SUPPLEMENTARY INFORMATION

Solution Structure of the 128 kDa Enzyme I Dimer from *Escherichia coli* and its 146 kDa Complex With HPr Using Residual Dipolar Couplings and Small and Wide Angle X-Ray Scattering

Charles D. Schwieters,[2] Jeong-Yong Suh,[1] Alexander Grishaev,[1] Rodolfo Ghirlando,[3] Yuki Takayama[1] and G. Marius Clore[1,*]

[1]Laboratory of Chemical Physics, Building 5, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520, U.S.A.

[2]Division of Computational Biosciences, Center for Information Technology, National Institutes of Health, Bethesda, MD 20892-5624

[3]Laboratory of Molecular Biology, Building 5, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of health, Bethesda, MD 20892-0530, U.S.A.

1. Sedimentation velocity

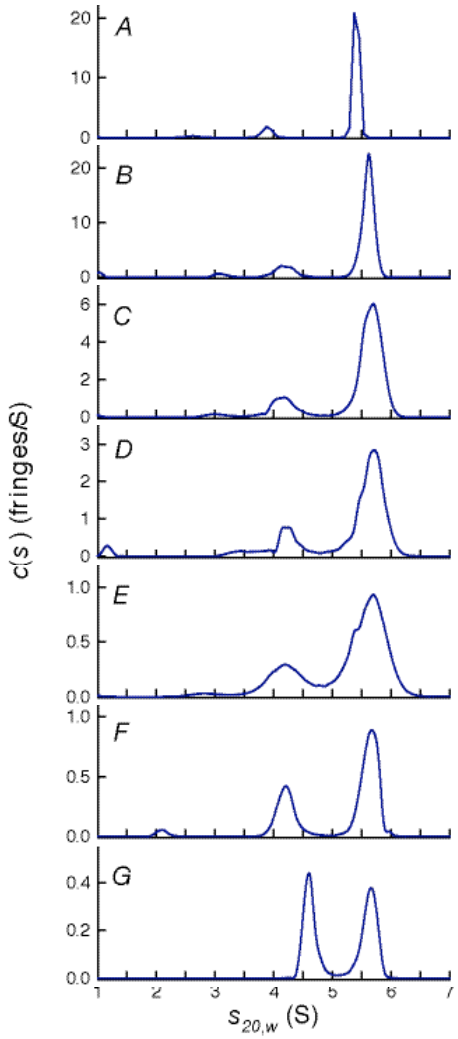

Figure S1. The self-association properties of EI. The *c*(*s*) distributions obtained for free EI at various loading concentrations based on sedimentation velocity interference data collected at 40 krpm and 20.0 ºC. Data were collected at loading concentrations of (A) 67, (B) 28, (C) 15, (D) 8.4, (E) 4.0, (F) 2.2 and (G) 1.0 µM. The buffer comprised 20 mM Tris pH 7.4, 100 mM NaCl, 10 mM dithiothreitol, 4 mM $MgCl_2$, and 1 mM EDTA. Based on these data, the sedimentation coefficients, obtained by extrapolation to zero concentration, are 5.68±0.02 for the EI dimer and 4.25 ±0.04 S for the EI monomer. The weighted average sedimentation coefficients combining contributions from the EI monomer and dimer were also used to determine the equilibrium constant. Using data collected at 1 to 15 µM EI, a K_{diss} of 0.6 µM was obtained, in good agreement with the value of 0.8±0.2 µM obtained from the signal intensities calculated by integration of the *c*(*s*) monomer and dimer peaks.

2. Structure determination protocols

Throughout the calculations, the two EIN domains (residues 1-254) are treated as rigid bodies, the two symmetry related EIC domains (residues 262-573) are held fixed in space, and the linker region (residues 255-261) is given torsional or Cartesian degrees of freedom. The calculational strategy makes use of a three step procedure comprising an initial conjoined rigid body/Cartesian simulated annealing step, an intermediate docking phase, and a final conjoined rigid body/torsion angle/Cartesian simulated annealing refinement step. The simulated annealing protocol employed for the first and third steps is identical with the exception that in the third step, side chains located at the EIN/EIC interface are given torsional degrees of freedom. The purpose of the intermediate docking phase is to ensure extensive sampling of conformational space. Details are provided below.

Conjoined rigid body/torsion angle/Cartesian simulated annealing. The protocol starts with a long molecular dynamics run of 800 ps or 8000 variable time steps (whichever comes first) at 3000 K employing RDC, SAXS/WAXS, bond, angle, improper, van der Waals (vdw) repulsion (E_{repel}), ϕ/ψ dihedral angle (within the EIN/EIC linker), multidimensional torsion angle database (E_{DB}), distance symmetry and non-crystallographic symmetry (NCS) energy terms. (The square-well backbone ϕ/ψ torsion angle restraints for residues 255-261 in the EIN/EIC linker serve to exclude disfavored as well as positive ϕ regions of the Ramachandran map). During the high-temperature phase, the non-bonded potential is configured such that only Cα atoms interact with a radius scale factor of 1.2 and force constant of 0.1 kcal.mol^{-1}.Å^{-4}. This is followed by simulated annealing during which the temperature is decreased to 25 K in 25 K increments. At each temperature molecular dynamics is performed for 0.2 ps or 100 variable time steps (whichever comes first) and the SAXS/WAXS solvent fit parameters and globic correction (Eq. 5, main text) are recomputed. During simulated annealing, the following parameters are geometrically ramped: van der Waals radius scale factor, 0.9→0.8; E_{repel} force constant, 0.004→4 kcal.mol^{-1}.Å^{-4}; E_{DB} force constant, 0.002→1; bond angle force constant: 200→500

kcal.mol^{-1}.rad^{-2}; improper angle force constant, 50→500 kcal.mol^{-1}.rad^{-2}; and torsion angle restraint force constant, 0.1→200 kcal.mol^{-1}.rad^{-2}. The following force constants are kept fixed throughout: SAXS/WAXS force constant, 400 kcal.mol^{-1}; RDC force constant, 1 kcal.mol^{-1}.Hz^{-2}; bond force constant: 1000 kcal.mol^{-1}.Å^{-2}; distance symmetry force constant, 100 kcal.mol^{-1}.Å^{-2}; and NCS force constant, 10 kcal.mol^{-1}.Å^{-2}. Powell gradient minimization is performed to obtain the final structures.

The docking phase. The following docking protocol was repeated 120 times and the resulting 120 structures were used as starting coordinates for the final simulated annealing calculations.

1. The initial values of the magnitude of the principal component of the alignment tensor, D_a^{NH}, and the rhombicity, η, are determined by fitting the observed RDCs to the EIN domain of subunit A alone.
2. Powell gradient minimization is used in the absence of the expensive E_{repel} term to produce a structure which is then annealed in the presence of the E_{repel} term in the following procedure:
    (a) The positions and orientations of the two EIN subunits are separately randomized, breaking the C'-N bond between residues 261 and 262 that connects each EIC domain to the EIN/EIC linker region. In addition, backbone ϕ/ψ torsion angles in the linker region (residues 255-261) are randomized.
    (b) The EIN domains are allowed to freely rotate and translate as rigid bodies, the full RDC alignment tensor is allowed to vary, and the linker region is given full torsion angle degrees of freedom. First two minimization runs of 1000 steps each are carried out with optimization of bond, distance symmetry and NCS terms, followed by two further minimization runs of 1000 steps each

with the RDC term added to the target function.

(c) The RDC R-factor (R_{inf} defined as $\{<(D_{obs}-D_{calc})^2>/(2<D_{obs}^2>)\}^{1/2}$ where D_{obs} and D_{calc} are the observed and calculated RDCs, respectively) for the resulting structure is computed and if this value is greater than a threshold set to 18.22%, slightly above the value obtained using SVD against a single EIN domain (18.02% for EI free), the structure is discarded and the procedure repeated, starting at step 2a.

(d) The orientations of the EIN domains and the RDC alignment tensor are then fixed, such that the RDC fit remains unchanged for the remainder of the calculation.

(e) With their orientations fixed, the EIN domains are allowed to translate, and the two linker regions (one for each subunit) are given torsion angle degrees of freedom to optimize the SAXS/WAXS data by 1000 steps gradient energy minimization optimizing bond, distance symmetry, NCS, RDC and R_{gyr} terms, followed by another 1000 steps of minimization optimizing bond, distance symmetry, NCS, RDC, R_{gyr} and explicit SAXS/WAXS terms.

(f) With the EIC and EIN domains fixed, the geometry of the linker region is optimized, allowing full Cartesian degrees of freedom for the linker regions. 4 minimization runs of 1000 steps each are carried out with optimization of bond and angle terms, followed by a further 4 cycles of minimization (1000 steps each) with optimization of bond, angle and improper terms.

(g) Structures with more bond violations (>0.05 Å) than the initial hybrid structure are discarded and the procedure is repeated starting at step 2a.

(h) A final gradient minimization, including bond, angle, improper, distance symmetry, NCS, RDC and SAXS/WAXS terms is performed allowing EIN translations and linker region torsion degrees of freedom.

(i) The procedure (steps 2a-h) is repeated four more times and the lowest energy

structure (considering the covalent geometry, NCS, distance symmetry, RDC and SAXS/WAXS terms) is selected for simulated annealing with the E_{repel} term included.

3. Simulated annealing with bath temperatures ramped from 3000→25 K in increments of 25 K is carried out allowing translations of the EIN domains and linker torsion degrees of freedom using uniform atomic masses of 100 amu. The treatment of the energy terms and the simulated annealing parameters are identical to those for the conjoined rigid body/torsion angle/Cartesian simulated annealing protocol, with the exception of the inclusion of one additional energy term for R_{gyr} with a constant force constant of 100 kcal.mol^{-1}.Å^{-2}.