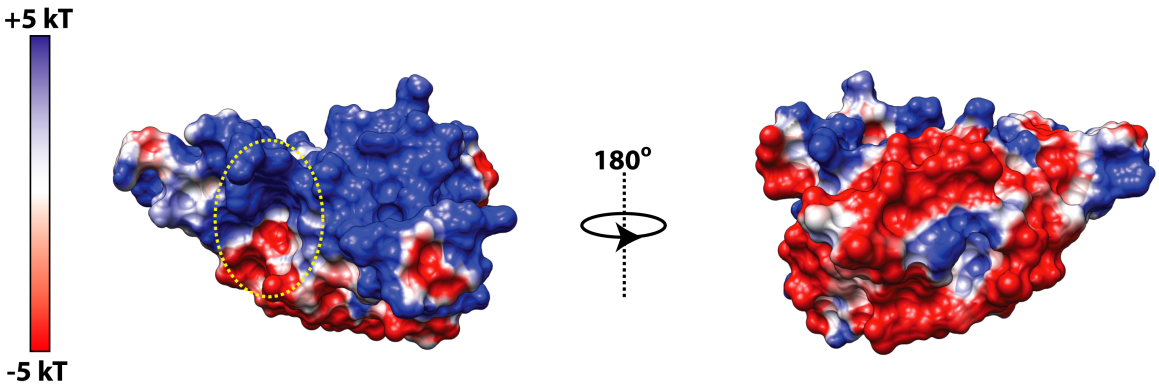

Figure S7. The Coulombic surface of *Pae*PE is shown. The charge distribution was calculated using UCSF Chimera (9) and includes only the core residues. The orientations shown are the same as in Figure 5 in the main text. The face bearing the catalytic residues (location of the catalytic site is indicated by the dotted oval, also see Figure 5a) presents a far more positively charged surface than the opposite face (right panel).

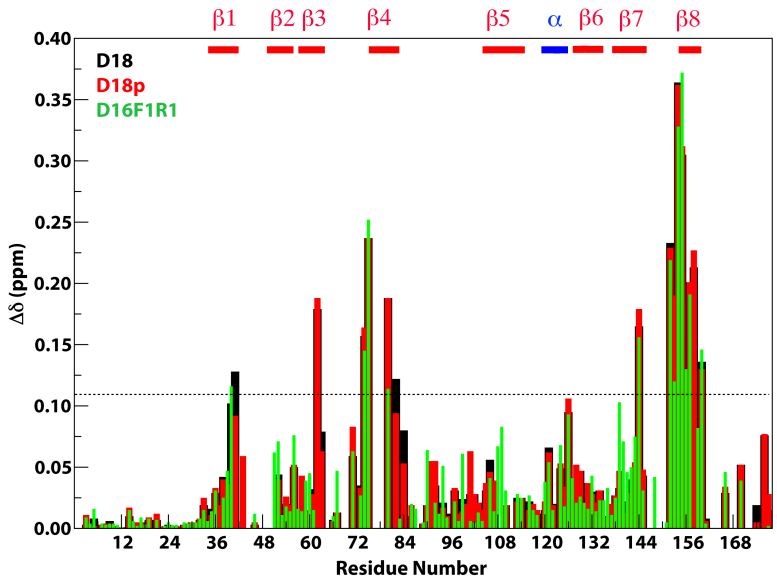

Figure S8. Chemical shift changes seen for *Pae*PE in the presence of each of the three primer-template constructs (D18, D18p and D16F1R1, shown in Figure 4a) at an equimolar protein to oligonucleotide ratio. The largest chemical shift changes in all cases are seen near the catalytic site and around strand β8. Residues that are broadened beyond detection (see Table S1) are not indicated. The 0.11 ppm threshold (used in the discussion in the main text) is indicated by the dotted line.

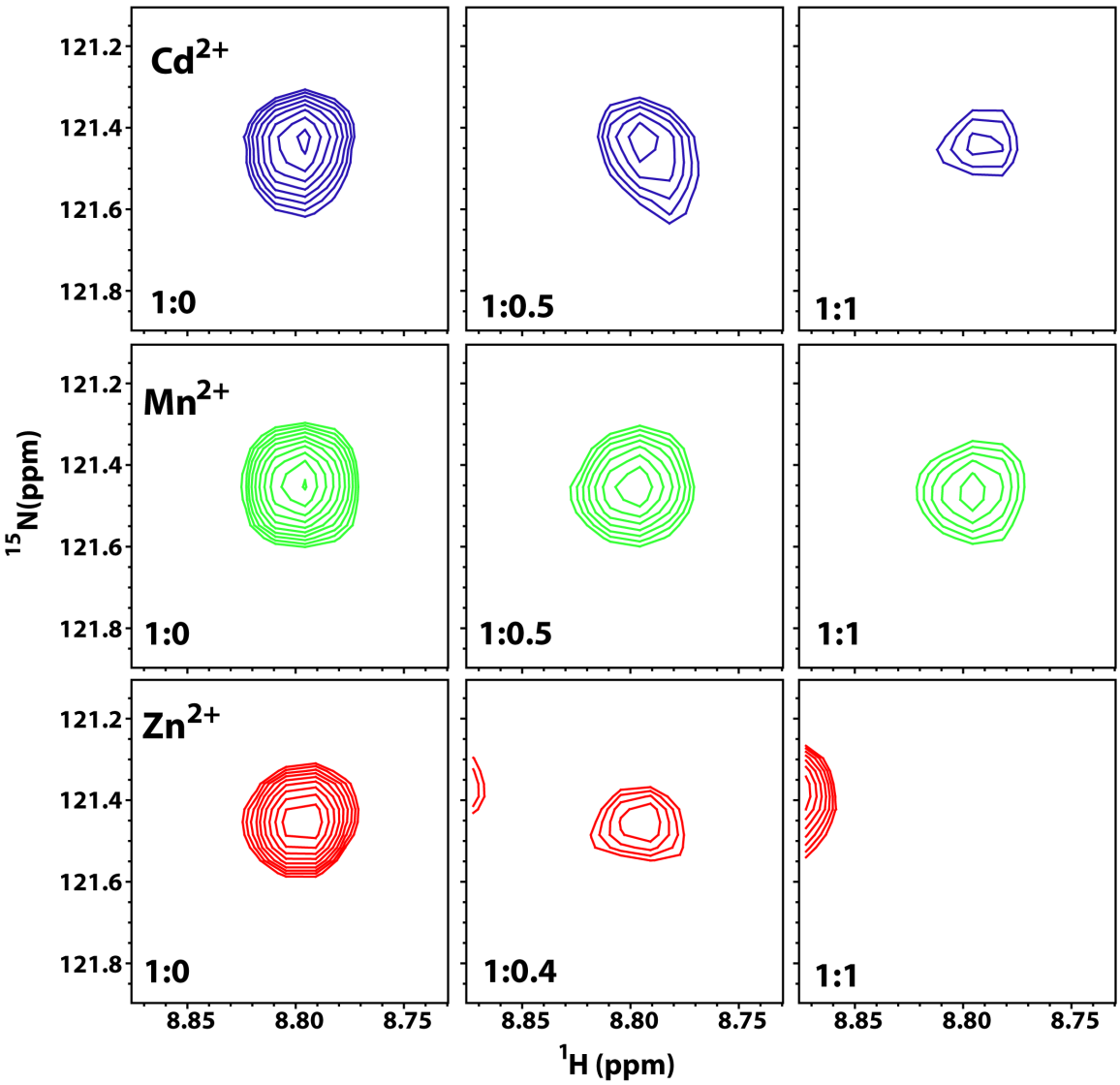

Figure S9. The resonance corresponding to Asp133 illustrates the nature of perturbations seen in $^{15}N,^{1}H$ HSQC spectra of *Pae*PE in the presence of each of Cd^{2+} (blue contours), Mn^{2+} (green contours), or Zn^{2+} (red contours). Typically, limited spectral perturbations were seen for the catalytic Mn^{2+} compared with either Cd^{2+} or Zn^{2+} (that showed the largest perturbations). In addition, spectra in the presence of Zn^{2+}, displayed the appearance of a second set of resonances (See Figure S11) concurrent with the disappearance of a subset of the original peaks during the titration course, suggesting exchange on the slow timescale and high affinity towards Zn^{2+}.

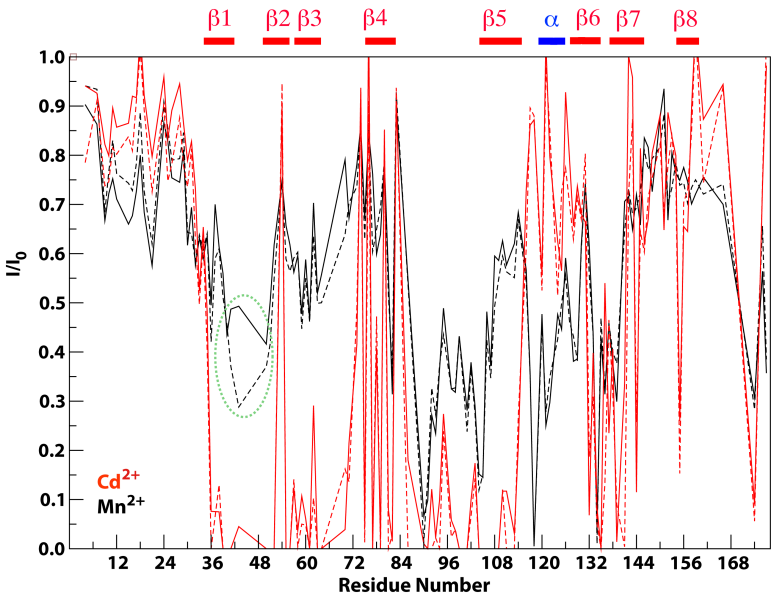

Figure S10. Ratio of resonance intensities (I) in $^{15}N,^{1}H$ HSQC spectra (700 MHz) of *Pae*PE in the presence of an equimolar concentration of bivalent metal ions (Mn^{2+}, black; Cd^{2+}; red) to that in their absence (I_0) is plotted against residue number. Data with (dashed line) and without (solid line) a 10-fold excess of phosphate, are shown. The region near the catalytic site that shows additional changes in the presence of phosphate in the Mn^{2+} case is indicated by the green dotted oval. No other significant differences with and without phosphate were seen in the presence of equimolar amounts of bivalent cations. Peak attenuation patterns for Zn^{2+} ions are qualitatively similar to those for Cd^{2+} ions and are not shown for simplicity.

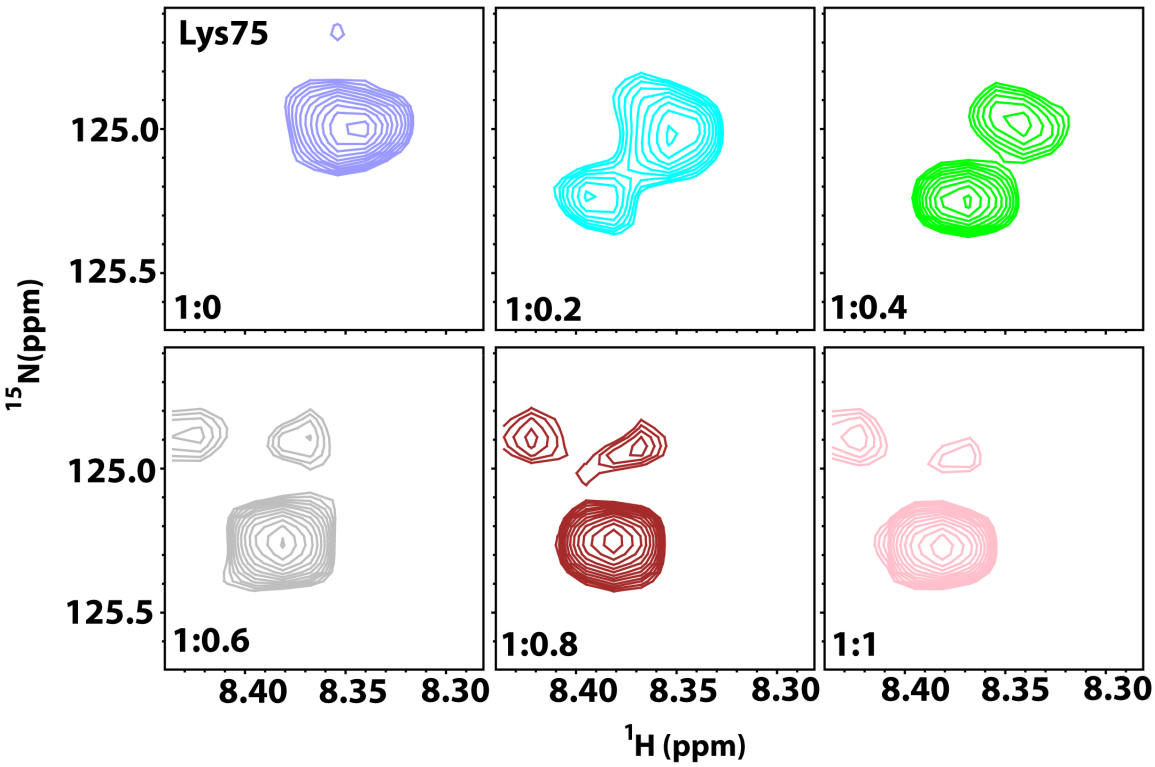

Figure S11. $^{15}N,^{1}H$ HSQC spectra of *Pae*PE displayed the attenuation of specific resonances and the concurrent appearance of a new set of resonances in the presence of increasing amounts of Zn^{2+}. A representative example is shown for the resonance corresponding to Lys75.

Table S1. Residues on *Pae*PAE that are broadened out during NMR titration with the primer-template duplexes[a]

Residue	Secondary Structure	D18p	D18	D16F1R1
Tyr37	β1	-	-	0
Lys41	β1	-	-	3
Asp43	β1–β2 loop	-	4	2
Asp50	β2	2	4	-
Phe51	β2	2[c]	2[c]	-
Trp62	β3	-	-	0[b]
Ala63	β3	-	-	2
Val64	β3–β4 loop	1	1	1
Leu70	β3–β4 loop	1	2	1
Arg76	β4	1	1	3
Leu77	β4	1	1	1
Ala78	β4	1[c]	2[c]	2[c]
Val79	β4	1[b]	4	2[c]
Val81	β4	2	2	0
Glu82	β4	-	-	3
His84	β4	-	-	3
Asp87	β4–β5 loop	3	2	-
Val105	β5	-	-	1
Trp108	β5	1	4	-
Asp109	β5	1	3	-
Gly111	β5	4	3	-
His127	β6	2	2	-
Lys136	β6–β7 loop	2	4	-
Arg140	β7	2	2	-
Ile144	β7	3[c]	-	-
Thr146	β7–β8 loop	1	1	1
Asn147	β7–β8 loop	1	2	2
Leu148	β7–β8 loop	4	4	-
Gly150	β7–β8 loop	1	1	3
Lys151	β7–β8 loop	0	2	-
Trp155	β8	2[b]	-	-
Trp158	β8	-	-	2[b]
Leu174	C-terminus	-	-	2
Glu176	C-terminus	-	-	5

[a] R = (concentration of primer-template construct)/(concentration of protein). **0**: Not seen in the presence of oligo; **1**: not seen at R > 0.05, **2**: not seen at R > 0.1; **3**: not seen at R > 0.2; **4**: not seen at R > 0.3 and **4**: not seen at R > 0.5. Protein concentration was 100 μM in all the titrations. The '-' indicates that the resonance corresponding to the residue was visible throughout the titration series up to R = 1 (an equimolar ratio of protein to oligo). Residues that correspond to the putative non-productive binding loci are shown in red font.

[b] Reappears at an R value of 0.5.

[c] Reappears at an R value of 1.0.

References

1. Zhu, H., Wang, L.K. and Shuman, S. (2005) Essential constituents of the 3'-phosphoesterase domain of bacterial DNA ligase D, a nonhomologous end-joining enzyme. *J. Biol. Chem.*, **280**, 33707-33715.
2. Koradi, R., Billeter, M. and Wuthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.*, **14**, 51-55, 29-32.
3. Smith, P., Nair, P.A., Das, U., Zhu, H. and Shuman, S. (2011) Structures and activities of archaeal members of the LigD 3'-phosphoesterase DNA repair enzyme superfamily. *Nucleic Acids Res.*, **39**, 3310-3320.
4. Nair, P.A., Smith, P. and Shuman, S. (2010) Structure of bacterial LigD 3'-phosphoesterase unveils a DNA repair superfamily. *Proc. Natl. Acad. Sci. USA*, **107**, 12822-12827.
5. Han, B., Liu, Y., Ginzinger, S.W. and Wishart, D.S. (2011) SHIFTX2: significantly improved protein chemical shift prediction. *J. Biomol. NMR*, **50**, 43-57.
6. Dutta, K., Natarajan, A., Nair, P.A., Shuman, S. and Ghose, R. (2011) Sequence-specific ^{1}H, ^{13}C and ^{15}N assignments of the phosphoesterase (PE) domain of *Pseudomonas aeruginosa* DNA ligase D (LigD). *Biomol NMR Assign*, **in press**.
7. Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673-4680.
8. Bugayevskiy, L.M. and Snyder, J.P. (1995) *Map projections:a reference manual*. Taylor and Francis, Philadelphia, PA.
9. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera – a visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605-1612.