SUPPLEMENTARY INFORMATION

Hybrid thermophilic/mesophilic enzymes reveal a role for conformational disorder in regulation of bacterial Enzyme I

Rochelle R. Dotas,[1] Trang T. Nguyen,[1] Charles E. Stewart Jr.,[2] Rodolfo Ghirlando,[3] Davit A. Potoyan*,[1,4] and Vincenzo Venditti*,[1,4]

[1] *Department of Chemistry, Iowa State University, Ames, Iowa 50011, USA.*

[2] *Macromolecular X-ray Crystallography Facility, Office of Biotechnology, Iowa State University, Ames, IA 50011, USA*

[3] *Laboratory of Molecular Biology, NIDDK, National Institutes of Health, Bethesda, Maryland 20892.*

[4] *Roy J. Carver Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, Iowa 50011, USA.*

* Address correspondence to:

Vincenzo Venditti, Department of Chemistry, Iowa State University, Hach Hall, 2438 Pammel Drive, Ames, IA 50011, USA. email: venditti@iastate.edu; Tel. 515-294-1044; Fax: 515-294-7550; ORCID 0000-0001-8734-0400.

Davit A. Potoyan, Department of Chemistry, Iowa State University, Gilman Hall, 2415 Osborn Drive, Ames, IA 50011, USA. email: potoyan@iastate.edu; Tel. 515-294-9971; Fax: 515-294-7550; ORCID 0000-0002-5860-1699.

X: active site X: conserved residues —EIN —EIC - - - active site loops *: mutations

```
eEI    1    MISGILASPGIAFGKALLLKEDEIVIDRKKISADQVDQEVERFLSGRAKASAQLETIKTK   60
tEI    1    MLKGVAASPGIAIGKAFLYTKEKVTINVEKIEESKVEEEIAKFRKALEVTQEEIEKIKEK   60

eEI    61   AGETFGEEKEAIFEGHIMLLEDEELEQEIIALIKDKHMTADAAAHEVIEGQASALEELDD   120
tEI    61   ALKEFGKEKAEIFEAHLMLASDPELIEGVENMIKTELVTADNAVNKVIEQNASVMESLND   120

eEI    121  EYLKERAADVRDIGKRLLRNILGLKIIDLSAIQDEVILVAADLTPSETAQLNLKKVLGFI   180
tEI    121  EYLKERAVDLRDVGNRIIENLLGVKSVNLSDLEEEVVVIARDLTPSDTATMKKEMVLGFA   180

eEI    181  TDAGGRTSHTSIMARSLELPAIVGTGSVTSQVKNDDYLILDAVNNQVYVNPTNEVIDKMR   240
tEI    181  TDVGGRTSHTAIMARSLEIPAVVGLGNVTSQVKAGDLVIVDGLEGIVIVNPDEKTVEDYK   240

eEI    241  AVQEQVASEKAELAKLKDLPAITLDGHQVEVCANIGTVRDVEGAERNGAEGVGLYRTEFL   300
tEI    241  SKKESYEKKVEGLKQLKDLPAETPDGKKVMLAANIGTPKDVASALANGAEGVGLFRTEFL   300
                                                                  β2α2 loop
eEI    301  FMDRDALPTEEEQFAAYKAVAEACGSQAVIVRTMDIGGDKELPYMNFPKEENPFLGWRAI   360
tEI    301  YMDRNSLPSEEEQFEAYKEVVEKMGGRPVTIRTLDIGGDKELPYLDMPKEMNPFLGYRAI   360
            (res. 296-309)                  β3α3 loop (res. 332-360)
eEI    361  RIAMDRKEILRDQLRAILRASAFGKLRIMFPMIISVEEVRALRKEIEIYKQELRDEGKAF   420
tEI    361  RLCLDRPDIFKTQLRAILRASAYGNVQIMYPMISSVEEVRKANSILEEVKAELDREGVKY   420

eEI    421  DESIEIGVMVETPAAATIARHLAKEVDFFSIGTNDLTQYTLAVDRGNDMISHLYQPMSPS   480
tEI    421  DKEIKVGIMVEIPSAAVTADILAKEVDFFSIGTNDLTQYTLAVDRMNEHVKEYYQPFHPA   480
                                            β6α6 loop (res. 454-477)
eEI    481  VLNLIKQVIDASHAEGKWTGMCGELAGDERATLLLLGMGLDEFSMSAISIPRIKKIIRNT   540
tEI    481  ILRLVKMVIDAAHKEGKFAAMCGEMAGDPLAAVILLGLGLDEFSMSATSIPEIKNIIRNV   540

eEI    541  NFEDAKVLAEQALAQPTTDELMTLVNKFIEEKTIC   575
tEI    541  EYEKAKEIAEKALNMSEAREIEKMMKDVI--KDIG   573
```

Figure S1. Sequence alignment of *e*EI and *t*EI. The amino acid sequences of *e*EI and *t*EI were aligned in BLAST. Active site residues are colored red. Conserved residues are colored green. The EIN and EIC domains are indicated with a blue and red line underneath the amino acid sequence, respectively. The location of the active site loops is shown with dashed black lines underneath the amino acid sequence. Positions of the 21 single-point mutations performed to hybridize the EIC domain are indicated by asterisks over the amino acid sequence. Note that residues 278 and 279 do not belong to the active site loops of EIC but were mutated in the hybrids because they face the β2α2 loop in the three-dimensional structure of the enzyme.

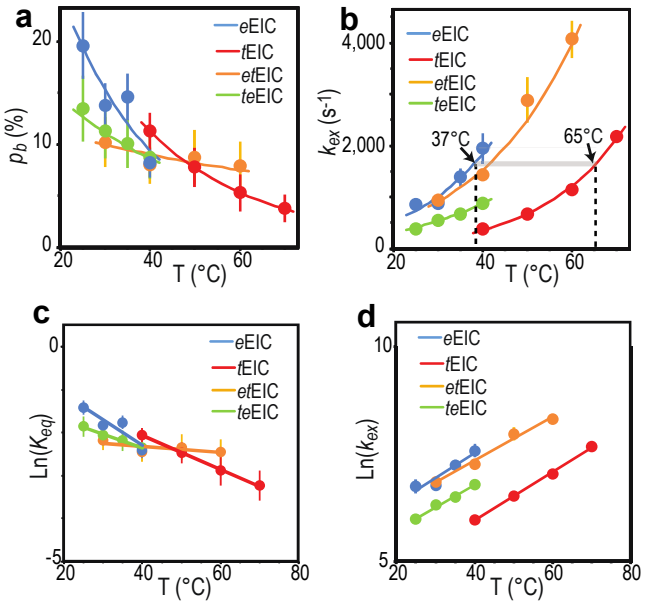

Figure S2. Fitting of RD data without enforcing Eyring and van't Hoff behavior. RD curves were fitted without enforcing the temperature dependence of k_{ex} and p_b to respect the Eyring and van't Hoff equations, respectively. Fitted p_b (**a**) and k_{ex} (**b**) for the expanded-to-compact equilibrium in apo *e*EIC (blue), *t*EIC (red), *et*EIC (orange), and *te*EIC (green) are plotted versus the experimental temperature. Corresponding values of ln(K_{eq}) and ln (k_{ex}) are shown in (**c**) and (**d**), respectively. Experimental data are represented by filled-in circles. Modeling of the experimental data using the van't Hoff (for p_b) and Eyring (for k_{ex}) equations are shown as solid lines. Vertical dashed lines are at the optimal PTS temperatures for *e*EIC (37 °C) and *t*EIC (65 °C). Given the good agreement between experimental and modelled data, Eyring and van't Hoff behaviors were enforced in the final fits (showed in main text) to reduce the number of variable parameters in the fitting protocol.

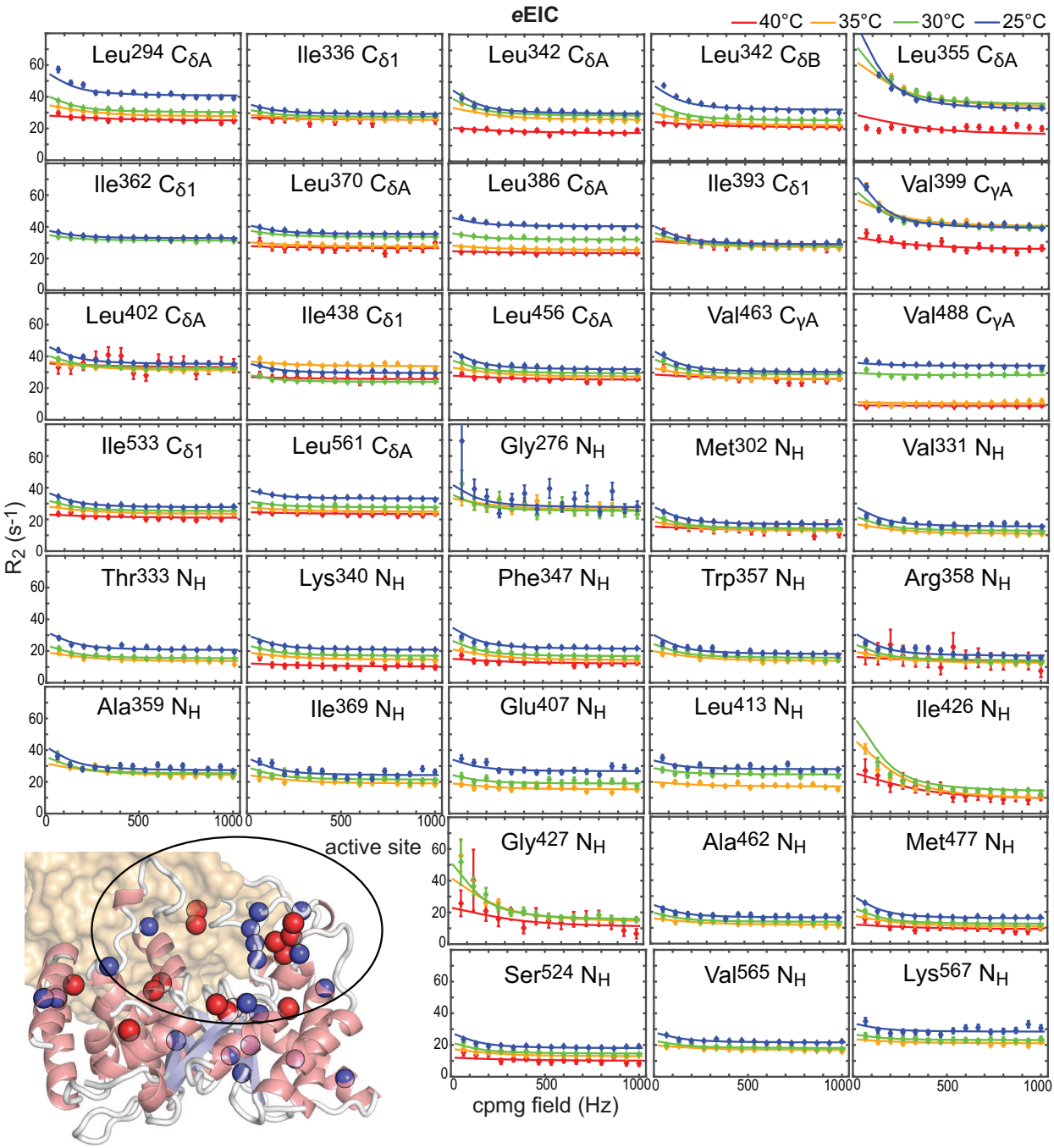

Figure S3. RD data on apo *e*EIC. Global fitting of the 36 ^{15}N and $^{13}C_{methyl}$ relaxation dispersion curves at 25 (blue), 30 (green) 35 (orange), and 40 (red) °C that describe µs-ms dynamics in apo *e*EIC. Experimental data are reported as circles. Results of the global fit are shown as solid lines. Relaxation dispersion curves measured at 800 MHz are shown. The location on the *e*EIC structure of the NMR peaks used in the fitting procedures (i.e. all peaks with $R_{ex} > 5$ s^{-1} at 25 °C) is shown

in the bottom-left corner. Amides and methyl groups are shown on one subunit as blue and red spheres, respectively. α-helices and β-strands are colored salmon and light blue, respectively. The location of the active site is indicated with a black circle. The second subunit is shown as transparent surface.

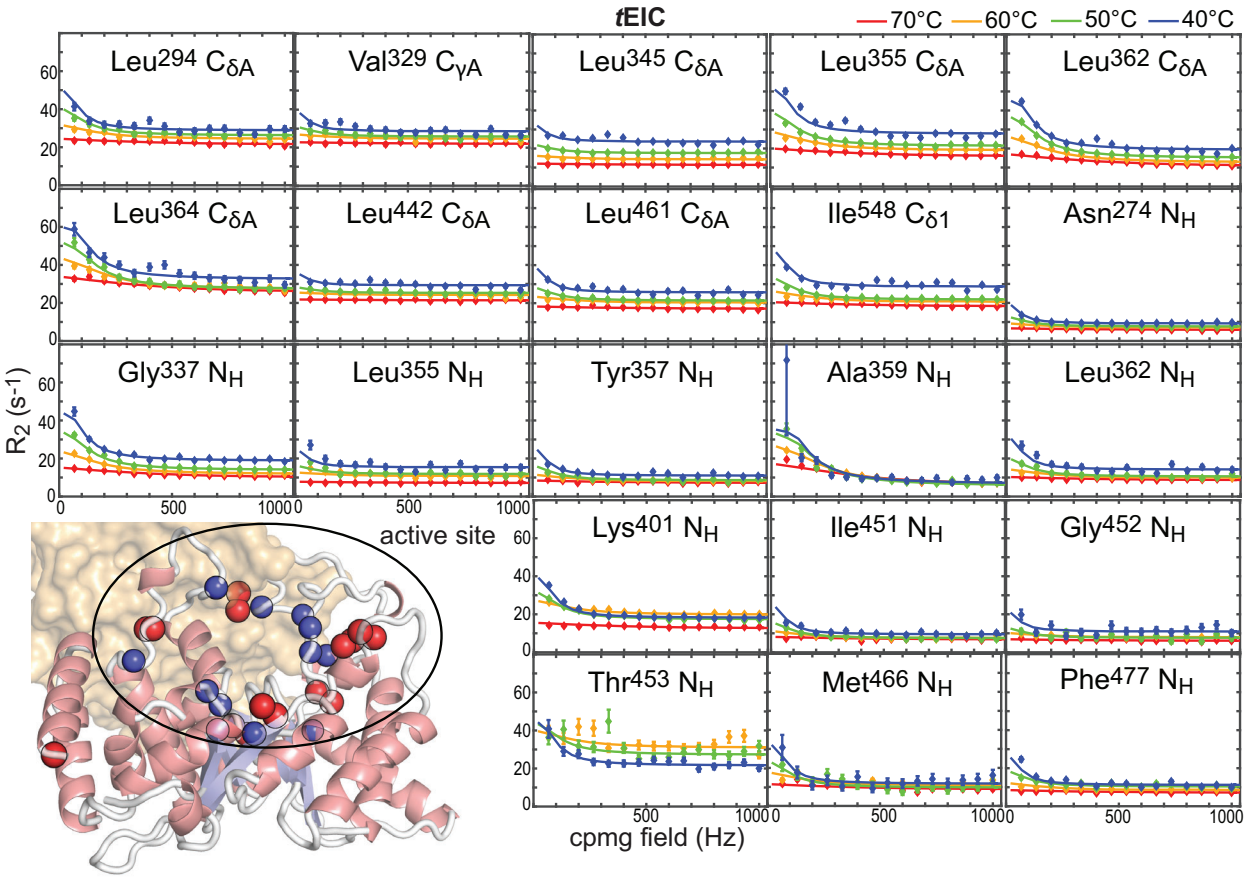

Figure S4. RD data on apo *t*EIC. Global fitting of the 21 ^{15}N and $^{13}C_{methyl}$ relaxation dispersion curves at 40 (blue), 50 (green) 60 (orange), and 70 (red) °C that describe µs-ms dynamics in apo *t*EIC. Experimental data are reported as circles. Results of the global fit are shown as solid lines. Relaxation dispersion curves measured at 800 MHz are shown. The location on the *t*EIC structure of the NMR peaks used in the fitting procedures (i.e. all peaks with $R_{ex} > 5$ s^{-1} at 40 °C) is shown in the bottom-left corner. Amides and methyl groups are shown on one subunit as blue and red spheres, respectively. α-helices and β-strands are colored salmon and light blue, respectively. The location of the active site is indicated with a black circle. The second subunit is shown as transparent surface.

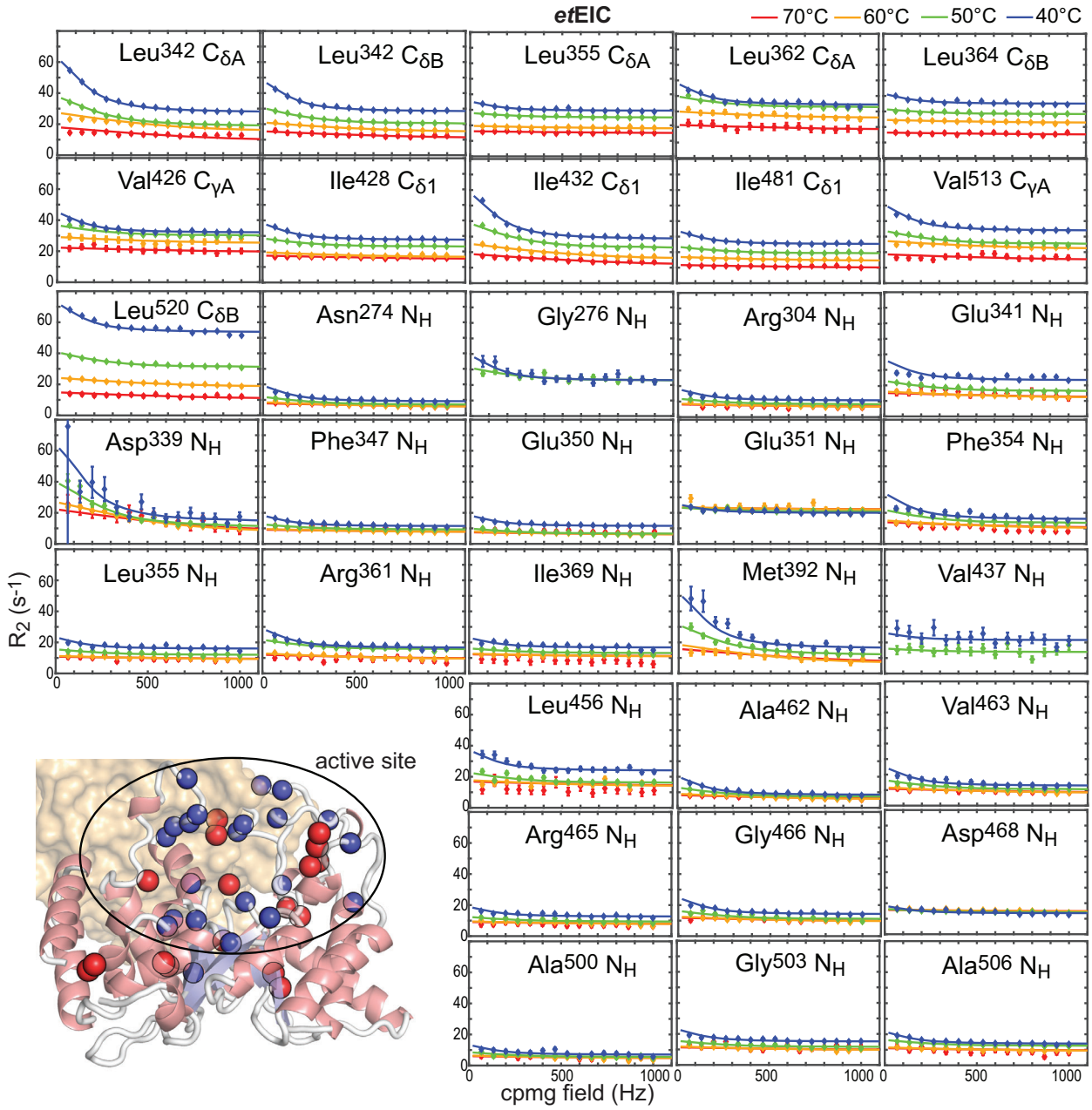

Figure S5. RD data on apo *et*EIC. Global fitting of the 34 ^{15}N and $^{13}C_{methyl}$ relaxation dispersion curves at 40 (blue), 50 (green) 60 (orange), and 70 (red) °C that describe µs-ms dynamics in apo *et*EIC. Experimental data are reported as circles. Results of the global fit are shown as solid lines. Relaxation dispersion curves measured at 800 MHz are shown. The location on the *et*EIC structure of the NMR peaks used in the fitting procedures (i.e. all peaks with $R_{ex} > 5$ s^{-1} at 40 °C) is shown in the bottom-left corner. Amides and methyl groups are shown on one subunit as blue and red spheres, respectively. α-helices and β-strands are colored salmon and light blue, respectively. The

location of the active site is indicated with a black circle. The second subunit is shown as transparent surface.

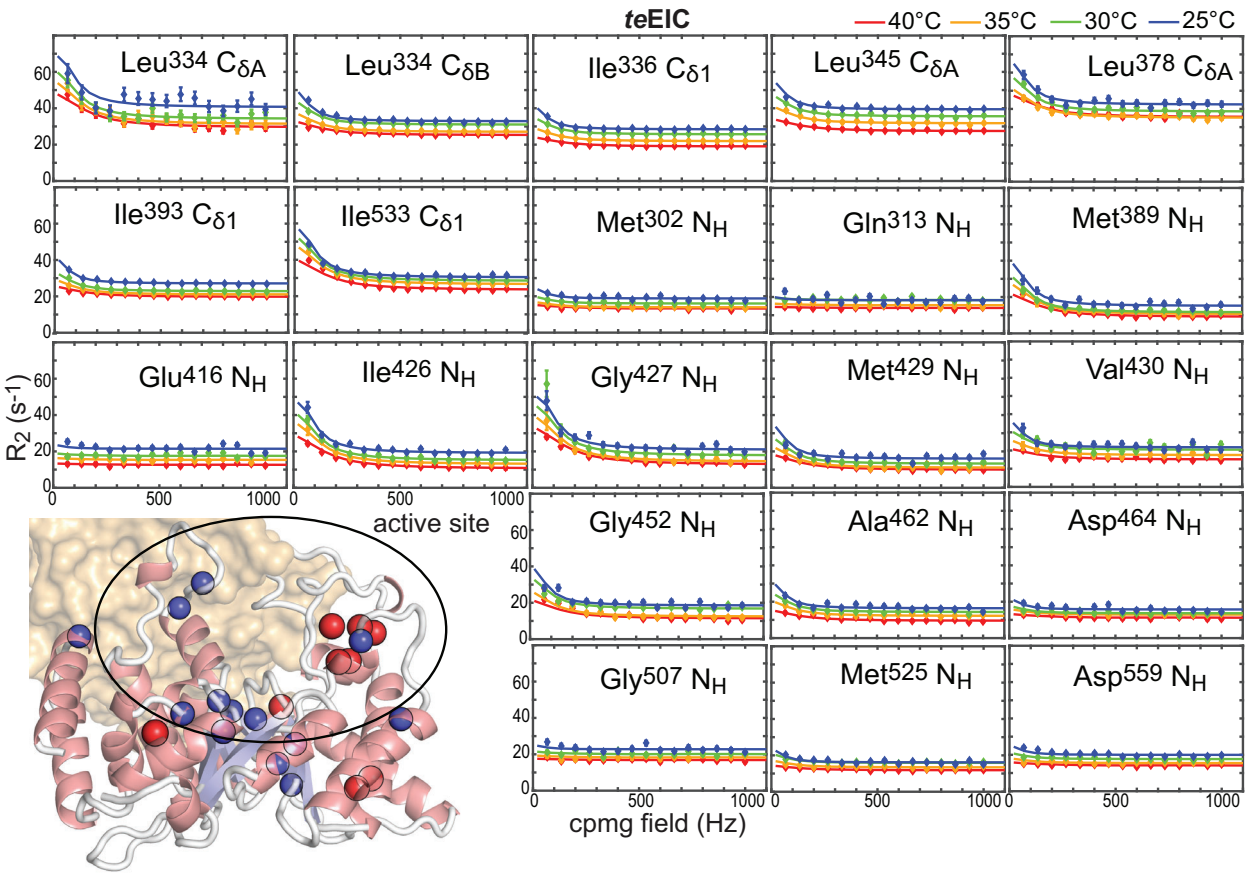

Figure S6. RD data on apo *te*EIC. Global fitting of the 21 ^{15}N and $^{13}C_{methyl}$ relaxation dispersion curves at 25 (blue), 30 (green) 35 (orange), and 40 (red) °C that describe µs-ms dynamics in apo *te*EIC. Experimental data are reported as circles. Results of the global fit are shown as solid lines. Relaxation dispersion curves measured at 800 MHz are shown. The location on the *te*EIC structure of the NMR peaks used in the fitting procedures (i.e. all peaks with $R_{ex} > 5$ s^{-1} at 25 °C) is shown in the bottom-left corner. Amides and methyl groups are shown on one subunit as blue and red spheres, respectively. α-helices and β-strands are colored salmon and light blue, respectively. The location of the active site is indicated with a black circle. The second subunit is shown as transparent surface.

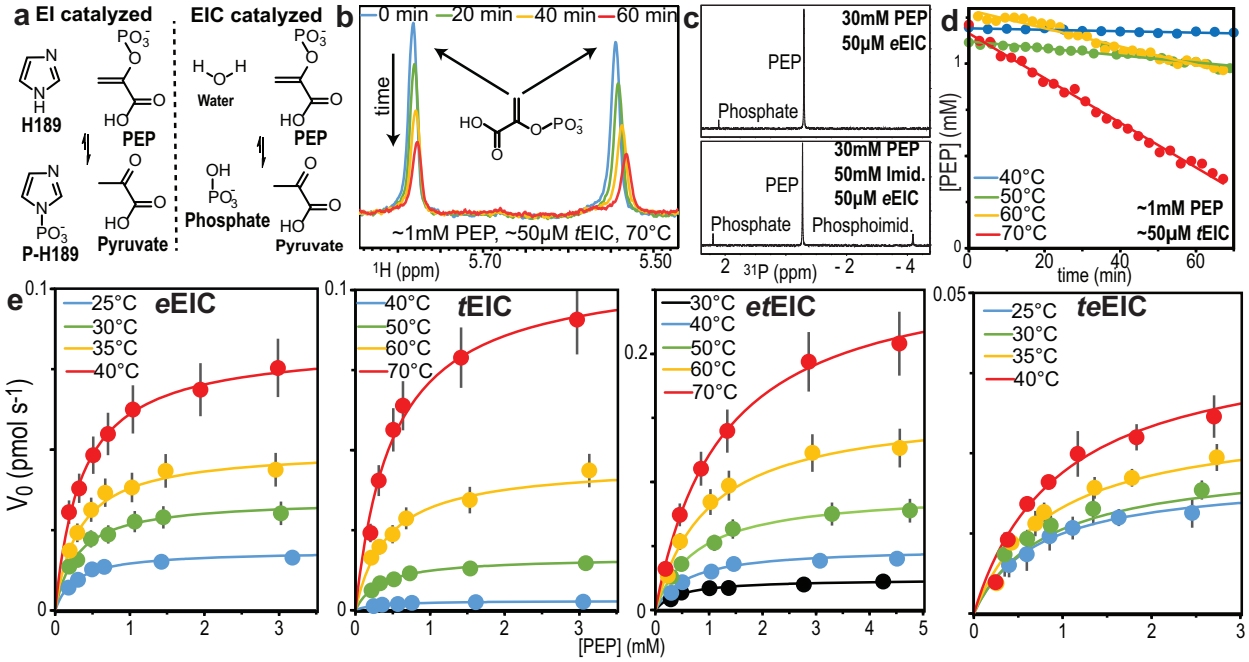

Figure S7. Enzymatic assay for EIC. (**a**) Enzymatic reactions catalyzed by full-length EI (left) and isolated EIC (right). (**b**) Region of the ^{1}H NMR spectrum of PEP displaying the signals of the two alkene protons. Spectra were measured at 70 °C in the presence of 1 mM PEP and 50 μM *t*EIC. Blue, green, orange, and red spectra were measured after 0, 20, 40, and 60 minutes of incubation, respectively. (**c**) ^{31}P NMR spectrum measured at 40 °C on samples containing 30 mM PEP and 50 μM *e*EIC (top) or 30 mM PEP, 50 mM imidazole, and 50 μM *e*EIC (bottom). Samples were incubated at 40 °C for ~30 minutes before acquisition of the NMR data. (**d**) PEP concentration versus time measured on samples containing ~1 mM PEP and 50 μM *t*EIC at 40 (blue), 50 (green), 60 (orange), and 70 (red) °C. Experimental data are shown as filled-in circles. Linear regressions of the data are shown as solid lines. (**e**) Michaelis-Menten kinetics for *e*EIC, *t*EIC, *et*EIC, and *te*EIC measured at different temperatures. Experiential data are shown as filled-in circles. Fitted curves are shown as solid lines.

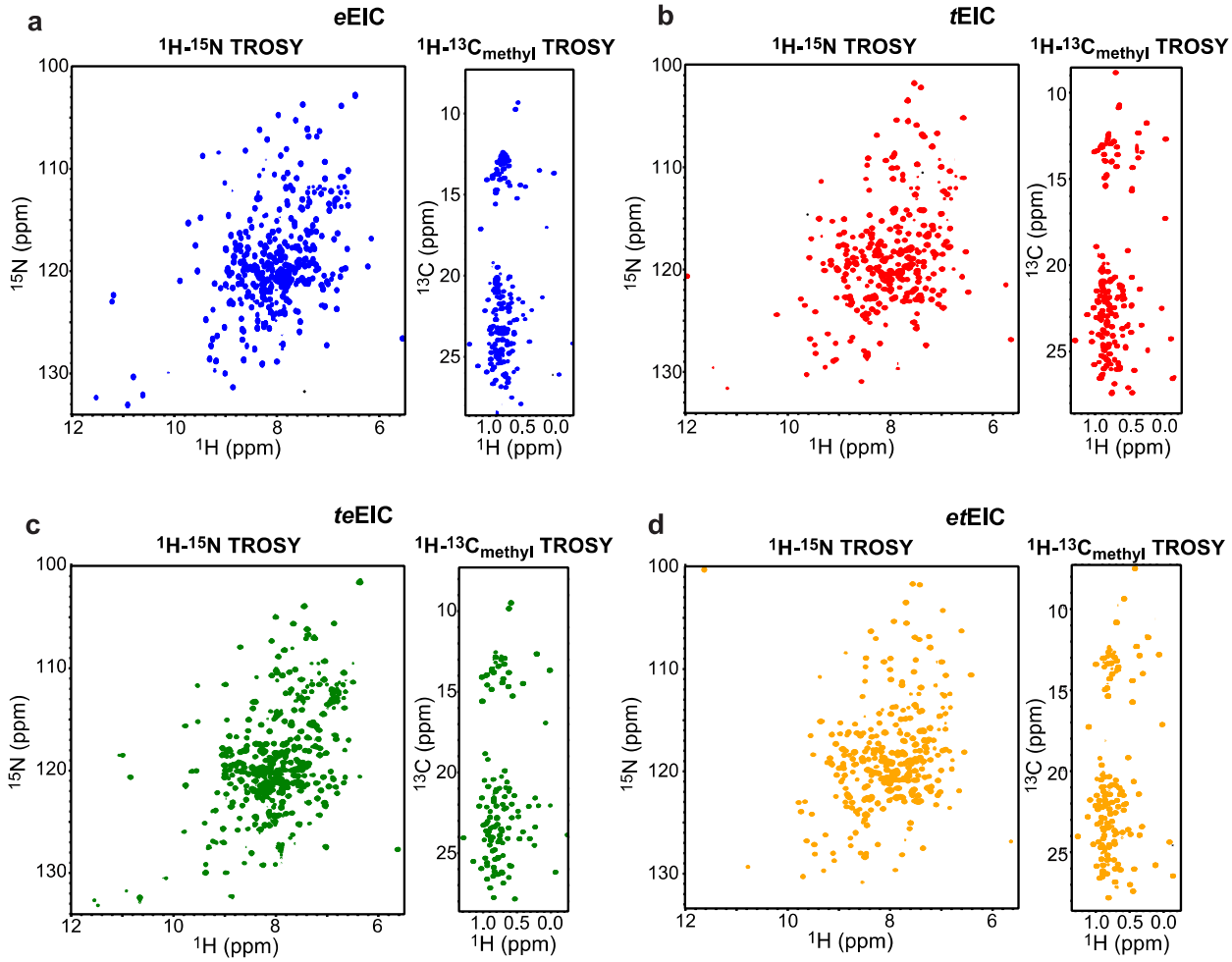

Figure S8. 800 MHz NMR spectra of EIC. ^{1}H-^{15}N TROSY (left) and ^{1}H-^{13}C$_{methyl}$ TROSY (right) spectra of (**a**) *e*EIC, (**b**) *t*EIC, (**c**) *te*EIC, and (**d**) *et*EIC.

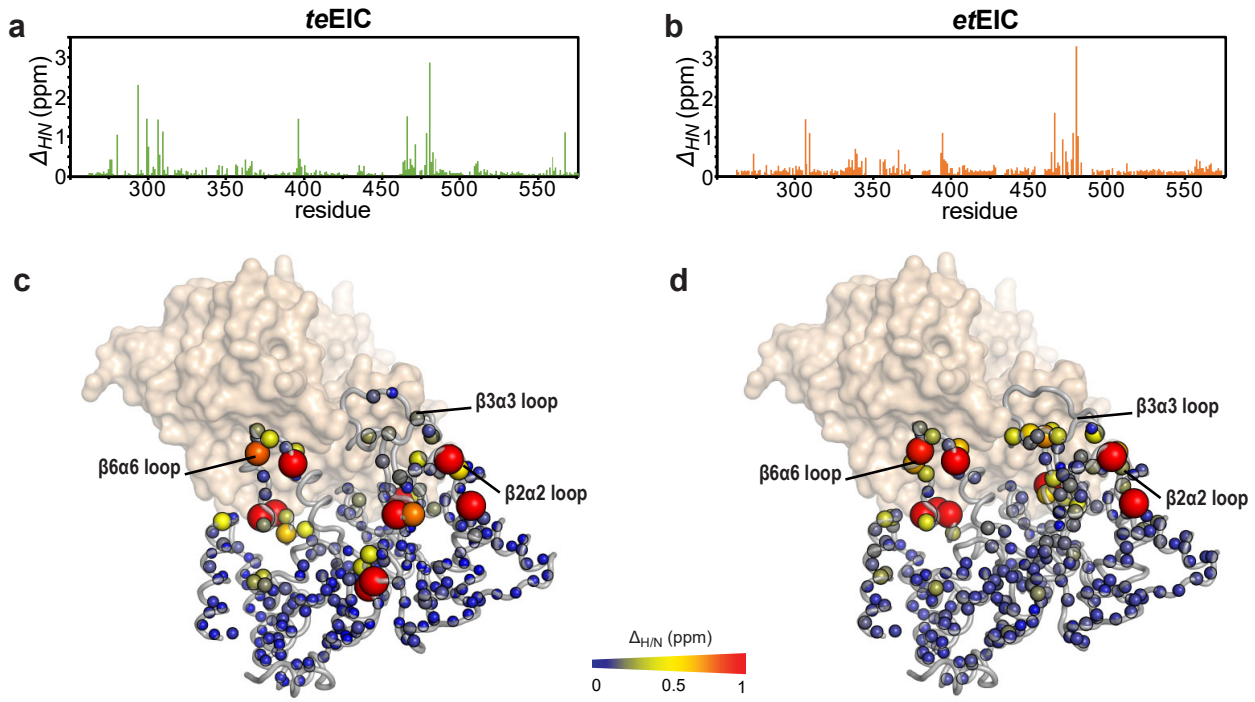

Figure S9. Effect of the 21-point mutations on the spectra of *e*EIC and *t*EIC. Weighted combined CSP ($\Delta_{H/N}$) induced by the 21 mutations of the ^{1}H-^{15}N TROSY spectra of (**a**) *e*EIC (i.e. the comparison between *e*EIC and *te*EIC spectra) and (**b**) *t*EIC (i.e. the comparison between *t*EIC and *et*EIC spectra). $\Delta_{H/N}$ values shown in (a) and (b) are plotted on the crystal structure of *t*EIC in (**c**) and (**d**), respectively. The relationship between size and color of each sphere and chemical shift perturbation is depicted by the color bar. The effect of the mutations on the NMR spectra of EIC is localized in the active site loops and adjacent regions.

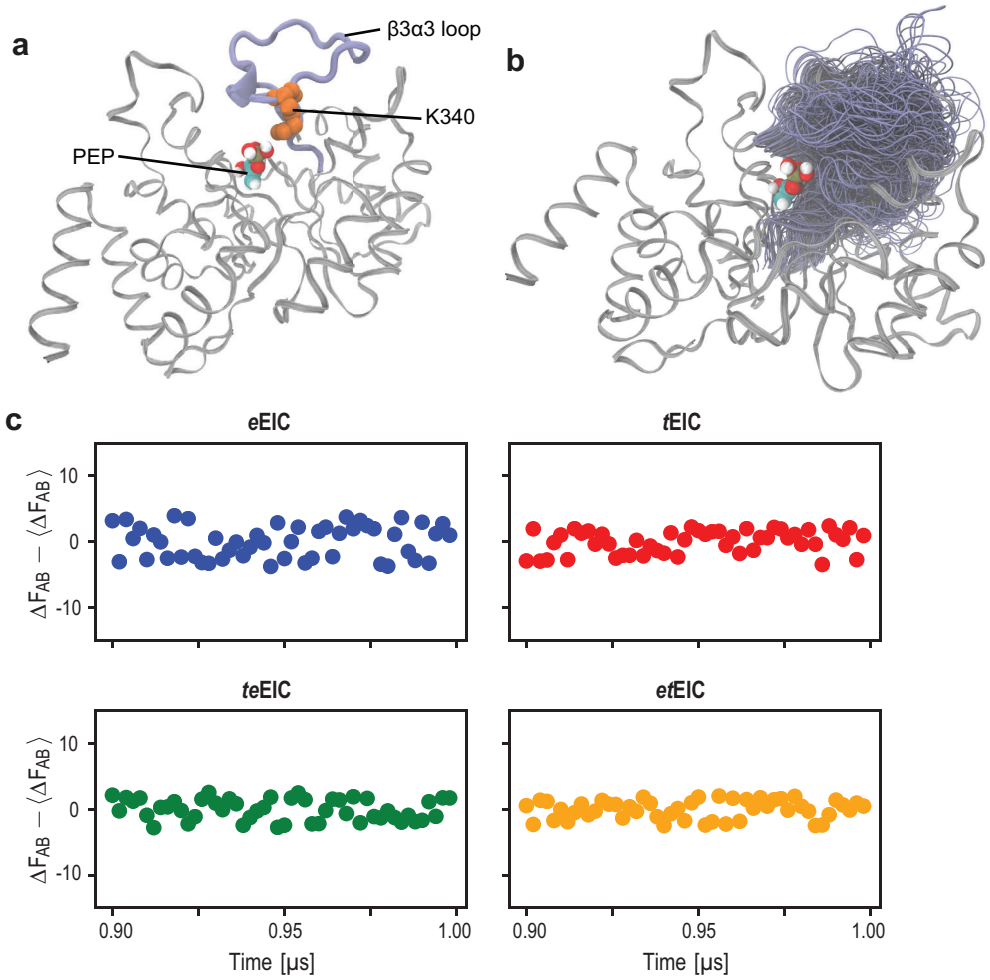

Figure S10. Convergence assessment for WT-MTD simulations. (**a**) Shown is the structure of *e*EIC highlighting the position of PEP and Lys^{340}. (**b**) Shown are the loop conformations sampled by *e*EIC during the 1 μs WT-MTD simulation. Structures were extracted at 10-ns intervals. (**c**) Shown are the de-meaned free energy differences between inactive (A) and active (B) conformations of *e*EIC (blue), *t*EIC (red), *te*EIC (green), and *et*EIC (orange) evaluated over the last 100 ns of the simulated trajectory. The stability of the mean over the last 100 ns simulation was used as a measure for assessing convergence.

Supplementary Tables

Table S1. Kinetics and thermodynamics of the expanded-to-compact equilibrium.

	k_{ab} / k_{ba} (s^{-1}) [a]							$\Delta^{\ddagger}H_{ab}$ / $\Delta^{\ddagger}H_{ba}$ [b]	$\Delta^{\ddagger}S_{ab}$ / $\Delta^{\ddagger}S_{ba}$ [b]	p_b (%) [c]							ΔH [d]	ΔS [d]
T (°C)	25	30	35	40	50	60	70	kJ mol^{-1}	J K^{-1} mol^{-1}	25	30	35	40	50	60	70	kJ mol^{-1}	J K^{-1} mol^{-1}
*e*EIC	68 539	77 740	87 1,002	97±7 1,342	- -	- -	- -	16±2 45±2	-156±6 -43±5	11	9	8	7	-	-	-	-29±2	-114±6
*t*EIC	- -	- -	- -	32 303	40 599	48 1,124	58 2,026	15±2 54±2	-168±4 -26±4	-	-	-	9	6	4	3	-39±2	-142±5
*et*EIC	- -	105 853	- -	111 1,321	115 1,975	119 2,867	- -	1±0.2 31±2	-203±5 -85±5	-	11	-	8	6	4	-	-30±2	-118±7
*te*EIC	39 342	45 459	50 608	57 797	- -	- -	- -	16±3 41±3	-160±10 -58±10	10	9	8	7	-	-	-	-25±2	-102±5
*te*EIC*	37 352	44 475	51 635	60 840	- -	- -	- -	22±3 42±3	-141±10 -54±10	10	9	8	7	-	-	-	-25±2	-102±5

[a] The expanded and compact states are referred to as *a* and *b*, respectively. k_{ab} and k_{ba} are the rate constants for the transition from *a* to *b* and from *b* to *a*, respectively, and are calculated from the values of the optimized parameters k_{ex} (= k_{ab} + k_{ba}) and p_b. For each entry, the upper and lower numbers refer to k_{ab} and k_{ba}, respectively. Errors for the reported *k*'s are < 15% of the reported value (see Figure 2).

[b] Activation enthalpies and entropies for the *a* to *b* and *b* to *a* transitions were calculated by fitting the temperature dependence of k_{ab} and k_{ba} to the Eyring equation, respectively. For each entry, the upper and lower numbers refer to *a* to *b* and *b* to *a* transition, respectively.

[c] Errors for the reported p_b's are < 15% of the reported value (see Figure 2).

[d] Enthalpy and entropy changes associated with the expanded-to-compact equilibrium were calculated by using the van't Hoff equation. The equilibrium constant (K_{eq}) at each temperature was calculated using the formula $K_{eq} = p_b / (1 - p_b)$.

Table S2. Michaelis-Menten parameters for PEP hydrolysis catalyzed by EIC.

	K_M (μM) [a]							k_{cat} x 10^3 (s^{-1}) [b]							$\Delta^{\ddagger}H$ [c]	$\Delta^{\ddagger}S$ [c]
T (°C)	25	30	35	40	50	60	70	25	30	35	40	50	60	70	kJ mol^{-1}	J K^{-1} mol^{-1}
*e*EIC	270	290	320	360	-	-	-	0.7	1.4	2.0	3.3	-	-	-	71	-65
*t*EIC	-	-	-	280	340	430	500	-	-	-	0.1	0.7	1.8	4.3	104	15
*et*EIC	-	440	-	590	770	908	1,240	-	1.0	-	1.9	3.7	6.3	10.8	50	-140
*te*EIC	710	810	860	970	-	-	-	0.4	0.5	0.6	0.9	-	-	-	37	-185

[a] Errors for the reported K_M's are < 30%.

[b] Errors for the reported k_{cat}'s are < 15%. Reported k_{cat} values are multiplied by 10^3 in the table.

[c] $\Delta^{\ddagger}H$ and $\Delta^{\ddagger}S$ values were calculated by fitting the temperature dependence of k_{cat} to the Eyring equation. Errors are < 10% and < 30% for $\Delta^{\ddagger}H$ and $\Delta^{\ddagger}S$, respectively.

Table S3. Xray data collection and refinement statistics.[a]

	***e*EIC**	***et*EIC**	***te*EIC**
PDB code	6VU0	6VBJ	6V9K
Resolution range	37.07-3.5 (3.62 - 3.5)	42.71-2.0 (2.07 - 2.0)	42.82 - 1.9 (1.97 - 1.9)
Space group	P 41 2 2	P 21 21 21	P 1 21 1
a, *b*, *c* (Å)	136.46, 136.46, 183.58	74.43, 85.42, 95.39	57.38, 69.53, 84.66
α, β, γ (°)	90 90 90	90, 90, 90	90, 108.73, 90
Total reflections	175027	481028	348351
Unique reflections	21070	40911	49346
Multiplicity	8.3 (3.2)	11.8 (5.3)	7.1 (6.9)
Completeness (%)	93.23 (85.27)	97.86 (84.78)	98.80 (99.54)
Mean I/sigma(I)	7.09 (1.52)	16.51 (2.06)	15.30 (2.42)
R_{meas}	0.395 (0.840)	0.108 (0.6594)	0.117 (0.8727)
CC1/2	0.994 (0.773)	0.998 (0.799)	0.998 (0.794)
No. of macromolecules in asymmetric unit	2	2	2
R_{work} / R_{free}	0.2308 / 0.2700	0.1910 / 0.2299	0.1604 / 0.2052
No. of protein residues	620	626	620
Root-mean-square deviation Bond lengths (Å)	0.003	0.004	0.005
Root-mean-square deviation Bond angles (°)	0.59	0.94	1.07
Ramachandran favored (%)	93.18	98.07	98.05
Ramachandran allowed (%)	6.66	1.93	1.95
Ramachandran outliers (%)	0.16	0.00	0.00
Rotamer outliers (%)	0.59	0.00	0.77
Clashscore	8.64	5.60	3.45
Average B-factor	108.98	34.68	26.02
macromolecules	108.93	34.51	24.99
ligands	132.09[b]	-	33.52[c]
solvent	-	36.68	34.75

[a] Statistics for the highest-resolution shell are shown in parentheses.

[b] Sulfate ions from crystallization condition.

[c] Mg^{2+} ion from crystallization condition.