Supplemental Data

Human DNA Ligase III Recognizes DNA Ends by Dynamic Switching Between Two DNA Bound States

Elizabeth Cotner-Gohara, In-Kwon Kim, Michal Hammel, John A. Tainer, Alan E. Tomkinson, and Tom Ellenberger

Figure S1. Pre-Step 2 Ligase-DNA Complex

Δ ZnF-LigIII β crystals were grown in the presence of both ATP and divalent cations to promote complete adenylation of the ligase (although the purified ligase is already 85-90% adenylation). The nicked DNA used for crystallization contained a 3' ddC in place of the 3' OH to prevent ligation. Though it was not intentionally blocked, DNA adenylation (Step 2) did not occur during crystal growth. The experimentally phased electron density map does not suggest that a 5'-5'-pyrophosphate linkage has formed between the AMP and the DNA 5'PO₄, whereas a composite OMIT map shows a clear electron density linking the side chain of Lys421 and the AMP phosphate, indicating that the AMP group was not transferred to the DNA and the complex crystallized represents a pre-step 2 intermediate in the DNA end joining reaction.

(A) The composite OMIT electron density map reveals a continuous density between Lys 421 and AMP. (B) The active site of LigIII- Δ ZnF (pink) closely resembles that of the Chlorella virus ligase (ChVLig, blue), which also crystallized in a Pre-Step 2

configuration (Nair et al.). Both LigIII and ChV Lig are covalently bound to AMP via the motif I lysine (Lys421 in LigIII β and Lys27 in ChV Lig). Residue numbers for LigIII are shown. (C) The AMP in an *anti* conformation and the 5'PO₄ of the nick are in similar orientations in both LigIII (pink) and ChV Lig (blue). We confirmed that the Δ ZnF enzyme does not adenylate DNA containing a 5'PO₄/3'ddC when incubated in a concentrated solution of ammonium sulfate mimicking the crystal growth condition (data not shown). The high ionic strength and/or the low pH of the concentrated ammonium sulfate solution may interfere with metal chelation in the active site and or otherwise disfavor DNA adenylation.

Figure S2. Interactions of the Catalytic Core with the Nicked DNA Substrate

The NTase and OBD domains together force the nicked DNA into an underground, distorted structure that widens the minor groove and exposes the nicked ends of the DNA. The OBD inserts into the widened minor groove, helping to position the nick

termini within the active site of the NTase domain. The five conserved motifs (magenta) of the NTase superfamily line the nucleotide binding pocket of the NTase domain (blue).

Figure S3. DBD-NTase Domain Interface Is Less Extensive in LigIII

The DBD-NTase domain interface of ligase III (A) is less extensive than that of ligase I (B). (C) Ligase III (magenta) lacks an extended loop present in the DBD of LigI (orange) that contacts the NTase domain.

Figure S4. Core DNA Interactions Impart Catalytic Specificity

(A) OBD-DNA interactions on the minor groove. Hydrogen bonds are shown as green, dashed lines. (B) Ligation of RNA-containing substrates by LigIIIβ. Ligation activities of full-length LigIIIβ for the substrates containing RNA downstream of the nick (center and right) and normal nicked DNA (left) are compared. Black and red bars indicate the DNA and RNA oligonucleotide, respectively.

Figure S5. SAXS analysis of LigIII β and ΔZnF LigIII

(A) Normalized pair distribution ($P(r)$) functions for the ΔZnF protein and protein-DNA complex reveal a large conformational transition of three domains of ligase III (DBD-NTase-OBD) from an extended to a compact structure. (B) Comparisons of $P(r)$ functions for LigIII β in the presence (black) and absence (green) of DNA show that ZnF-containing ligase III proteins are elongated even when bound to DNA, supporting flexible conformations of the ZnF during end joining.

Supplemental Experimental Procedures

Rigid body modeling – BILBOMD and Minimal Ensemble Search (MES)

Rigid body modeling was performed using the BILBOMD (Pelikan et al., 2009). The crystal structures of LigIII β in complex with DNA (PDB ID: 3L2P) and the solution structure of the ZnF domain (PDB ID: 1UW0) were used to construct the initial atomic models. Regions of the protein missing from the LigIII β crystal structure were manually modeled. These atomic models were used as templates for the rigid body modeling. The program CHARMM, version 33.b (Brooks et al., 1983), with the all-atom CHARMM force field version 27.0 (MacKerell et al., 1998), was used for molecular dynamics (MD) simulation. The initial atomic models were taken as the starting point of the simulations. In all cases, only the atoms of the linkers in between LigIII β domains and the ZnF domain were allowed to move, while the domains were treated as rigid bodies with no

internal motion. The simulations were performed in vacuum with a time step of 1fs. All simulations were performed using the following protocol. The system was subjected to energy minimization with harmonic constraints on the protein atoms, followed by MD heating up to 1500K while keeping the protein atoms fixed. During the subsequent “phase of production” the resulting structures of the protein were recorded every 0.5ps in a trajectory file. For each registered conformation the theoretical SAXS profile, the R_G , D_{max} and the corresponding fit to the experimental data were calculated using the approach developed by Forster et al. (Forster et al., 2008). To select an appropriate minimal ensemble from pool of 10000 MD - generated conformations, Minimal Ensemble Search (MES) was employed (Pelikan et al., 2009). The scattering curves from all the structures in the MD – pool were first computed and the subsequent genetic algorithm-selection operators were performed using these patterns and not the structures (Bernado et al., 2007). The final model was calculated to best fit the experimental curve $I_{exp}(q)$ minimizing the discrepancy χ^2 between the experimental and calculated multiconformational curve

$$\chi^2 = \frac{1}{K-1} \sum_{j=1}^K \left[\frac{\mu I(q_j) - I_{exp}(q_j)}{\sigma(q_j)} \right]^2$$

where K is the number of experimental points, $\sigma(q)$ are standard deviations, and μ is a scaling factor. The BILBOMD modeling and Minimal Ensemble Search (MES) programs are available on the website http://bl1231.als.lbl.gov/saxs_protocols/.

Supplemental References

- Bernado, P., Mylonas, E., Petoukhov, M.V., Blackledge, M., and Svergun, D.I. (2007). Structural characterization of flexible proteins using small-angle X-ray scattering. *Journal of the American Chemical Society* 129, 5656-5664.
- Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., and M., K. (1983). A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *Journal of Computational Chemistry* 4, 187-217.
- Forster, F., Webb, B., Krukenberg, K.A., Tsuruta, H., Agard, D.A., and Sali, A. (2008). Integration of small-angle X-ray scattering data into structural modeling of proteins and their assemblies. *Journal of Molecular Biology* 382, 1089-1106.
- MacKerell, J.A.D., Bashford, D., Bellott, M., Dunbrack, R.L.j., Evenseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., *et al.* (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The Journal of Physical Chemistry B* 102, 3586 -3616.
- Pelikan, M., Hura, G.L., and Hammel, M. (2009). Structure and Flexibility within proteins as identified through small angle X-ray scattering. *General Physiology and Biophysics* 28, 174 - 189.