

Supplemental Table S1: Data collection and refinement statistics for M.HaeIII C71S / DNA pre-extrusion complex InC.

| Data collection | |
|---|---|
| Space group | P2 ₁ 2 ₁ 2 ₁ |
| Cell dimensions | |
| <i>a</i> , <i>b</i> , <i>c</i> (Å) | 57.56, 129.59, 132.38 |
| α, β, γ (°) | 90, 90, 90 |
| Resolution (Å) | 17-2.5 (2.6-2.5) ^a |
| <i>R</i> _{sym} or <i>R</i> _{merge} (%) ^b | 12.2 (37) |
| <i>I</i> / <i>σI</i> | 7.9 (3.4) |
| Completeness (%) | 90.4 (70.4) |
| Redundancy | 3.4 (2.5) |
| Refinement | |
| Resolution (Å) | 17-2.5 |
| No. reflections | 31673 |
| <i>R</i> _{work} / <i>R</i> _{free} ^c | 18.2/22.1 |
| No. atoms ^d | |
| Protein/DNA | 7845/711 |
| Ligand/ion | 119/3 |
| Water | 580 |
| B-factors | |
| Protein/DNA | 20.0/37.5 |
| Ligand/ion | 51.0/31.4 |
| Water | 26.9 |
| R.m.s deviations | |
| Bond lengths (Å) | 0.002 |
| Bond angles (°) | 0.642 |
| Ramachandran plot (residues, %) | |
| Most favored | 90.4 |
| Additionally allowed | 9.6 |

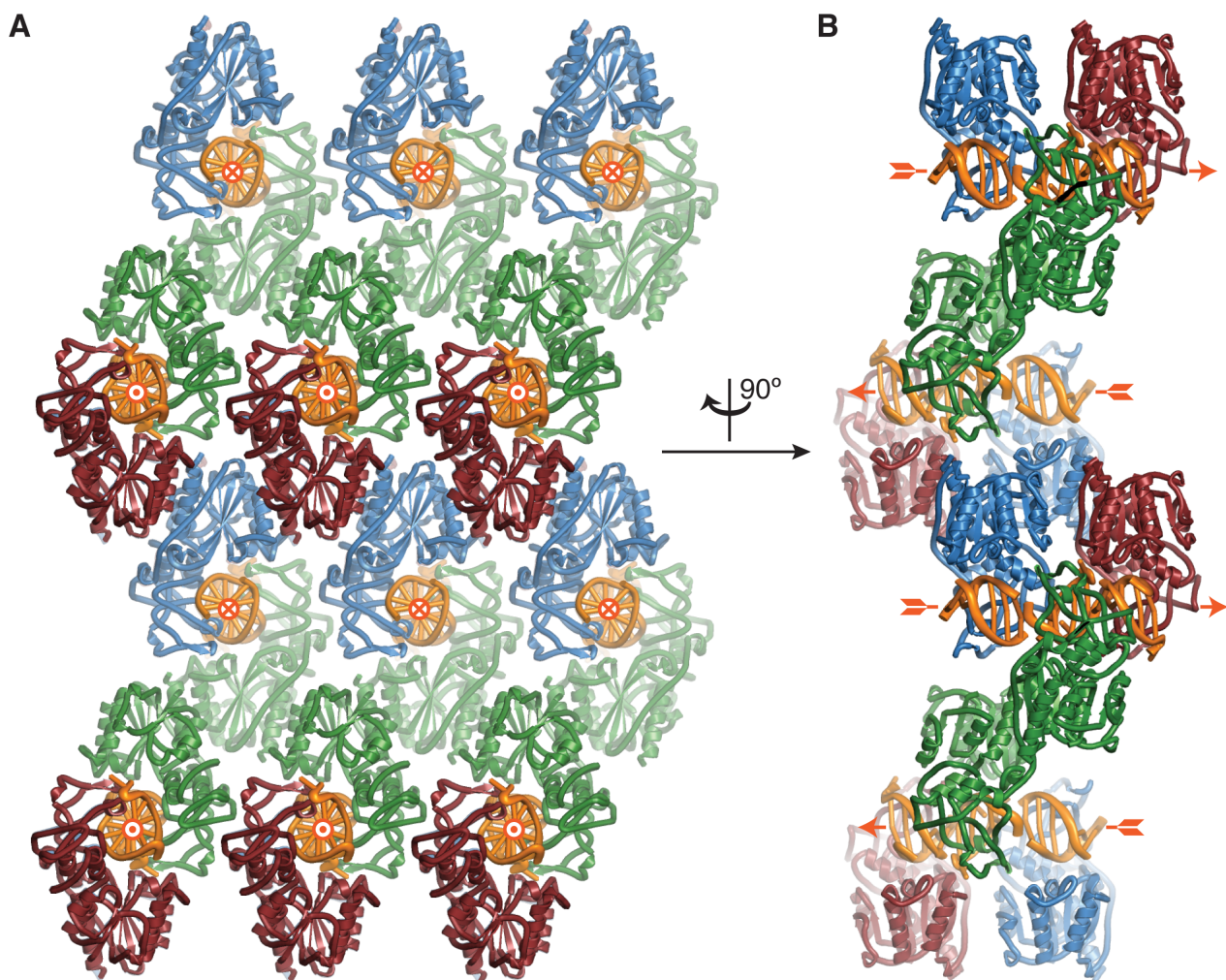
Diffraction data from 6 crystals were merged to produce the dataset.

^a The highest resolution shell is shown in parenthesis.

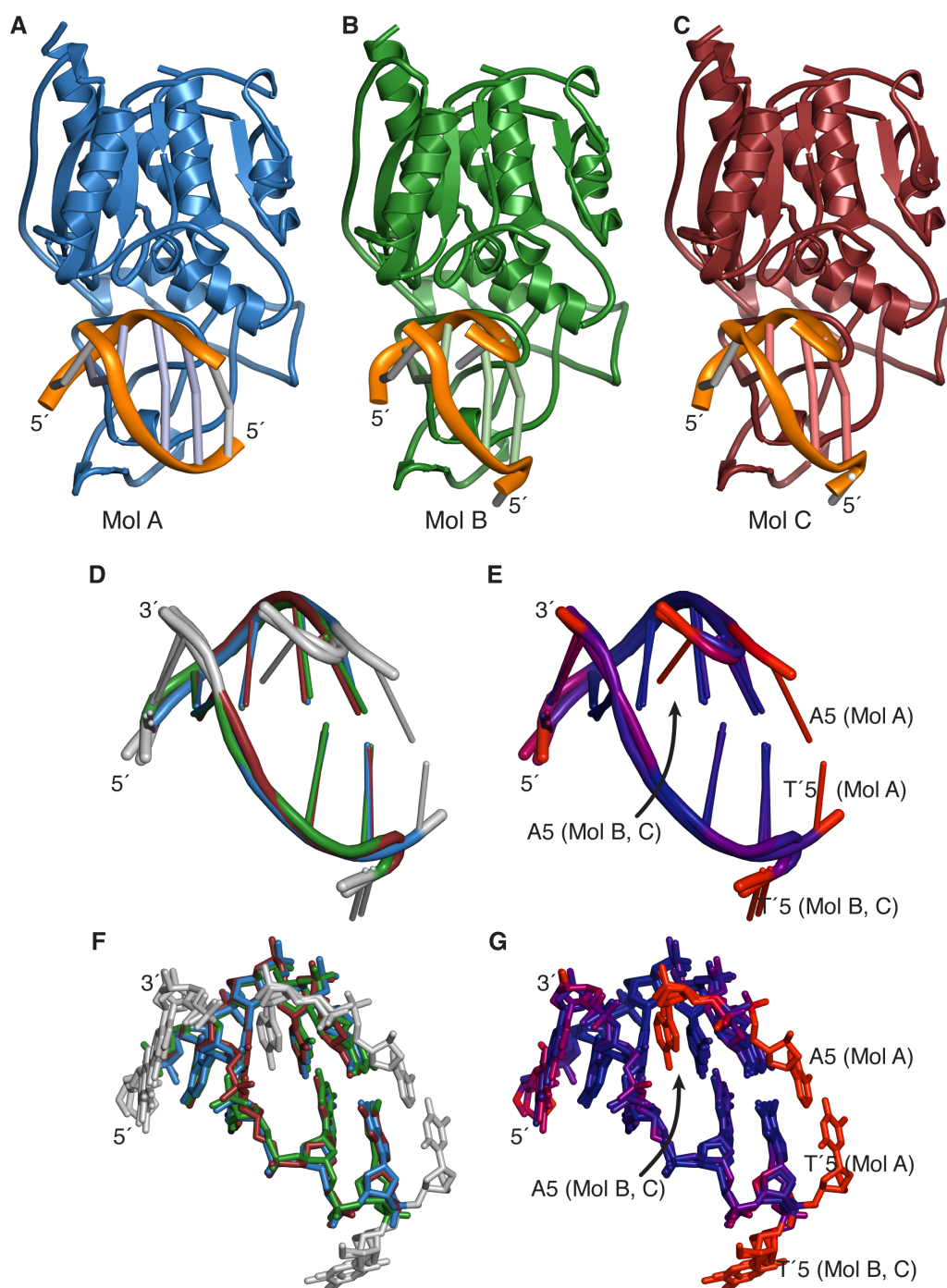
^b $R_{\text{sym}} = R_{\text{merge}} = \frac{\sum_{\text{hkl}} \sum_i |\langle I_{\text{hkl}} \rangle - I_{\text{hkl},i}|}{\sum_{\text{hkl}} \sum_i I_{\text{hkl},i}}$; where $\langle I_{\text{hkl}} \rangle = \frac{\sum_i I_{\text{hkl},i}}{n_{\text{hkl}}}$, $i = 1..n_{\text{hkl}}$.

^c *R*_{free} calculated using randomly selected subset of the data (4.8%).

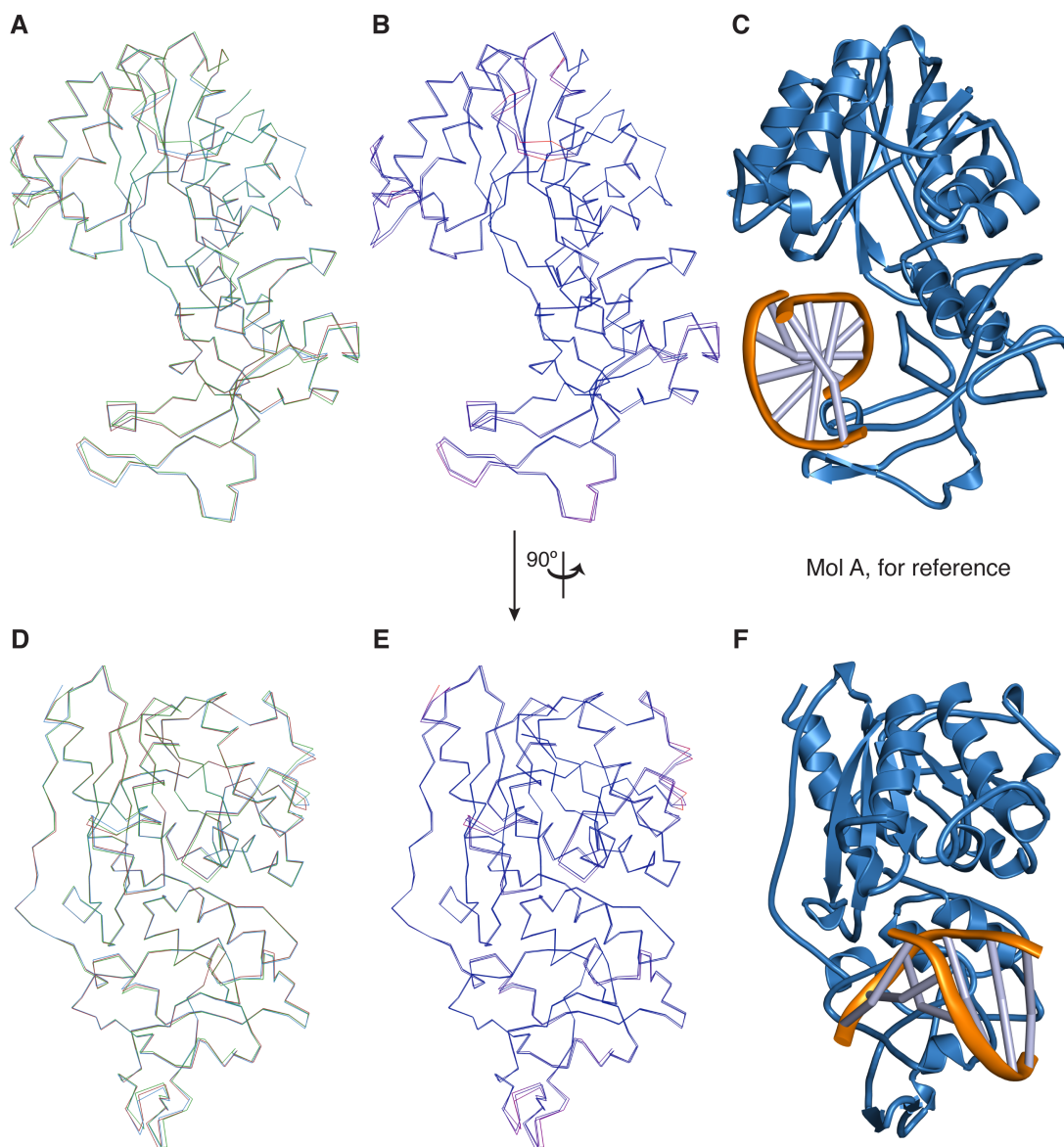
^d Including alternative conformations.



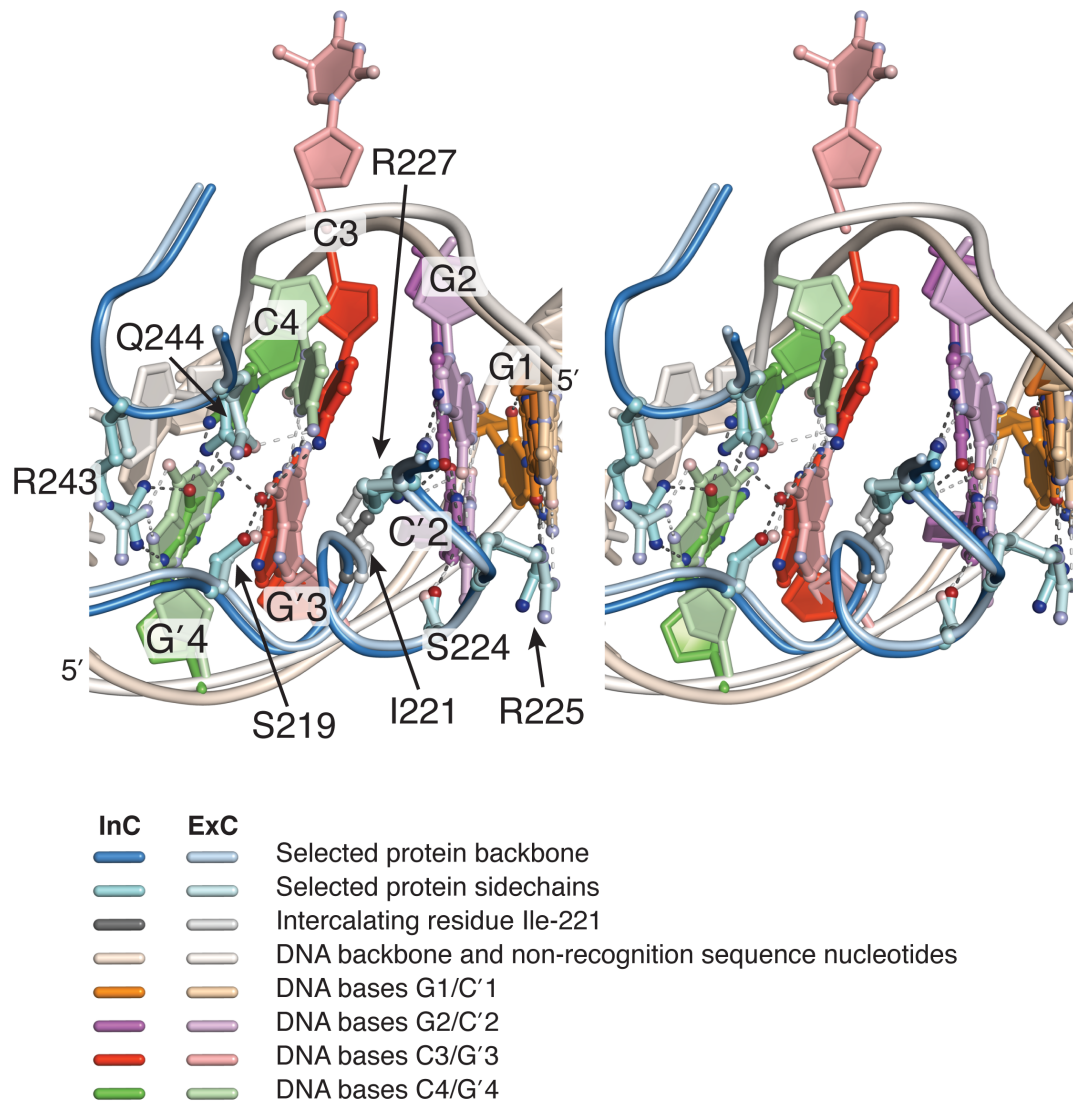
Supplemental Figure S1. InC crystal packing. M.HaeIII in Mol A – blue, in Mol B – green, in Mol C – dark-red, DNA – orange. One asymmetric unit contains Mol A–C and three double stranded DNA oligonucleotides. The global helical axes of pseudo-continuous DNA helices in the crystal are shown in dark-orange.



Supplemental Figure S2. Comparison of the three DNA molecules in the asymmetric unit of InC crystal. *A, B, C*, M.HaeIII in Mol A – blue, in Mol B – green, in Mol C – dark-red. *D, F*, DNA in Mol A – blue, in Mol B – green, in Mol C – dark-red. Cartoon representation in *D*, sticks representation in *F*. *E, G*, Same representation of DNA in Mol A–C as in *D, F*; corresponding atoms in Mol A–C are colored according to their deviation from a common centre of mass. Only the sugar-phosphate backbone of the 3'-terminal adenine nucleotide (A5) of the target DNA strand (top strand on this figure) in Mol C was included in the model; the corresponding adenine nucleobase was not modeled due to its poor electron density.



Supplemental Figure S3. Comparison of the three M.HaeIII molecules in the asymmetric unit of InC crystal. *A, D*, M.HaeIII in Mol A – blue, in Mol B – green, in Mol C – dark-red. *B, E*, Same orientation as in *A* and *D*; corresponding atoms in Mol A–C are colored according to their deviation from a common centre of mass. *C, F*, M.HaeIII-DNA complex of Mol A is shown for reference. The molecules are represented as C_{α} ribbons in *A, B, D, E* and as cartoon in *C, F*. M.HaeIII in Mol B and Mol C were superimposed with M.HaeIII in Mol A using all-atom least-squares fit function of Coot.



Supplemental Figure S4. Overlay of the M.HaeIII DNA sequence specific contacts in InC and ExC. Stereoview of the overlay of the protein-DNA interface in InC (saturated colors) and ExC (light colors), focusing on sequence-specific contacts between the major groove surface of the recognition site in DNA and amino acid side-chains of M.HaeIII. Coding-coding, representation, and orientation as in Fig. 3A, C.