Electronic Supplementary Material for the manuscript:

Fine tuning of the catalytic activity of colicin E7 nuclease domain by systematic N-terminal mutations

Eszter Németh,[1,2] Tamás Körtvélyesi,[2] Peter W. Thulstrup,[3] Hans E.M. Christensen,[4] Milan Kožíšek,[5] Kyosuke Nagata,[6] Anikó Czene,[7] and Béla Gyurcsik[1,7]*

Table S1. Sequences of the primers applied for the amplification reactions of the genes of NColE7 and its mutants with the restriction endonuclease cleavage sites underlined.

Primer name	**Sequence (5' → 3')**
KGNK-5'	GGAATTCAAAGGGAATAAGCCAGGGAA
KGNG-5'	GGAATTCAAAGGGAATGGGCCAGGGAAGGCAACAGG
GGNK-5'	GGAATTCGGAGGGAATAAGCCAGGGAAGGCA
GGNG-5'	GGAATTCGGAGGGAATGGGCCAGGGAAGGCAACAGG
NColE7-5'	GGAATTCAAACGGAATAAGCCAGGGAAG
Im7-3'	AATCAGCCTCGAGTTAGTGATGGTGATGGTGATGA

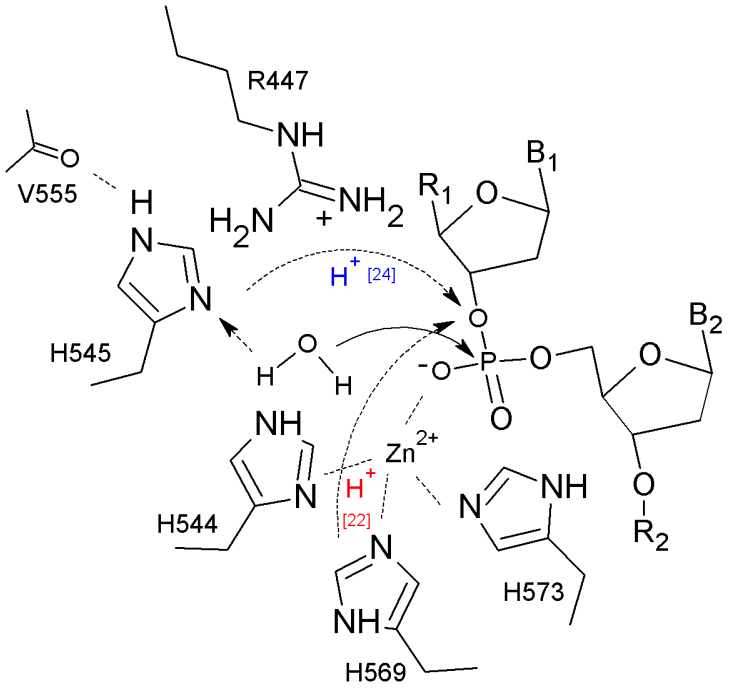


Figure S1. Schematic representation proposed mechanism of DNA-cleavage by NColE7. H545 as a general base deprotonates the water molecule, and the resulted OH performs the nucleophilic attack at the phosphorous atom. Zn^{2+}-ion is playing role in the substrate binding, electrostatic activation and stabilization of the transition state. The source of the proton allowing for protonation of the leaving alcoholate is not identified. Suggestions are a proton channel involving H569, or the shuttle mechanism between H545 and the leaving group.

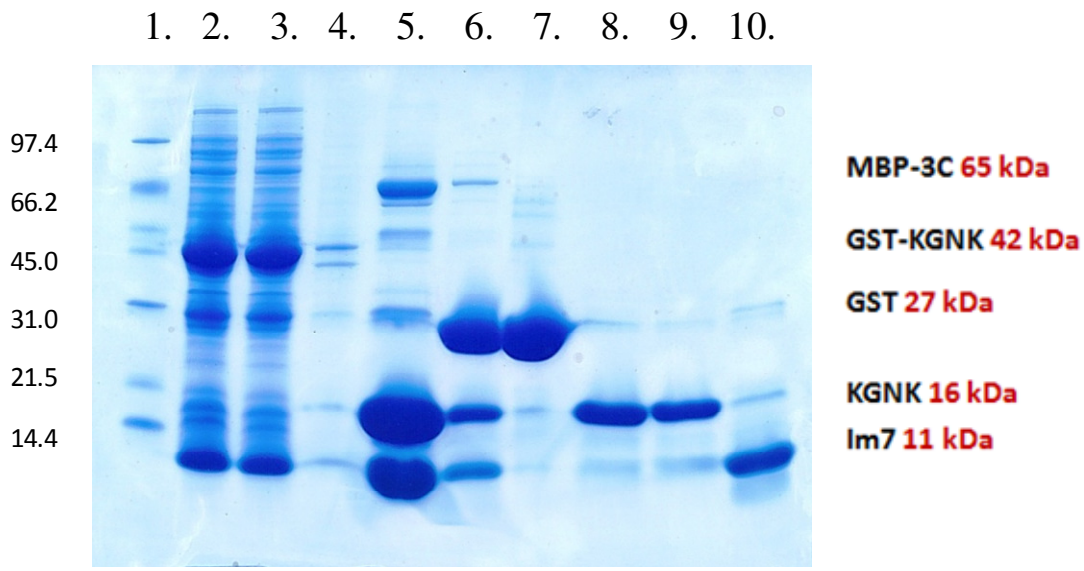
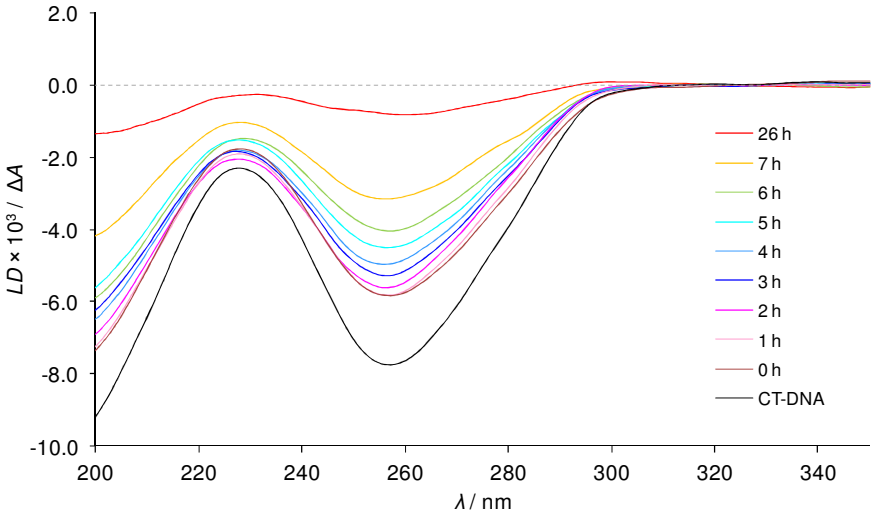


Figure S2. The SDS-PAGE bands of the KGNK mutant protein at different stages of purification. Explanation of the Lane contents: 1. Protein marker (low range, 14-97 kDa, BioRad), 2. Raw extract, 3. Soluble phase of the raw extract, 4. Pellet – insoluble proteins, 5-6. Wash fractions of the GST-affinity chromatography after the cleavage with Human rhinovirus C3 protease, containing the cleaved protein and the protease, 7. Elution fraction of the GST and uncleaved GST-KGNK/Im7, 8-9. Fractions from the cation exchange purification around 1 M NaCl, 10. The immunity protein eluted with PBS containing ~ 2 M NaCl.

A)



B)

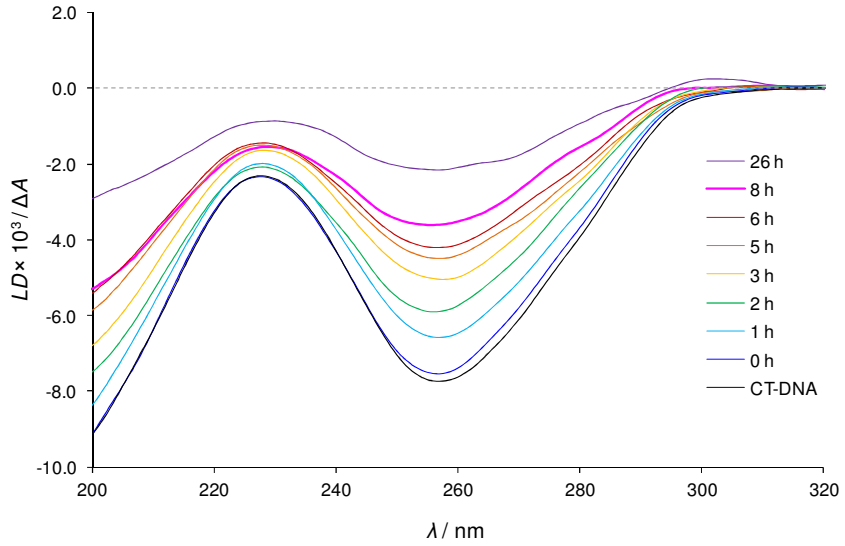


Figure S3. CT-DNA cleavage (130 μM for bp) by KGNG (0.5 μM, **A**) and GGNG (0.5 μM, **B**) in the presence of Zn^{2+} (0.5 μM) monitored by FLD spectroscopy. The baseline corrected spectra are plotted. The first spectrum (denoted by 0 h) was recorded immediately after mixing the reagents. The following spectra were recorded after incubation of the reaction mixture at 37 °C up to 8 hours, and the last spectrum was recorded after 26 hours of incubation.

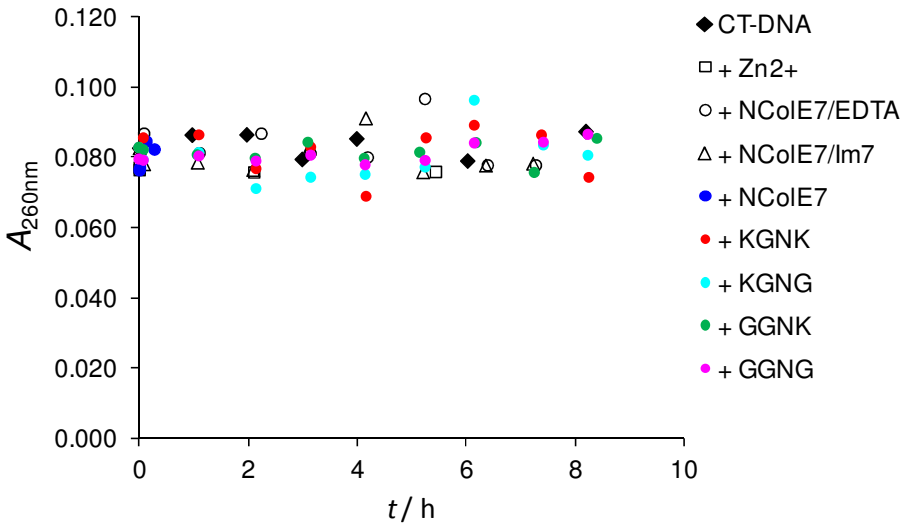
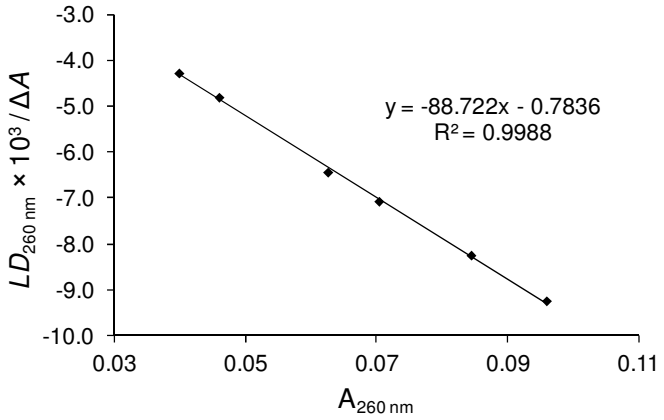


Figure S4. Light absorption at 260 nm during the DNA cleavage reaction. The cleavage of 130 µM (base pairs) CT-DNA by different nucleases (0.5 µM) was followed by flow linear dichroism spectroscopy and the absorption of the samples was obtained by the converted of the HT values. Proteins were pre-incubated for 30 min in the presence of one equivalent $ZnCl_2$. The samples were incubated at 37 °C.

A)



B)

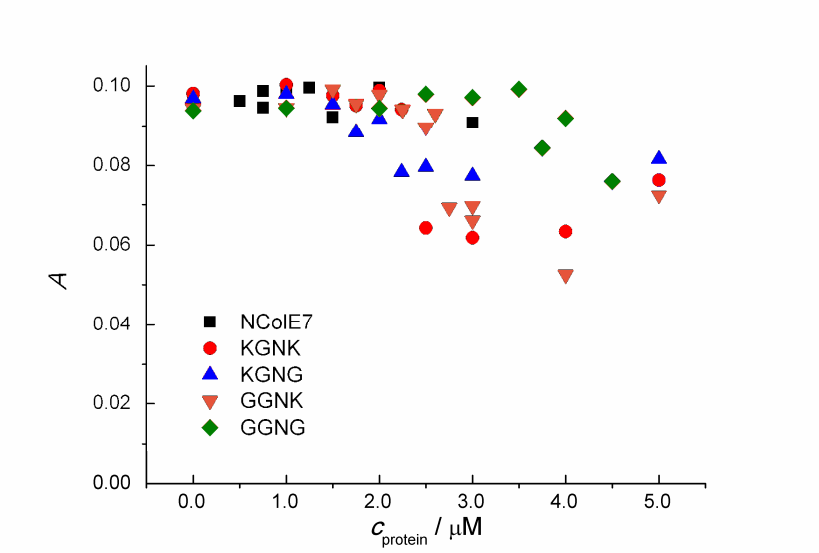


Figure S5. A) Flow linear dichroism calibration curve of CT-DNA in the 55-130 μM concentration range. **B)** Absorption changes during the titration of CT-DNA with NColE7 and its mutants. At high concentrations of DNA, e.g. > 300 μM (bp) and studied protein/DNA ratio (1/100-1/10) the precipitation was observed. The prior addition of one equivalent of Im7 protein to the nuclease prevented this. Thus, the DNA-nuclease complex precipitated. The absorbance measured during LD experiments showed that there was at maximum ~ 20% precipitation at NColE7 variant protein concentrations higher than 2.5 μM. Based on the calibration curve of CT-DNA (Figure S4.a.) 20% decrease in absorbance resulted in loss of 10% of the LD signal. Thus the precipitation did not significantly influence the results.

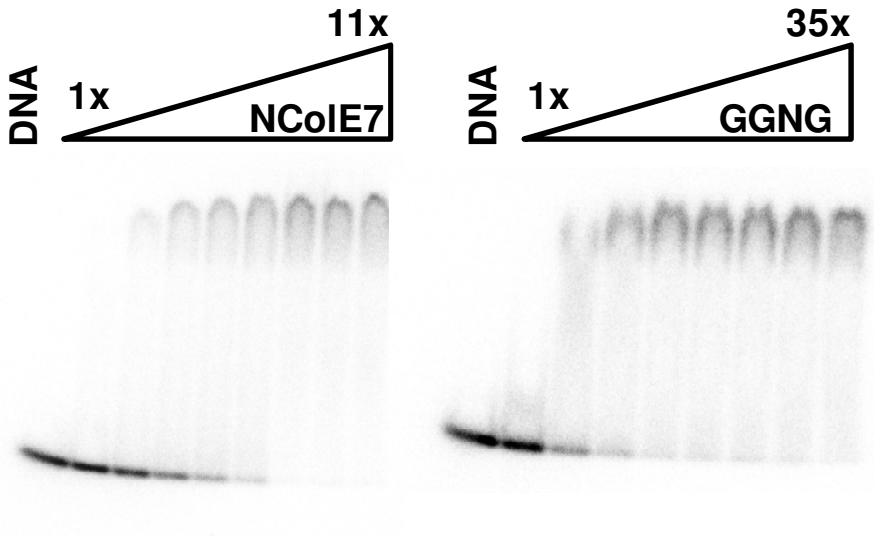
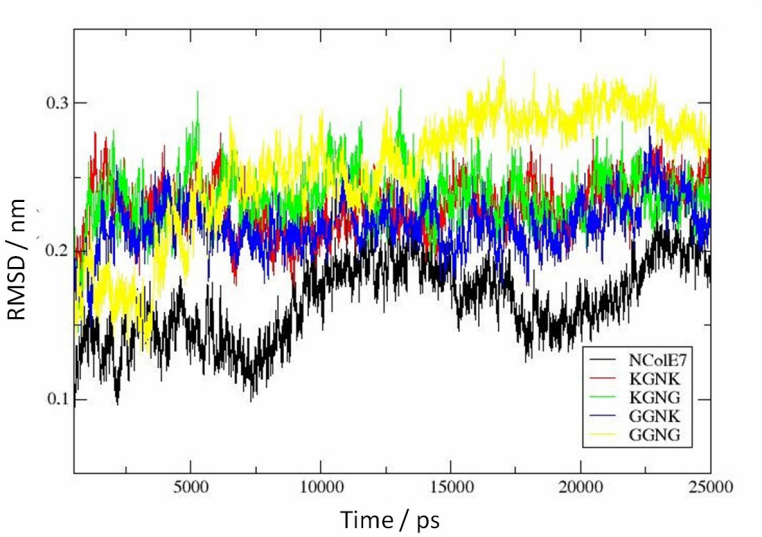


Figure S6. Titration of a radiolabelled 13 bp DNA with **A)** NColE7 and **B)** GGNG. The final concentration of DNA was 0.20 µM in each lanes, while protein concentration was varied as 0, 0.2, 0.4, 0.6, 0.8. 1.0, 1.4, 1.8, 2.2 µM for NColE7 and 0, 0.2, 0.6, 1.0, 1.4, 1.8, 2.2, 4.0, 7.0 µM for GGNG. The samples contained 100 µM EDTA to inhibit nuclease digestion.

A)



B)

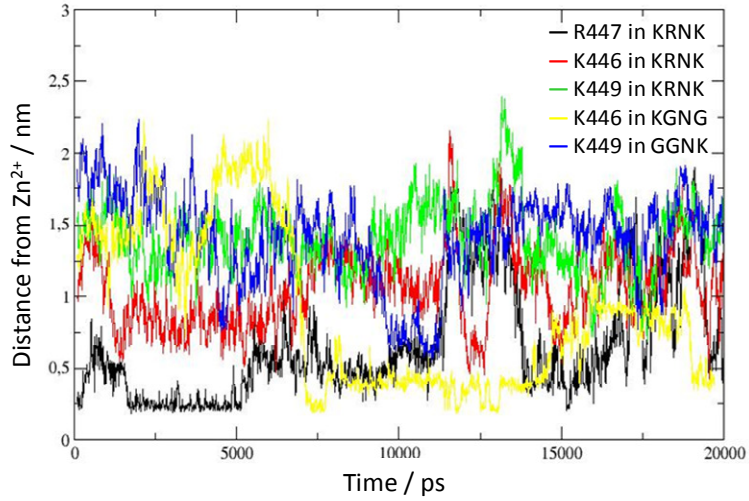


Figure S7. A) RMSD (root mean square deviation) of the studied proteins, as compared to the starting structure of the simulation. **B)** Distance of the positively charged residues from the metal ion in the active centre in the MD simulation of the zinc-bound proteins containing a phosphate ion in the active center. The distance of positive charges from the active center (measured from Zn^{2+}) was analyzed as a function of time. K446 and K449 in NColE7 had an average distance from the Zn^{2+} centre around 1.5 nm, while R447 around 0.5 nm. However, when the arginine was removed, both K446 and to a lesser extent K449 approached the Zn^{2+}-ion supporting that lysines might be able to partially take over the role of the arginine in the function of the protein. This is comparable to the experimental finding that the KGNG and GGNK mutants still had higher activity than the GGNG protein.