IOWA STATE UNIVERSITY
Digital Repository

2020

Characterization of the C-terminal binding domain from bacterial Enzyme I

Rochelle Rea Dotas
Iowa State University

Characterization of the C-terminal binding domain from bacterial Enzyme I

by

Rochelle Rea Dotas

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Chemistry

Program of Study Committee:
Vincenzo Venditti, Major Professor
Amy Andreotti
Arthur Winter
Dipali Sashital
Levi Stanley

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2020

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I would like to thank my committee chair, Vincenzo Venditti, and my committee members, Levi Stanley, Arthur Winter, Amy Andreotti, and Dipali Sashital, for their guidance and support throughout the course of this research. Thank you for believing in me. A special thank you to Amy Andreotti who gave me confidence and encouragement when I needed it the most.

In addition, I would also like to thank my friends, colleagues, the department faculty and staff, for all the opportunities I was given, professional relationships I was able to build, the safe and comfortable learning environment you provided, and for overall making my time as a graduate student at Iowa State University a memorable experience.

To each and every one of my research group members, who became a second family to me, thank you. I will never forget your unwavering support, kindness, encouragement, helpfulness, knowledge, patience, understanding, humor, and positivity you provided daily. I am grateful to have worked with some of the best people.

I look back on my time as a graduate student feeling thankful for the relationships I gained, the huge education opportunity I was blessed with, the confidence and voice I built, the immense personal growth that took place, and most importantly, learning perseverance and endurance through overcoming obstacles, failure, and rejection.

To the women I have been fortunate to meet along the way, you know who you are, I cherish all of those long nights together. Thank you for allowing me to enter your sisterhood and offering me a creative outlet. Thank you for sharing in the up's and down's of life, the laughs, tears, fun times, and hard times. Always remember you are capable of greatness and can do anything you set your mind to.

To my cats, Kit Kat (20+, 10/13/19), Morris (18, 3/13/20), and Max (16, 4/8/20), who I adopted upon moving to Iowa, thank you for making my apartment in Iowa a home and your faithful 5+ years of companionship. You have been there right beside me every step up the way offering me comfort and unconditional love during this season of my life. I will love all three of you forever and remember all the sweet memories together.

To my family, specifically my wonderful parents Kent and Sheri Dotas, thank you for being my biggest encouragers, for the sacrifices you made, guidance you gave me, life lessons you taught me, and the humble work ethic you instilled in me. Thank you for being my rock, keeping me grounded, your consistency, reliability, and most importantly providing me with endless emotional support throughout this journey and my life as a young adult. You celebrated with me in my successes and comforted me during my trials. In many ways I feel privileged yet blessed for all that I have been able to accomplish.

Finally, I would like to thank God for making all things possible and for always having a plan for my life. Thank you for continually reminding me to trust in your plan and process because your timing is perfect. I would not have accomplished what I have without my parents and God by my side.

ABSTRACT

Modulation of enzyme structure and flexibility by substrate/ligand binding provides an important source of enzyme function regulation. Unfortunately, our understanding of the fundamental mechanisms coupling protein dynamics to biological function is still largely incomplete, therefore limiting our ability to harness protein conformational dynamics in order to regulate enzymatic activity. Here we couple variable temperature (VT) NMR, particularly relaxation dispersion experiments, X-ray crystallography, computer simulations, protein engineering, and enzyme kinetic assays to explore the role of structural heterogeneity and conformational disorder in regulation of the C-terminal substrate binding domain (EIC) of bacterial Enzyme I (EI). In particular, we investigate the relationship between structure, conformational dynamics, and biological function of four EIC constructs: the wild type mesophilic enzyme (*e*EIC), a thermophilic homologue (*t*EIC), and two hybrid constructs engineered by incorporating the active site loops of the mesophilic enzyme into the scaffold of the thermophilic enzyme (*et*EIC), and vice versa (*te*EIC). Through this characterization we provide evidence that the four EIC constructs are structurally similar and that holo EIC undergoes an exchange between a disordered expanded inactive state and a more ordered compact active state. Furthermore, we report that the population of the active state dictates the effective turnover number and that this functional regulation is achieved by tuning the thermodynamic balance between the active and inactive states providing rational for thermal adaption (i.e. why thermophilic homologues exhibit lower activity than their mesophilic counterpart at low temperatures but increased activity comparable to its mesophilic homologue at higher temperatures). We demonstrate that altering thermal stability, conformational flexibility,

and enzymatic activity through the hybridization of mesophilic/thermophilic enzyme pairs is a

promising strategy for protein engineering in the field of biotechnology.

CHAPTER 1.   GENERAL INTRODUCTION

Dissertation Structure

This dissertation is composed of five chapters based on manuscripts in which I am the primary author. These manuscripts were prepared by myself with the help of my major professor and collaborators.

Chapter 1 is a general introduction of the work described in subsequent chapters. This chapter provides an introduction to the topics explored and an overview of my research work conducted while at Iowa State University.

Chapter 2 is a modified manuscript published in Biomolecular NMR titled "^{1}H, ^{15}N, ^{13}C backbone resonance assignment of the C-terminal domain of Enzyme I from *Thermoanaerobacter tengcongensis*". In this publication I report the backbone chemical shift assignment obtained at 70°C by heteronuclear multidimensional NMR of the 70 kDa EIC dimer from a thermophilic bacterium. The reported assignment allows for detailed structural and thermodynamic characterization of the thermophilic homologue of EIC. This work serves as a basis for the backbone assignment performed and reported for two sequent hybrid enzymes described and characterized in the manuscript presented in Chapter 3. In addition, the data from this publication assisted in the work detailed in Chapter 4.

Chapter 3 is a modified manuscript in submission to Science Advances titled "Hybrid thermophilic/mesophilic enzymes reveal a role for conformational disorder in the regulation of bacterial Enzyme I". In this study I designed two chimeric enzymes by hybridizing EIC from mesophilic and thermophilic organisms and provide a detailed characterization for structure, dynamics, and biological function of the four constructs by NMR, crystallography, molecular dynamics simulations, and enzyme kinetic assays.

Chapter 4 is a modified manuscript published in Biomolecular NMR titled "Resonance assignment of the 128 kDa Enzyme I dimer from *Thermoanaerobacter tengcongensis*." In this study I assigned and reported the ^1H_N, ^{15}N, $^{13}C'$, $^1H_{methyl}$, and $^{13}C_{methyl}$ chemical shifts at 70°C of the 128 kDa homodimer EI from the thermophilic bacterium *Thermoanaerobacter tengcongensis*. The chemical shift assignment was achieved through heteronuclear multidimensional NMR of the thermophilic full-length EI in combination with the isolated EIC and EIN backbone and side chain resonance assignments. Backbone assignments and side chain assignments of isolated EIC are reported in Chapter 2 and Chapter 3 of this dissertation, respectively. Additionally, in this work described in Chapter 4, the isolated EIN backbone and side chain assignments were performed and reported for the ^1H_N, ^{15}N, $^{13}C_\alpha$, $^{13}C_\beta$, $^{13}C'$, $^1H_{methyl}$, and $^{13}C_{methyl}$ chemical shifts obtained at 70°C. This contribution serves as a method for the backbone and side chain chemical shift assignments of full-length EI hybrids proposed as a continuation of this project briefly described in Chapter 5.

Chapter 5 is a general conclusion which provides a summary of the major findings and outcomes from manuscripts provided in Chapters 2-4 accomplished by the research work I conducted as a graduate student. In this chapter, future avenues currently being pursued for the extension of this research project are discussed.

Introduction

Significance

Obtaining a comprehensive, atomic resolution understanding of the role played by conformational dynamics in regulation of protein function remains a challenge of great importance. Folded biological proteins are described as an ensemble of conformational states within a unique energy landscape,[1] suggesting they can occupy numerous sub-states simultaneously who's interconversion rates are dictated by the energy barriers intrinsic to the

ensemble. [2, 3] Therefore, protein motion can occur on a wide range of timescales.[4] The landscape of the protein ensemble is dynamic and can be altered by external conditions such as temperature substrate/ligand binding, mutations, pH, pressure, etc.[4, 5] Perturbations of protein dynamics effect a wide variety of biological functions processes such as enzymatic catalysis,[6-11] ligand binding,[12] allostery,[13] thermostability,[14-16] and signaling pathways.[17, 18] Understanding the delicate balance between structure, function, and dynamics that governs biological function will increase our ability to manipulate or completely redirect protein function by selectively stabilizing particular conformations in the ensemble.[19] In addition, accurate knowledge of the fundamental forces that trigger and drive conformational changes in proteins will open new opportunities in drug design, diagnostics, and protein engineering.[20]

Background

Enzyme I (EI) is the first protein of the bacterial phosphotransferase system (PTS), a signal transduction pathway that is involved in both transport and phosphorylation of a large number of carbohydrates in the movement of cells towards there carbon sources (chemotaxis), in biofilm production, in the regulation of interactions between carbon and nitrogen metabolisms, and in the regulation of a number of other metabolic pathways, including catabolic gene expression, potassium transport, and inducer exclusion.[21, 22] The signal for all these different regulatory processes is provided by the phosphorylation state of the PTS components, which varies according to the intracellular availability of phosphoenolpyruvate (PEP).[21, 23] PEP acts as a phosphoryl donor for EI, which initiates a phosphorylation cascade allowing for the sugar uptake in cells. It was shown that inhibition of the EI autophosphorylation reaction impaired PTS-mediated regulatory mechanisms and resulted in reduction of bacterial grown and virulence.[22, 24-29]

EI is a large 128 kDa multidomain homodimer comprising an N-terminal domain (EIN, residues 1-249) that contains the phosphorylation site (His[189]) and the binding site for HPr (the second PTS protein)[30, 31] followed by a C-terminal domain (EIC, residues 261-575) that is responsible for the dimerization, contains the binding site for PEP, and activates PEP for catalysis.[12, 32] The EIN and EIC domains are connected by a short helical linker (residues 250-260).[30, 31] EI catalyzes an autophosphorylation reaction in which a phosphoryl group is transferred from PEP, bound to EIC, to the His[189] residue, located on EIN.[33, 34] The activity of EI depends on coupling among intradomain, interdomain, and intersubunit conformational equilibria. Large scale conformational rearrangements reported for EI include a monomer-dimer transition,[35, 36] an expanded-to-compact conformational change in EIC,[12, 37] and an open-to-close transition reorienting EIN relative to EIC.[30, 31, 37, 38] PEP binding to EIC shifts the conformational equilibria towards the catalytically competent form and activates the enzyme for catalysis.[35, 37]

Rationale

Within the past decade, NMR spectroscopy in conjunction with other structural methods has helped to realize how conformational dynamics assist protein function. These studies have been largely limited to low molecular weight systems while the understanding of how coupling among multiple conformational equilibria determines the activity of large multidomain systems continues to lag. Based on an increasing body of literature that supports the role for structural heterogeneity and conformational disorder in several native enzymes,[19, 39-44] we hypothesize that local equilibria between ordered and disordered conformational states mediate signaling, allostery, and catalysis in complex multidomain proteins. We propose to investigate self-regulatory mechanisms in EI. Clearly EI is a complex system which must be deconvoluted in order to interpret the implications each individual equilibrium has in the regulation of its biological function. Our belief is that intradomain conformational equilibria involved in

regulation of EI are at play in isolated EIN and EIC. In this research work reported we characterize the structure/dynamics/function relationship in the isolated EIC domain in order to gain insight on the expanded-to-compact equilibrium without interference of EIN.[12, 45] Substrate binding stabilizes the compact state of EIC through side chain interactions of active site loops allowing catalysis to take place.[12, 31, 37, 46, 47] We suspect that flexibility of the active site loops, structural heterogeneity, and conformational disorder are key in understanding regulation in EIC.

Approach

We characterized the structure/dynamics/function relationship of EIC from two organisms: a mesophilic bacterium (*Escherichia coli*) and a thermophilic organism (*Thermoanaerobacter tengcongensis*), referred to as *e*EIC and *t*EIC respectively. The mesophilic/thermophilic homologue pair share a similar sequence (overall identity 54%; active site identity 100%) and 3D structure.[31, 46-48] They were reported to have optimal PTS activity at 37°C and 65°C respectively.[46] The higher thermostability of *t*EIC allows us to probe a broader temperature range. Additionally, we designed two hybrid enzymes of EIC to determine how the degree of plasticity of the active site loops affects the conformational dynamics. One hybrid was engineered by incorporating the more flexible active site loops of *e*EIC with the thermostability of the *t*EIC scaffold and is referred to as (*et*EIC), and alternatively the other hybrid was engineered by incorporating the more rigid *t*EIC active site loops with the scaffold of the *e*EIC and is referred to as (*te*EIC).

Innovation

EI is a multidomain oligomeric protein whose phosphoryl transfer activity depends on the coupling among intradomain, interdomain, and intersubunit conformational equilibria. Through an approach combining, NMR, enzyme kinetics, crystallization, and computer simulations, with lower resolution structural data we are able to provide an in-depth atomic resolution

characterization on the temperature dependence of the structure/dynamics/function relationship in isolated EIN and EIC.[49] We propose that this multidisciplinary protocol can be implemented to untangle the specific contributions of each equilibrium to EI and its biological function. In this work we focus our studies on isolated EIC. Designing and characterizing mesophilic/thermophilic EI hybrids (which was done for EIC) allows us to analyze the effect of perturbations of a selected intradomain equilibrium on the structure, dynamics, and biological activity of EI. We believe that the concept of perturbing intradomain dynamics selectively by mesophilic/thermophilic hybrid enzymes is highly innovative and transferable to other complex systems. No molecular system with the complexity exhibited by EI has been previously deconvoluted to provide a robust interpretation of the implication of each conformational equilibrium in regulation of its biological activity. Our hope is to describe with intricacy how multiple equilibria synergistically converge to provide enzyme regulation.

References

1.   Tokuriki, N.; Tawfik, D. S., Protein Dynamism and Evolvability. *Science* **2009,** *324* (5924), 203.

2.   Dill, K. A., Polymer principles and protein folding. *Protein science : a publication of the Protein Society* **1999,** *8* (6), 1166-1180.

3.   Tang, C.; Schwieters, C. D.; Clore, G. M., Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. *Nature* **2007,** *449* (7165), 1078-1082.

4.   Kleckner, I. R.; Foster, M. P., An introduction to NMR-based approaches for measuring protein dynamics. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2011,** *1814* (8), 942-968.

5.   Uversky, V. N., Chapter One - Protein intrinsic disorder and structure-function continuum. In *Progress in Molecular Biology and Translational Science*, Uversky, V. N., Ed. Academic Press: 2019; Vol. 166, pp 1-17.

6.   Bhabha, G.; Lee, J.; Ekiert, D. C.; Gam, J.; Wilson, I. A.; Dyson, H. J.; Benkovic, S. J.; Wright, P. E., A Dynamic Knockout Reveals That Conformational Fluctuations Influence the Chemical Step of Enzyme Catalysis. *Science* **2011,** *332* (6026), 234.

7.   Eisenmesser, E. Z.; Millet, O.; Labeikovsky, W.; Korzhnev, D. M.; Wolf-Watz, M.; Bosco, D. A.; Skalicky, J. J.; Kay, L. E.; Kern, D., Intrinsic dynamics of an enzyme underlies catalysis. *Nature* **2005,** *438* (7064), 117-121.

8.   Fraser, J. S.; Clarkson, M. W.; Degnan, S. C.; Erion, R.; Kern, D.; Alber, T., Hidden alternative structures of proline isomerase essential for catalysis. *Nature* **2009,** *462* (7273), 669-673.

9.   Henzler-Wildman, K. A.; Lei, M.; Thai, V.; Kerns, S. J.; Karplus, M.; Kern, D., A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* **2007,** *450* (7171), 913-916.

10.   Henzler-Wildman, K. A.; Thai, V.; Lei, M.; Ott, M.; Wolf-Watz, M.; Fenn, T.; Pozharski, E.; Wilson, M. A.; Petsko, G. A.; Karplus, M.; Hübner, C. G.; Kern, D., Intrinsic motions along an enzymatic reaction trajectory. *Nature* **2007,** *450* (7171), 838-844.

11.   Tuttle, L. M.; Dyson, H. J.; Wright, P. E., Side-chain conformational heterogeneity of intermediates in the Escherichia coli dihydrofolate reductase catalytic cycle. *Biochemistry* **2013,** *52* (20), 3464-3477.

12.   Venditti, V.; Clore, G. M., Conformational Selection and Substrate Binding Regulate the Monomer/Dimer Equilibrium of the C-terminal domain of Escherichia coli Enzyme I. *Journal of Biological Chemistry* **2012,** *287* (32), 26989-26998.

13.   Motlagh, H. N.; Wrabl, J. O.; Li, J.; Hilser, V. J., The ensemble nature of allostery. *Nature* **2014,** *508* (7496), 331-339.

14.   Liu, Z.; Lemmonds, S.; Huang, J.; Tyagi, M.; Hong, L.; Jain, N., Entropic contribution to enhanced thermal stability in the thermostable P450 CYP119. *Proceedings of the National Academy of Sciences* **2018,** *115* (43), E10049.

15.   Stafford, K. A.; Robustelli, P.; Palmer, A. G., III, Thermal Adaptation of Conformational Dynamics in Ribonuclease H. *PLOS Computational Biology* **2013,** *9* (10), e1003218.

16.   Nguyen, V.; Wilson, C.; Hoemberger, M.; Stiller, J. B.; Agafonov, R. V.; Kutter, S.; English, J.; Theobald, D. L.; Kern, D., Evolutionary drivers of thermoadaptation in enzyme catalysis. *Science* **2017,** *355* (6322), 289.

17.   Tsai, C.-J.; Del Sol, A.; Nussinov, R., Protein allostery, signal transmission and dynamics: a classification scheme of allosteric mechanisms. *Molecular bioSystems* **2009,** *5* (3), 207-216.

18.   Clore, G. M.; Venditti, V., Structure, dynamics and biophysics of the cytoplasmic protein-protein complexes of the bacterial phosphoenolpyruvate: sugar phosphotransferase system. *Trends in biochemical sciences* **2013,** *38* (10), 515-530.

19.   Bonk, B. M.; Weis, J. W.; Tidor, B., Machine Learning Identifies Chemical Characteristics That Promote Enzyme Catalysis. *Journal of the American Chemical Society* **2019,** *141* (9), 4108-4118.

20.   van den Bedem, H.; Fraser, J. S., Integrative, dynamic structural biology at atomic resolution--it's about time. *Nature methods* **2015,** *12* (4), 307-318.

21.   Deutscher, J.; Aké, F. M. D.; Derkaoui, M.; Zébré, A. C.; Cao, T. N.; Bouraoui, H.; Kentache, T.; Mokhtari, A.; Milohanic, E.; Joyet, P., The Bacterial Phosphoenolpyruvate:Carbohydrate Phosphotransferase System: Regulation by Protein Phosphorylation and Phosphorylation-Dependent Protein-Protein Interactions. *Microbiology and Molecular Biology Reviews* **2014,** *78* (2), 231.

22.   Postma, P. W.; Lengeler, J. W.; Jacobson, G. R., Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiological reviews* **1993,** *57* (3), 543-594.

23.   Hogema, B. M.; Arents, J. C.; Bader, R.; Eijkemans, K.; Yoshida, H.; Takahashi, H.; Aiba, H.; Postma, P. W., Inducer exclusion in Escherichia coli by non-PTS substrates: the role of the PEP to pyruvate ratio in determining the phosphorylation state of enzyme IIAGlc. *Molecular Microbiology* **1998,** *30* (3), 487-498.

24.   Huang, K.-J.; Lin, S.-H.; Lin, M.-R.; Ku, H.; Szkaradek, N.; Marona, H.; Hsu, A.; Shiuan, D., Xanthone derivatives could be potential antibiotics: virtual screening for the inhibitors of enzyme I of bacterial phosphoenolpyruvate-dependent phosphotransferase system. *The Journal of Antibiotics* **2013,** *66* (8), 453-458.

25.   Kok, M.; Bron, G.; Erni, B.; Mukhija, S., Effect of enzyme I of the bacterial phosphoenolpyruvate:sugar phosphotransferase system (PTS) on virulence in a murine model. *Microbiology* **2003,** *149* (9), 2645.

26.   Edelstein, P. H.; Edelstein, M. A. C.; Higa, F.; Falkow, S., Discovery of virulence genes of Legionella pneumophila by using signature tagged mutagenesis in a guinea pig pneumonia model. *Proceedings of the National Academy of Sciences* **1999,** *96* (14), 8190.

27.   Hava, D. L.; Camilli, A., Large-scale identification of serotype 4 Streptococcus pneumoniae virulence factors. *Molecular Microbiology* **2002,** *45* (5), 1389-1406.

28.   Jones, A. L.; Knoll, K. M.; Rubens, C. E., Identification of Streptococcus agalactiae virulence genes in the neonatal rat sepsis model using signature-tagged mutagenesis. *Molecular Microbiology* **2000,** *37* (6), 1444-1455.

29.   Lau, G. W.; Haataja, S.; Lonetto, M.; Kensit, S. E.; Marra, A.; Bryant, A. P.; McDevitt, D.; Morrison, D. A.; Holden, D. W., A functional genomic analysis of type 3 Streptococcus pneumoniae virulence. *Molecular Microbiology* **2001,** *40* (3), 555-571.

30.   Schwieters, C. D.; Suh, J.-Y.; Grishaev, A.; Ghirlando, R.; Takayama, Y.; Clore, G. M., 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. *Journal of the American Chemical Society* **2010,** *132* (37), 13026-13045.

31.   Teplyakov, A.; Lim, K.; Zhu, P.-P.; Kapadia, G.; Chen, C. C. H.; Schwartz, J.; Howard, A.; Reddy, P. T.; Peterkofsky, A.; Herzberg, O., Structure of phosphorylated enzyme I, the phosphoenolpyruvate:sugar phosphotransferase system sugar translocation signal protein. *Proceedings of the National Academy of Sciences* **2006,** *103* (44), 16218.

32.   Chauvin, F.; Brand, L.; Roseman, S., Enzyme I: the first protein and potential regulator of the bacterial phosphoenolpyruvate: glycose phosphotransferase system. *Research in Microbiology* **1996,** *147* (6), 471-479.

33.   Weigel, N.; Waygood, E. B.; Kukuruzinska, M. A.; Nakazawa, A.; Roseman, S., Sugar transport by the bacterial phosphotransferase system. Isolation and characterization of enzyme I from Salmonella typhimurium. *Journal of Biological Chemistry* **1982,** *257* (23), 14461-14469.

34.   Weigel, N.; Kukuruzinska, M. A.; Nakazawa, A.; Waygood, E. B.; Roseman, S., Sugar transport by the bacterial phosphotransferase system. Phosphoryl transfer reactions catalyzed by enzyme I of Salmonella typhimurium. *Journal of Biological Chemistry* **1982,** *257* (23), 14477-14491.

35.   Nguyen, T. T.; Ghirlando, R.; Venditti, V., The oligomerization state of bacterial enzyme I (EI) determines EI's allosteric stimulation or competitive inhibition by α-ketoglutarate. *Journal of Biological Chemistry* **2018**.

36.   Patel, H. V.; Vyas, K. A.; Savtchenko, R.; Roseman, S., The Monomer/Dimer Transition of Enzyme I of the Escherichia coli Phosphotransferase System. *Journal of Biological Chemistry* **2006,** *281* (26), 17570-17578.

37.   Venditti, V.; Tugarinov, V.; Schwieters, C. D.; Grishaev, A.; Clore, G. M., Large interdomain rearrangement triggered by suppression of micro- to millisecond dynamics in bacterial Enzyme I. *Nature Communications* **2015,** *6* (1), 5960.

38.   Venditti, V.; Schwieters, C. D.; Grishaev, A.; Clore, G. M., Dynamic equilibrium between closed and partially closed states of the bacterial Enzyme I unveiled by solution NMR and X-ray scattering. *Proceedings of the National Academy of Sciences* **2015,** *112* (37), 11565.

39.   Palombo, M.; Bonucci, A.; Etienne, E.; Ciurli, S.; Uversky, V. N.; Guigliarelli, B.; Belle, V.; Mileo, E.; Zambelli, B., The relationship between folding and activity in UreG, an intrinsically disordered enzyme. *Scientific Reports* **2017,** *7* (1), 5977.

40.   Larion, M.; Miller, B.; Brüschweiler, R., Conformational heterogeneity and intrinsic disorder in enzyme regulation: Glucokinase as a case study. *Intrinsically Disordered Proteins* **2015,** *3* (1), e1011008.

41.   DeForte, S.; Uversky, V. N., Not an exception to the rule: the functional significance of intrinsically disordered protein regions in enzymes. *Molecular BioSystems* **2017,** *13* (3), 463-469.

42.   Ghose, R., Nature of the Pre-Chemistry Ensemble in Mitogen-Activated Protein Kinases. *Journal of Molecular Biology* **2019,** *431* (2), 145-157.

43.   Glowacki, D. R.; Harvey, J. N.; Mulholland, A. J., Taking Ockham's razor to enzyme dynamics and catalysis. *Nature Chemistry* **2012,** *4* (3), 169-176.

44.   Liao, Q.; Kulkarni, Y.; Sengupta, U.; Petrović, D.; Mulholland, A. J.; van der Kamp, M. W.; Strodel, B.; Kamerlin, S. C. L., Loop Motion in Triosephosphate Isomerase Is Not a Simple Open and Shut Case. *Journal of the American Chemical Society* **2018,** *140* (46), 15889-15903.

45.   Dotas, R. R.; Venditti, V., 1H, 15N, 13C backbone resonance assignment of the C-terminal domain of enzyme I from Thermoanaerobacter tengcongensis. *Biomolecular NMR Assignments* **2018,** *12* (1), 103-106.

46.   Navdaeva, V.; Zurbriggen, A.; Waltersperger, S.; Schneider, P.; Oberholzer, A. E.; Bähler, P.; Bächler, C.; Grieder, A.; Baumann, U.; Erni, B., Phosphoenolpyruvate: Sugar Phosphotransferase System from the Hyperthermophilic Thermoanaerobacter tengcongensis. *Biochemistry* **2011,** *50* (7), 1184-1193.

47.   Oberholzer, A. E.; Bumann, M.; Schneider, P.; Bächler, C.; Siebold, C.; Baumann, U.; Erni, B., Crystal Structure of the Phosphoenolpyruvate-binding Enzyme I-Domain from the Thermoanaerobacter tengcongensis PEP: Sugar Phosphotransferase System (PTS). *Journal of Molecular Biology* **2005,** *346* (2), 521-532.

48.   Evangelidis, T.; Nerli, S.; Nováček, J.; Brereton, A. E.; Karplus, P. A.; Dotas, R. R.; Venditti, V.; Sgourakis, N. G.; Tripsianes, K., Automated NMR resonance assignments and structure determination using a minimal set of 4D spectra. *Nature Communications* **2018,** *9* (1), 384.

49.   Dotas, R. R.; Venditti, V., Resonance assignment of the 128 kDa enzyme I dimer from Thermoanaerobacter tengcongensis. *Biomolecular NMR Assignments* **2019,** *13* (2), 287-293.

CHAPTER 2.   1H, 15N, 13C BACKBONE RESONANCE ASSIGNEMENT OF THE C-TERMINAL DOMAIN OF ENZYME I FROM *THERMOANAEROBACTER TENGCONGENSIS*

Rochelle Dotas[1] and Vincenzo Venditti[1, 2]

Author Affiliations

[1] Department of Chemistry, Iowa State University

[2] Roy J. Carver Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University

Journal Information

Abstract

Phosphoenolpyruvate (PEP) binding to the C-terminal domain (EIC) of Enzyme I (EI) of the bacterial phosphotransferase system (PTS) initiates a phosphorylation cascade that results in sugar translocation across the cell membrane and controls a large number of essential pathways in bacterial metabolism. EIC undergoes an expanded to compact conformational equilibrium that is regulated by ligand binding and determines the phosphorylation state of the overall PTS. Here, we report the backbone 1H, 15N and 13C chemical shift assignments of the 70 kDa EIC dimer from the thermophilic bacterium Thermoanaerobacter tengcongensis. Assignments were obtained at 70°C by heteronuclear multidimensional NMR spectroscopy. In total, 90 % of all backbone resonances were assigned, with 264 out of a possible 299 residues assigned in the 1H-15N TROSY spectrum. The secondary structure predicted from the assigned backbone resonance using the program TALOS+ is in good agreement with the X-ray crystal structure of T. tengcongensis EIC. The reported assignments will allow detailed structural and thermodynamic investigations on the coupling between ligand binding and conformational dynamics in EIC.

Biological Context

Enzyme I (EI) is the first protein of the bacterial phosphoenolpyruvate (PEP):carbohydrate phosphotransferase system (PTS) (Clore and Venditti 2013), a phosphorylation cascade that is involved in both transport and phosphorylation of a large number of carbohydrates (PTS carbohydrates), in the movement of cells toward these carbon sources (chemotaxis), in biofilm formation, in the regulation of interactions between carbon and nitrogen metabolisms, and in the regulation of a number of other metabolic pathways, including catabolic gene expression, potassium transport, and inducer exclusion (Deutscher et al. 2014; Postma et al. 1996). All the regulatory functions of PTS are controlled by the intracellular concentration of PEP, which determines the phosphorylation state of the PTS components by acting as phosphoryl donor for EI (Deutscher et al. 2014). Therefore, EI plays a central role in the control and regulation of bacterial metabolism, and represents a promising target for antimicrobial research.

The 3D structure of EI shows that the enzyme comprises two structurally and functionally distinct domains separated by a helical linker (Schwieters et al. 2010; Teplyakov et al. 2006). The N-terminal domain (EIN) contains the phosphorylation site (His^{189}) and is responsible for transferring the phosphoryl group to the second PTS protein (the phosphocarrier protein HPr). The C-terminal domain (EIC) is responsible for protein dimerization and contains the binding site for PEP. Biophysical and biochemical investigations on the *E. coli* EI revealed that the enzyme is functionally regulated by an intricate series of intra- and inter-domain conformational equilibria including: 1) a monomer/dimer equilibrium via EIC (Patel et al. 2006), 2) a compact/expanded equilibrium within EIC (Venditti and Clore 2012; Venditti et al. 2015b), 3) a state A/state B equilibrium within EIN (Schwieters et al. 2010; Teplyakov et al. 2006; Venditti et al. 2015b), and 4) an open/closed equilibrium that involves reorientation of EIN relative to EIC (Schwieters et al. 2010; Teplyakov et al. 2006; Venditti et al. 2015b). The state of

each equilibrium is tightly controlled by PEP binding to EIC that perturbs the compact/expanded equilibrium and promotes an overall shift toward the activated EI structure (Venditti and Clore 2012; Venditti et al. 2013).

Here we report the ^{1}H, ^{15}N, ^{13}C backbone resonance assignment of the 70 kDa dimer of EIC from *Thermoanaerobacter tengcongensis*. The extended thermal stability of the *T. tengcongensis* protein (melting point = 90°C) (Navdaeva et al. 2011; Oberholzer et al. 2005) compared to its *E. coli* analogue will enable future investigations on the temperature dependence of the interplay between ligand binding and EIC conformational dynamics that has been proposed to regulate the activity of EI and to control the phosphorylation state of the overall PTS (Venditti and Clore 2012; Venditti et al. 2013; Venditti et al. 2015a; Venditti et al. 2015b).

Materials and Methods

Protein Expression and Purification

A construct including *T. tengcongensis* EIC (residue 261-573), a N-terminal Maltose Binding Protein (MBP)-tag, and a N-terminal His-tag (His$_6$-MBP-EIC) has been cloned into a pET11a vector (Novagen). The MBP and EIN portion of the construct were separated by the TEV-protease consensus sequence ENLYFQ\S (where '\' denotes the cleaved peptide bond). The plasmid was introduced into *E. coli* strain BL21star(DE3) (Invitrogen) and the transformed bacteria were plated onto a LB-agar plate containing ampicillin (100 µg/ml) for selection. Cells were grown at 37°C in D$_2$O-based minimal medium. ^{15}NH$_4$Cl and ^{13}C$_6$-glucose were used as the sole nitrogen and carbon sources, respectively. Uniformly deuterated ^{13}C$_6$-glucose was used as the carbon source in D$_2$O-based cultures. At A$_{600}$ ~ 0.8 the temperature was reduced to 20°C and expression was induced with 1 mM isopropyl-D-thiogalactopyranoside (IPTG). Cells were harvested by centrifugation (40 min at 4,000g) after 16 h of induction and the pellet was resuspended in 20 ml of 20 mM Tris (pH 8.0) and 100mM NaCl. The suspension was lysed

using a microfluidizer and centrifuged at 50,000g for 40 min. The supernatant was filtered and loaded onto two connected HisTrap columns (5 ml each; GE Healthcare), and the protein was eluted with a 100 ml gradient of 0.35 M imidazole. The fractions containing the protein were confirmed by SDS-polyacrylamide gel electrophoresis and incubated at room temperature with 1 mg of TEV protease overnight. The solution was heated to 80°C for two hours, and then centrifuged at 50,000g for 40 min. The protein was further purified by gel filtration on a Superdex-75 column (GE Healthcare) equilibrated with 20 mM Tris, pH 7.4, 200 mM NaCl, 2 mM DTT, and 1 mM EDTA.

NMR Spectroscopy

All the NMR samples were prepared in 20 mM Tris buffer, pH 7.4, 100 mM NaCl, 4 mM MgCl$_2$, 1 mM EDTA, 2 mM DTT, and 90% H$_2$O/10% D$_2$O (v/v). The protein concentration (in subunits) was 0.5-1.0 mM. NMR spectra were recorded at 70°C on Bruker 700 and 600 MHz spectrometers equipped with a z-shielded gradient triple resonance cryoprobe. Spectra were processed using NMRPipe (Delaglio et al. 1995) and analyzed using the program SPARKY (http://www.cgl.ucsf.edu/home/sparky). Sequential ^{1}H/^{15}N/^{13}C backbone assignment of the EIC domain was carried out using transverse relaxation optimized (TROSY) versions (Pervushin et al. 1997; Tugarinov et al. 2002) of conventional 3D triple resonance correlation experiments (HNCO, HNCA, HNCACB, HN(CO)CA, and HN(CO)CACB) (Clore and Gronenborn 1998).

<div align="center">

Resonance Assignment and Data Deposition

</div>

Backbone ^{1}H$_N$, ^{15}N, ^{13}C$_\alpha$, ^{13}C$_\beta$ and ^{13}C' chemical shifts of *T. tengcongensis* EIC were assigned using triple resonance methodology (Clore and Gronenborn 1998) and deposited in the BioMagResBank (http://www.bmrb.wisc.edu/) (Ulrich et al. 2008) under the BMRB accession code 27183. Excluding the 14 proline residues, 264 out of a total of 299 residues were assigned in the ^{1}H-^{15}N TROSY spectrum of EIC (Fig. 1). In total, 90 % of all backbone resonances were

assigned (89 % of ^1H_N, 89 % of ^{15}N, 93 % of $^{13}C_\alpha$, 92 % of $^{13}C_\beta$ and 89 % of $^{13}C'$ nuclei). The 35 unassigned peaks in the ^1H-^{15}N TROSY spectrum predominantly cluster at the dimer interface (K^{349}-N^{352}, and L^{484}-I^{489}) and at the PEP binding site (N^{454}-Y^{459}, Fig. 2a). The fact that millisecond timescale conformational dynamics were observed at the dimer interface and at the PEP binding site in *E. coli* EIC (Venditti and Clore 2012; Venditti et al. 2015b) suggests that these resonances are likely broadened out beyond detection in the ^1H-^{15}N TROSY spectrum.

The secondary structure propensity of EIC was predicted from the assigned backbone ^1H_N, ^{15}N, $^{13}C_\alpha$, $^{13}C_\beta$ and $^{13}C'$ chemical shifts using the program TALOS+ (Shen et al. 2009). The predicted secondary structure is compared to the crystal structure of *T. tengcongensis* EIC in Fig. 2b. The excellent agreement between NMR and crystallographic data indicates that tertiary structure of EIC in solution is essentially the same as that in the crystal state, and provides confidence in the obtained resonance assignment.

Acknowledgements

References

Clore GM, Gronenborn AM (1998) Determining the structures of large proteins and protein complexes by NMR. Trends Biotechnol 16:22-34

Clore GM, Venditti V (2013) Structure, dynamics and biophysics of the cytoplasmic protein-protein complexes of the bacterial phosphoenolpyruvate: sugar phosphotransferase system. Trends Biochem Sci 38:515-530 doi:10.1016/j.tibs.2013.08.003

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277-293

Deutscher J, Ake FM, Derkaoui M, Zebre AC, Cao TN, Bouraoui H, Kentache T, Mokhtari A, Milohanic E, Joyet P (2014) The bacterial phosphoenolpyruvate:carbohydrate phosphotransferase system: regulation by protein phosphorylation and phosphorylation-dependent protein-protein interactions. Microbiol Mol Biol Rev 78:231-256 doi:10.1128/MMBR.00001-14

Navdaeva V, Zurbriggen A, Waltersperger S, Schneider P, Oberholzer AE, Bahler P, Bachler C, Grieder A, Baumann U, Erni B (2011) Phosphoenolpyruvate: sugar phosphotransferase system from the hyperthermophilic Thermoanaerobacter tengcongensis. Biochemistry 50:1184-1193

Oberholzer AE, Bumann M, Schneider P, Bachler C, Siebold C, Baumann U, Erni B (2005) Crystal structure of the phosphoenolpyruvate-binding enzyme I-domain from the Thermoanaerobacter tengcongensis PEP: sugar phosphotransferase system (PTS). J Mol Biol 346:521-532

Patel HV, Vyas KA, Savtchenko R, Roseman S (2006) The monomer/dimer transition of enzyme I of the Escherichia coli phosphotransferase system. J Biol Chem 281:17570-17578 doi:10.1074/jbc.M508965200

Pervushin K, Riek R, Wider G, Wuthrich K (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci U S A 94:12366-12371

Postma PW, Lengeler JW, Jacobson GR (1996) Phosphoenolpyruvate:Carbohydrate Phosphotransferase Systems. In: Neidhardt FC, Lin EC, Curtiss R (eds) Escherichia coli and Salmonella: Cellular and Molecular Biology. pp 1149-1174

Schwieters CD, Suh JY, Grishaev A, Ghirlando R, Takayama Y, Clore GM (2010) 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. J Am Chem Soc 132:13026-13045

Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J Biomol NMR 44:213-223

Teplyakov A, Lim K, Zhu PP, Kapadia G, Chen CC, Schwartz J, Howard A, Reddy PT, Peterkofsky A, Herzberg O (2006) Structure of phosphorylated enzyme I, the phosphoenolpyruvate:sugar phosphotransferase system sugar translocation signal protein. Proc Natl Acad Sci U S A 103:16218-16223

Tugarinov V, Muhandiram R, Ayed A, Kay LE (2002) Four-dimensional NMR spectroscopy of a 723-residue protein: chemical shift assignments and secondary structure of malate synthase g. J Am Chem Soc 124:10025-10035

Ulrich EL, Akutsu H, Doreleijers JF, Harano Y, Ioannidis YE, Lin J, Livny M, Mading S, Maziuk D, Miller Z, Nakatani E, Schulte CF, Tolmie DE, Kent Wenger R, Yao H, Markley J (2008) BioMagResBank. Nucleic Acids Res 36:D402-D408

Venditti V, Clore GM (2012) Conformational selection and substrate binding regulate the monomer/dimer equilibrium of the C-terminal domain of Escherichia coli enzyme I. J Biol Chem 287:26989-26998 doi:10.1074/jbc.M112.382291

Venditti V, Ghirlando R, Clore GM (2013) Structural basis for enzyme I inhibition by alpha-ketoglutarate. ACS Chem Biol 8:1232-1240 doi:10.1021/cb400027q

Venditti V, Schwieters CD, Grishaev A, Clore GM (2015a) Dynamic equilibrium between closed and partially closed states of the bacterial Enzyme I unveiled by solution NMR and X-ray scattering. Proc Natl Acad Sci U S A 112:11565-11570 doi:10.1073/pnas.1515366112

Venditti V, Tugarinov V, Schwieters CD, Grishaev A, Clore GM (2015b) Large interdomain rearrangement triggered by suppression of micro- to millisecond dynamics in bacterial Enzyme I. Nat Commun 6:5960 doi:10.1038/ncomms6960
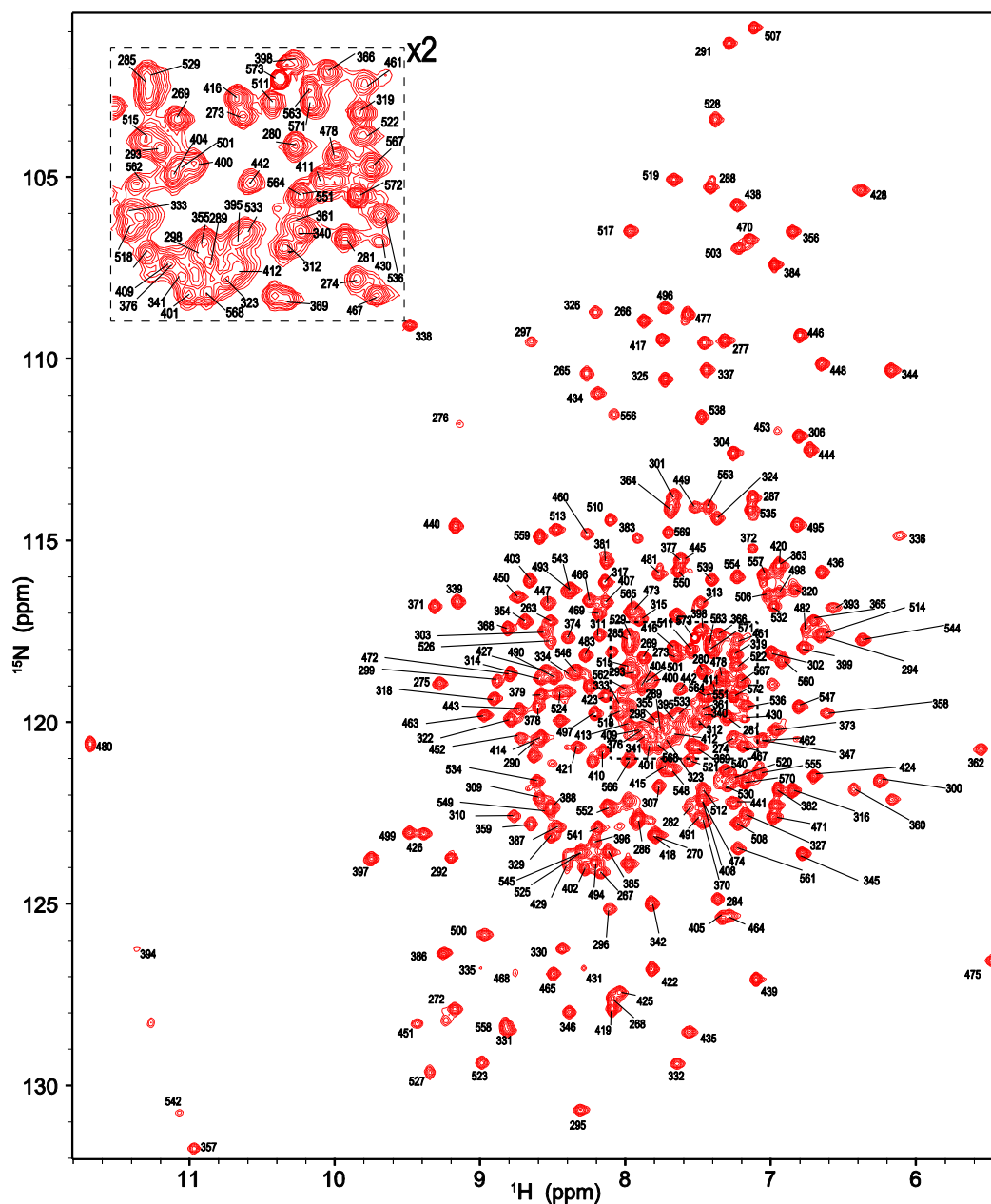
Figures

Fig. 1 700 MHz ^{1}H-^{15}N TROSY spectrum of the 70 kDa dimer of *T. tengcongensis* EIC at pH 7.4 and 70°C. Assigned cross-peaks are labeled with the corresponding residue numbers. The central region of the spectrum (central box) is expanded (x2) in the top left corner
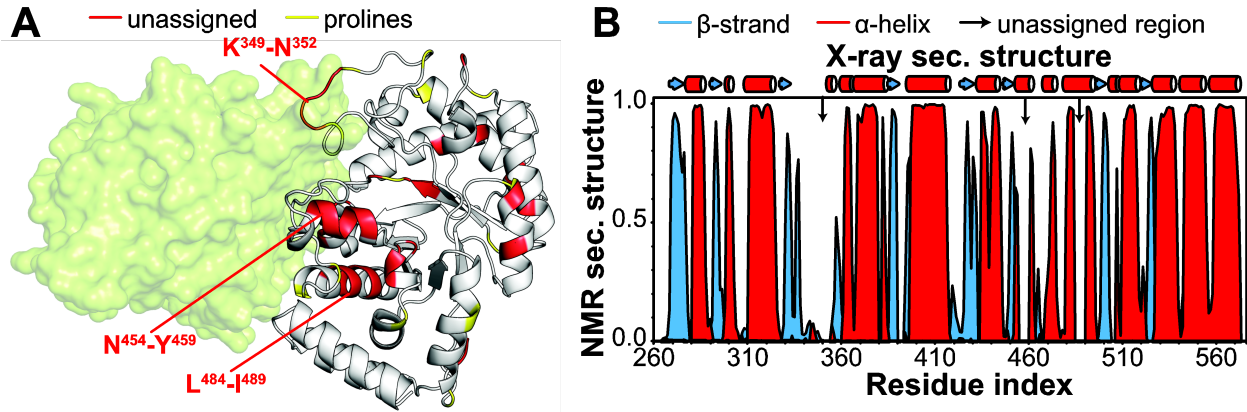
Fig. 2 (**a**) Crystallographic structure of the *T. tengcongensis* EIC (PDB code:2BG5). Proline residues are colored yellow, and all unassigned residues are colored red. The second monomeric subunit is shown as transparent green surface. (**b**) Secondary structure propensity versus residue index. Secondary structure propensity was calculated using the assigned backbone resonances and the software TALOS+. Blue and red curves denote β-strand and α-helix propensities, respectively. Arrows highlight stretches of contiguous unassigned residues (K^{349}-N^{352}, N^{454}-Y^{459}, and L^{484}-I^{489}). The secondary structure derived from the X-ray crystal structure of *T. tengcongensis* EIC is also shown (top) for comparison

CHAPTER 3.   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]

Author Affiliations

[1] Department of Chemistry, Iowa State University

[2] Macromolecular X-ray Crystallography Facility, Office of Biotechnology, Iowa State

University

[3] Laboratory of Molecular Biology, NIDDK, National Institutes of Health

[4] Roy J. Carver Department of Biochemistry, Biophysics and Molecular Biology, Iowa State

University

Abstract

Conformational disorder is emerging as an important feature of biopolymers, regulating a

vast array of cellular functions, including signaling, phase separation, and enzyme catalysis. Here

we combine NMR, crystallography, computer simulations, protein engineering, and functional

assays to investigate the role played by conformational heterogeneity in determining the activity

of the C-terminal domain of bacterial Enzyme I (EIC). In particular, we design chimeric proteins

by hybridizing EIC from thermophilic and mesophilic organisms, and we characterize the

resulting constructs for structure, dynamics, and biological function. We show that EIC exists as

a mixture of active and inactive conformations and that functional regulation is achieved by

tuning the thermodynamic balance between active and inactive states. Interestingly, we also

present a hybrid thermophilic/mesophilic enzyme that is thermostable and more active than the wild-type thermophilic enzyme, suggesting that hybridizing thermophilic and mesophilic proteins is a valid strategy to engineer thermostable enzymes with significant low-temperature activity.

Introduction

Enzyme I (EI) is the first protein in the phosphoenolpyruvate (PEP)-dependent sugar phosphotransferase system (PTS), a signal transduction pathway that controls sugar uptake in bacteria and regulates several other cellular functions, including catabolic gene expression, coupling between central nitrogen and carbon metabolism, chemotaxis, and biofilm formation, via phosphorylation dependent protein-protein interactions.[1-5] The phosphorylation state of the PTS components is controlled by the intracellular level of PEP, which initiates the PTS by acting as phosphoryl donor for EI.[2, 6] As such, inhibition of the EI autophosphorylation reaction was shown to impair PTS-mediated regulatory mechanisms and to result in reduced bacterial growth and virulence.[3, 7-12]

EI is a 128 kDa homodimer composed of two structurally and functionally distinct domains separated by a 11-residues helical linker.[13] The phosphoryl-transfer N-terminal domain (EIN, residues 1-249) contains the phosphorylation site (His[189]) and the binding site for HPr (the second PTS protein). The C-terminal domain (EIC, residues 261-575) contains the binding site for PEP, is responsible for EI dimerization,[13] and activates PEP for catalysis.[14] In particular, EIC locks the substrate into a high energy conformation in which the phosphate group lays outside of the plane formed by the three conjugated sp^2 carbons, and forms salt-bridges with the small molecule that provide stabilization to the transition state of the phosphoryl-transfer reaction (Figure 1a). Functional regulation of EI is achieved through synergistic coupling of multiple intra and interdomain conformational equilibria that are modulated by substrate and cofactor binding

(Figure 1b). Specifically, EI was shown to undergo (i) a monomer-dimer equilibrium,[15, 16] (ii) a compact-to-expanded equilibrium within the EIC domain,[14, 17] (iii) a g^+-to-g^- equilibrium within the rotameric state of the His[189] side chain,[18] (iv) a state A-to-state B equilibrium within the EIN domain,[19, 20] and (v) an open-to-close equilibrium describing a reorientation of EIN relative to EIC.[17, 19-21] Binding of PEP to EIC promotes transition to the dimer/compact/g^-/state B/closed form and activates the enzyme for catalysis.[15, 17] Therefore, in addition to being a promising pharmaceutical target, EI is also an important model system to study the interplay between ligand binding, post-translational modifications, and conformational dynamics that determines the activity of complex multidomain enzymes.

In this contribution, we investigate the role played by EIC intradomain conformational dynamics in activation of PEP. In particular, we engineer hybrid thermophilic/mesophilic EIC constructs with modulated thermal stability and active site flexibility, and we test their ability to catalyze hydrolysis of PEP. Although non-physiological, the latter enzymatic reaction has the advantage of being independent from the EIN domain,[14] therefore allowing us to assess the contribution of EIC to PEP activation without interference from the open-to-close equilibrium of the full-length enzyme. Our data demonstrate that, in the presence of substrate, EIC does not adopt a single, catalytically active conformation but exists as a collection of active and inactive states. The thermodynamic balance between active and inactive states is finely controlled by the amino acid sequence of EIC, with the mesophilic enzyme favoring active conformations and the thermophilic enzyme favoring inactive states. Rescaling the turnover number by the fractional population of active EIC is sufficient to recapitulate the difference in activity observed among all EIC variants analyzed here. Our results provide evidence for a regulatory role of conformational disorder in activation of PEP by EIC, which supports an increasing body of literature suggesting

that conformational disorder and structural heterogeneity control the activity of several native enzymes.[22-28]

Results and Discussion

In this work, we investigate the structure/dynamics/function relationship of EIC from two organisms: a mesophilic bacterium (*Escherichia coli*) and a thermophilic organism (*Thermoanaerobacter tengcongensis*). The two proteins will be referred to as *e*EIC and *t*EIC, respectively, throughout the manuscript. Similar notation will be used for the full-length enzymes (*e*EI and *t*EI, respectively). *e*EI and *t*EI share a similar sequence (overall identity 54%; EIC identity 60%; active site identity 100%, Supplementary Figure S1) and 3D structure.[20, 29-31] While the temperature dependence of *e*EI and *t*EI activities has not been investigated yet, the *E. coli* and *T. tengcongensis* PTS's were shown to be optimally active at 37 °C and 65 °C, respectively.[30]

Temperature-Dependence of the Expanded-to-Compact Equilibrium

EIC displays a $(\beta/\alpha)_8$-barrel fold with active site residues localized at the C-terminal ends of the β-barrel domain, and within the β2α2 (residues 296-309), β3α3 (residues 332-360), and β6α6 (residues 454-477) loops (Figure 2a and Supplementary Figure S1). The active site loops of *e*EIC were shown to undergo conformational dynamics in the apo form between a lowly populated state that closely resembles the PEP-bound conformation, and a major state in which the active site is likely more expanded.[14, 17] Binding of PEP was observed to shift this expanded-to-compact conformational equilibrium toward the compact form.[14, 17] In the full-length *e*EI, quenching of the C-terminal domain dynamics induced by PEP-binding triggers the open-to-close interdomain rearrangement and activates the enzyme for the autophosphorylation reaction (Figure 1b).[17] Here, the temperature dependence of the expanded-to-compact equilibrium in *e*EIC and *t*EIC was investigated by means of ^{15}N and ^{13}C$_{\text{methyl}}$ relaxation dispersion (RD)

NMR.[32] Experiments were acquired on U-[^{2}H,^{15}N]/Ile(d$_1$)-^{13}CH$_3$/Val, Leu-(^{13}CH$_3$/^{12}C$_2$H$_3$)-labeled EIC at two static fields (600 and 800 MHz) and four different temperatures (25, 30, 35, and 40 °C for eEIC, and 40, 50, 60, and 70 °C for tEIC). Simultaneous investigation of RD data measured at multiple temperatures returns a deep characterization of the kinetics and thermodynamics of conformational exchange processes between species with distinct chemical shifts occurring on a timescale ranging from ~100 μs to ~10 ms, by providing enthalpy (ΔH), entropy (ΔS), activation enthalpy ($\Delta^{\ddagger}H$), and activation entropy ($\Delta^{\ddagger}S$) for the conformational equilibrium.[33]

As expected, several residues within and in the surrounding of the active site loops of eEIC and tEIC display exchange contribution to the transverse relaxation rate (R_{ex}) significantly larger than zero (Figure 2b), which is indicative of conformational dynamics in the μs-ms timescale. Of note, while for tEIC residues with positive R_{ex} are confined within the area of the active site loops, μs-ms timescale dynamics in eEIC propagate to the β-barrel region of the enzyme (Figure 2b and Supplementary Figure S2). All RD curves with $R_{ex} > 5$ s^{-1} at 25 and 40 °C for eEIC and tEIC, respectively, were selected for further analysis and used to obtain a quantitative description of the detected conformational equilibria. Interestingly, for both the mesophilic and thermophilic constructs, all the analyzed ^{15}N and ^{13}C$_{methyl}$ RD curves measured at multiple temperatures and static magnetic fields could be fit simultaneously to a model describing the interconversion of two conformational states (Figure 2c and Supplementary Figures S2 and S3), confirming that the active site in apo EIC undergoes concerted motion between two conformations. In this global fitting procedure, the activation ($\Delta^{\ddagger}G$) and standard (ΔG) free energy of the conformational equilibrium were optimized as global parameters, whereas the ^{15}N and ^{13}C chemical shift differences between the two conformational states ($\Delta\omega_N$

and $\Delta\omega_C$, respectively) were treated as peak-specific and temperature independent parameters. The exchange rate (k_{ex}) and the fractional population of the minor conformational state (p_b) were calculated at each temperature from the fitted values of $\Delta^{\ddagger}G$ and ΔG using the general form of the Eyring and reaction isotherm equations, respectively (see Methods). This fitting procedure effectively reduces the number of optimized parameters and is justified if the heat capacity of activation remains constant over the experimental temperature range[34] (25-40 °C and 40-70 °C for eEIC and tEIC, respectively). An example of the fit is provided in Figure 2c. The full set of experimental RD curves used in the fitting procedure is shown in Supplementary Figures S2 and S3 for eEIC and tEIC, respectively. The good agreement between experimental and back-calculated data indicates that no additional variable parameter is required to fulfill the experimental RD curves. Values for the forward and backward rate constants (k_{ab} and k_{ba}, respectively) at the different experimental temperatures were calculated from corresponding values of k_{ex} (= k_{ab} + k_{ba}) and p_b. The temperature dependence of p_b, k_{ab}, and k_{ba} was modeled using the van't Hoff (for p_b) and Eyring (for k_{ab} and k_{ba}) equations to obtain ΔH, ΔS, $\Delta^{\ddagger}H$, and $\Delta^{\ddagger}S$ of the expanded-to-compact conformational equilibrium. All thermodynamic, kinetic, and chemical shift parameters obtained for the two enzymes are summarized in Figure 2d,e, 3b, and Supplementary Table S1. From these data it is apparent that the expanded-to-compact equilibrium occurs with negative ΔH and ΔS (Supplementary Table 1), which explains the decrease in p_b (i.e. the population of compact state) with increasing temperature (Figure 1d), and is consistent with the hypothesis that the compact state is more ordered and structurally stabilized by a higher number of interatomic contacts than the expanded conformation.[14, 17]

Addition of substrate to tEIC causes line-broadening beyond detection level for the NMR peaks of the residues directly facing PEP. However, large ^{1}H/^{15}N ($\Delta_{H/N}$) and ^{1}H/^{13}C$_{methyl}$ ($\Delta_{H/C}$)

chemical shift perturbations can still be observed within the active site of tEIC upon substrate binding (Figure 3a), suggesting that structural rearrangements are occurring within this region of the enzyme. As previously reported for eEIC,[14, 17] the perturbations on the ^{15}N (Δ_N) and ^{13}C$_{methyl}$ (Δ_C) chemical shifts of tEIC observed upon PEP binding are in good agreement with the $\Delta\omega_N$ and $\Delta\omega_C$ values obtained from the RD measurements on apo tEIC (Figure 3b), supporting the hypothesis that binding of PEP shifts the expanded-to-compact equilibrium detected in apo EIC toward the compact state.[14, 17] In this respect, it is also interesting to note that the measured $\Delta_{H/N}$ and $\Delta_{H/C}$ values increase with increasing temperature and are inversely proportional to the population of compact state in apo tEIC (Figure 3c). As chemical shifts in fast-exchanging systems are population weighted quantities (note that free and PEP-bound EIC are in fast exchange on the chemical shift timescale, Figure 3d), these observations further confirm that PEP-binding shifts the pre-existing equilibrium between the expanded and compact EIC conformations. Consistent with this conformational-shift model, analysis of the structural models of holo EIC[14, 17, 20, 30] shows that salt bridges between the phosphate group of PEP and the side chains of Arg^{296}, Arg^{332}, Lys^{340}, Arg^{358}, and Arg^{465} provide structural stabilization to the active site loops in the compact form (Supplementary Figure S1b). Interestingly, these electrostatic interactions also result in near complete suppression of the R_{ex}'s measured for tEIC (Figure 3e). While these results indicate unambiguously that PEP binding has a dramatic effect on the expanded-to-compact equilibrium, it is not possible to ascertain if the R_{ex} suppression observed in the RD data upon substrate binding is the result of (i) a complete quench of the expanded-to-compact equilibrium, (ii) a shift of the kinetics for the expanded-to-compact equilibrium outside the μs-ms timescale regime detectable by RD experiments, or (iii) a drastic reduction of $\Delta\omega_N$ and

$Δω_C$ between the expanded and compact conformations in holo EIC compared to the apo enzyme.

Engineering of Hybrid *e*EIC/*t*EIC Constructs

The analysis of the temperature dependence of the expanded-to-compact equilibrium has revealed similar thermodynamic parameters for the conformational change in apo eEIC and tEIC ($ΔH$ and $ΔS$ in Supplementary Table S1), which is reflected in the similar temperature dependence of p_b obtained for the two EIC constructs (Figure 2d). On the other hand, the RD NMR data indicate the existence of significant differences in the kinetics of the conformational equilibrium in the mesophilic and thermophilic enzyme, with the expanded-to-compact transition occurring with a ~4-fold higher k_{ex} in *e*EIC than in *t*EIC at 40 °C (Figure 2e). Interestingly, plotting k_{ex} versus temperature reveals that the two enzymes display similar exchange rate constant at their optimal PTS temperature (37 and 65 °C for eEIC and tEIC, respectively) (Figure 2e), which suggests a possible role for the expanded-to-compact equilibrium in functional regulation of EIC.

To investigate the relationship between activity and active site conformational dynamics in EIC we have engineered hybrid thermophilic/mesophilic constructs by merging the scaffold of one enzyme with the active site loops of the other enzyme (Figure 2a). In particular, the hybrid formed by the active site loops of the mesophilic protein and the scaffold of the thermophilic enzyme is referred to as *et*EIC, while the hybrid comprised of the active site loops of the thermophilic protein and the scaffold of the mesophilic enzyme is referred to as *te*EIC. This protein engineering effort required 21 single-point mutations per hybrid construct. A list of the performed mutations and their localization within the EIC structure are provided in Supplementary Figure S1. The 21 mutations point away from the catalytic site (Supplementary Figure S1c) and do not perturb the fold and oligomeric state of the enzyme, as revealed by X-ray

crystallography and analytical ultracentrifugation (AUC) (Figure 4). On the other hand, RD NMR data indicate that the mutations have a profound effect on the μs-ms timescale conformational dynamics of eEIC and tEIC. Indeed, incorporating the mesophilic active site loops into the scaffold of the thermophilic enzyme is sufficient to increase the rate of the active site dynamics at the same level observed for the mesophilic enzyme (note that eEIC and etEIC have very similar temperature dependence of k_{ex}, Figure 2e). In contrast, implanting the active site loops of tEIC onto the scaffold of eEIC increases the activation energy for the expanded-to-compact conformational change and results in slower k_{ex} compared to the mesophilic enzyme (Figure 2e). However, it should be noted that the exchange kinetics of teEIC don't follow the same temperature trend observed for tEIC (Figure 2e). This finding is likely due to the fact that the conformational dynamics in eEIC are not localized exclusively in the active site loops but extend to the β-barrel region of the enzyme (note, for example, the large R_{ex} values measured for residues Ile[426] and Gly[427] in eEIC and teEIC, Supplementary Figures S2 and S5). Coupling the β-barrel from eEIC and the loops from tEIC results in intermediate kinetics for the expanded-to-compact equilibrium in teEIC. In this regard, it is also important to notice that these results are not an artifact of the global fitting procedure, as repeating the fit using only the RD data from loop residues (Met[302], Leu[334], Ile[336], Leu[345], Ala[462], and Asp[464]) returns identical thermodynamic and kinetic parameters for the conformational equilibrium (Supplementary Table S1).

The temperature-induced unfolding of all EIC constructs was monitored by circular dichroism (CD), and the transition temperatures (T_m) were determined from the first derivatives of the unfolding curves.[35] T_m values of 54, 57, 80, and 91 °C were obtained for eEIC, teEIC, etEIC, and tEIC, respectively (Figure 2f), indicating that the 21 single-point mutations have a small but opposite effects on the temperature stability of eEIC and tEIC. In particular,

introducing the active site loops from the thermophilic construct into mesophilic EIC results in a 3 °C increase in T_m (compare T_m values measured for eEIC and teEIC), while incorporating the mesophilic active site loops into thermophilic EIC results in a 11 °C decrease in T_m (compare T_m values measured for tEIC and etEIC). As the analysis of the crystal structures of eEIC, teEIC, etEIC, and tEIC indicate that thermophilic active site loops are better packed and establish a larger number of interatomic contacts than the mesophilic ones (Figure 4e), we ascribe the small changes in melting temperature to small variations in the ΔH of folding among the investigated EIC constructs.

Temperature-Dependence of PEP Hydrolysis Catalyzed by EIC

EI catalyzes an autophosphorylation reaction in which a phosphoryl group is transferred from PEP, bound to C-terminal domain, to the His[189] residue, located on the N-terminal domain of the enzyme (Figure 1a).[13] In this catalytic process, EIC does not act as mere scaffold for PEP binding, but plays a role in the enzymatic reaction by activating PEP for catalysis. Indeed, isolated EIC was shown to catalyze hydrolysis of PEP in the absence of the EIN domain.[14] This reaction is six orders of magnitude slower than the direct phosphoryl transfer to EIN catalyzed by the full-length enzyme, but it occurs via a similar mechanism (Supplementary Figure S6a).[14] As His[189] is the only residue from the EIN domain that is in direct contact with PEP in activated, closed EI (Figure 1a), investigating the PEP hydrolysis reaction is a valuable strategy to study the role of EIC in PEP activation without interferences due to the presence of the EIN domain.

Here, we have established a method using real-time NMR to detect degradation of PEP and assay the activity of isolated EIC over a wide temperature range (5-70 °C; these limits are determined by the temperature range accessible to our NMR probe) (Supplementary Figure S6). Results of our enzymatic assay reveal that the hydrolysis of PEP by eEIC, tEIC, etEIC, and teEIC follows Michaelis-Menten kinetics (Supplementary Figure S6e). Fitted values for the

Michaelis constant (K_M) and turnover number (k_{cat}) at different temperatures (25, 30, 35, and 40 °C for eEIC; 40, 50, 60, and 70 °C for tEIC; 30, 40, 50, 60, and 70 °C for etEIC; and 25, 30, 35, and 40 °C for teEIC) are reported in Figure 5 and Supplementary Table S2 for all investigated EIC constructs.

K_M values measured for eEIC and tEIC range between 300 and 500 µM (Supplementary Table S2), which agrees well with previously reported K_M values for the EI-PEP interaction (~ 350 µM).[14] On the other hand, K_M's obtained for the PEP complexes with etEIC and teEIC (in the 400-1,300 µM range; Supplementary Table S2) are larger by a factor of ~2 compared to the ones obtained for the wild-type enzymes. This increase in K_M indicates that swapping the mesophilic with the thermophilic active site loops or, vice versa, the thermophilic with the mesophilic active site loops in EIC produces a similar effect on the EIC-PEP interaction, with both hybrid constructs binding PEP less tightly than the wild-type proteins.

Analysis of the temperature-dependence of k_{cat} indicates that eEIC is more active toward PEP hydrolysis than tEIC at low temperature (Figure 5a), suggesting that eEIC is more suited than tEIC at activating PEP for the phosphoryl-transfer reaction. Interestingly, etEIC (i.e. the hybrid construct incorporating the active site loops of eEIC and the scaffold of tEIC) has increased low-temperature turnover compared to tEIC, while teEIC (i.e. the hybrid construct incorporating the active site loops of tEIC and the scaffold of eEIC) has overall lower k_{cat} compared to eEIC (Figure 5a). These results provide evidence of a direct involvement of the active site loops in functional regulation of EIC, specifically that the amino acid sequence of the active site loops of mesophilic EIC is better suited for activation of PEP, resulting in faster turnover for PEP hydrolysis.

Conformational Disorder in Functional Regulation of EIC

The catalytic core of EIC consists of 12 residues (Leu[294], Arg[296], Arg[332], Asp[335], Lys[340], Arg[358], Met[429], Glu[431], Asn[454], Asp[455], Arg[465], and Cys[502]) that are in direct contact with PEP and the Mg^{2+} ion in the holo enzyme (Supplementary Figure S1) and are highly conserved across species. As these residues are fully conserved in the four EIC constructs investigated in this work (Supplementary Figure S1), the observation that *e*EIC, *t*EIC, *et*EIC, and *te*EIC catalyze PEP hydrolysis with different turnover numbers (Figure 5a) implies a regulatory role for residues that are not in direct contact with the substrate or metal cofactor. Interestingly, the NMR analysis of the expanded-to-compact equilibrium and the investigation of the enzyme kinetics reported above revealed that k_{ex} and k_{cat} are linearly correlated (Figure 5b), suggesting that functional regulation of EIC might be achieved via modulation of conformational dynamics within the active site. Indeed, similar linear correlations between k_{cat} and k_{ex} were previously reported for enzymatic reactions in which enzyme's conformational dynamics are rate-limiting for catalysis.[36-41] Such catalytic systems undergo conformational transitions that are essential for catalysis (either in the enzyme activation or product release process) on a time scale that is slow compared to the chemical step, causing k_{cat} to be similar to the exchange rate constant, k_{ex}. Therefore, the fact that for the PEP hydrolysis reaction investigated here k_{cat} is six orders of magnitude smaller than k_{ex} (Figure 5b) excludes the hypothesis that the expanded-to-compact conformational change regulates EIC by being the slow step in catalysis. Consistent with this observation, addition of imidazole to the reaction mixture speeds up degradation of PEP by EIC (Figure 6a). Imidazole is an analogue of the His[189] side-chain and better acceptor of the phosphoryl group than water. Therefore, we expect formation of phospho-imidazole (confirmed by ^{31}P NMR, Supplementary Figure S6c) to occur with a lower activation energy than PEP hydrolysis. The fact that the turnover rate for enzymatic degradation of PEP (i.e. the combination

of the PEP hydrolysis and imidazole phosphorylation reactions measured by our ^{1}H NMR assay (Supplementary Figure S6b) increases linearly with the concentration of imidazole (Figure 6a) indicates that k_{cat} is independent on the k_{ex} of the expanded-to-compact equilibrium and is determined by the chemical step.

Next, we have evaluated the possibility that EIC is functionally regulated by conformational disorder in the active site. Indeed, it is now widely accepted that proteins need to be considered as conformational ensembles of a range of interconverting microstates.[42, 43] As different microstates may exhibit different catalytic ability,[44, 45] the observed $k_{cat} = \sum_i p_i k_{cat}^i$, where p_i and k_{cat}^i are the fractional population and turnover number of microstate i, respectively.[25-28, 46] Inspection of the available structural models for holo EIC[14, 17, 20, 30] suggests that the EIC-PEP complex exists in two major conformations, in which the β3α3 loop adopts an open or closed conformation, respectively (Figure 6c). The side-chains of Lys^{340} and Arg^{358} coordinate the phosphate group of PEP and stabilize the transition state in the closed form of the enzyme, but do not contact the substrate in the open conformation. Therefore, we hypothesize that only microstates associated to closed EIC can efficiently catalyze PEP degradation. Consistent with this hypothesis, a single point mutation (Lys^{340}Ala) that breaks the interaction between PEP and Lys^{340} abolishes the enzymatic activity of EIC (Figure 6d), indicating that PEP-Lys^{340} contacts are crucial for catalysis and that modulating the thermodynamics of the open-to-close equilibrium within the β3α3 loop can provide functional regulation to the enzyme. It is important to notice that this conformational equilibrium differs from both (i) the open-to-close equilibrium depicted in Figure 1b for the full-length enzyme, which involves reorientation of EIN relative to EIC, and (ii) the expanded-to-compact equilibrium observed in the apo enzyme, which involves concerted dynamics of all active site loops. Also, no conformational

dynamics are detected by RD experiments on holo EIC (Figure 3e), indicating that the kinetics of the open-to-close transition of the β3α3 loop fall outside of the μs-ms timescale detectable by RD NMR and/or that the difference in chemical shift between the open and closed conformations is too small to produce detectable R_{ex} values.

As the thermodynamics of the open-to-close equilibrium of the β3α3 loop cannot be characterized by RD NMR, free-energy profiles for the conformational equilibrium in the PEP complexes with *e*EIC, *t*EIC, *et*EIC, and *te*EIC were estimated at 40 °C by Well-tempered Metadynamics (WT-MTD), a method for enhancing the sampling of collective variables of interest in constant temperature Molecular Dynamics (MD) simulations.[47] As a result of a WT-MTD run we obtain a free-energy surface as a function of selected collective variables.[48] In particular 1-μs long MTD's were performed on the four EIC-PEP complexes by using the distance between the phosphate group of PEP and the ε-ammonium group of Lys[340] ($r_{K340-PEP}$) as the collective variable. In addition, as in the available structural models for the EIC-PEP complex $r_{K340-PEP}$ ranges between 5 and 10 Å, a wall potential was applied to restrain sampling of the collective variable within the 2-15 Å range. Additional details on the construction of the starting structures, equilibration steps, and set up and analysis of the MTD runs are provided in Methods.

Inspection of the energy profiles constructed for the open-to-close conformational change by WT-MTD reveals the presence of two major energy wells corresponding to the closed ($r_{K340-PEP} < 7$ Å) and open ($r_{K340-PEP} > 7$ Å) states of the β3α3 loop (Figure 6e). While the boundaries of the closed conformation are well defined by our simulations, the energy well corresponding to the open state shows more variability and spreads over different $r_{K340-PEP}$ values in the four EIC-PEP complexes (Figure 6e), reflecting the higher disorder associated with the open conformation

(in which the side chain of Lys[340] is not structurally stabilized by the salt-bridge with PEP). It is also interesting to note that for the *e*EIC-PEP and *et*EIC-PEP complexes the energy minimum corresponding to the closed state has a lower energy than the minima associated to the open conformation, while the opposite is true for the PEP complexes with *t*EIC and *te*EIC (Figure 6e). This finding indicates that the mesophilic β3α3 loop (present in *e*EIC and *et*EIC) has a higher propensity than the thermophilic one (present in *t*EIC and *te*EIC) to sample the catalytically active, closed conformation.

The contribution of the thermodynamics of the open-to-close equilibrium to regulation of EIC was estimated by assuming that only closed conformations with $r_{K340\text{-}PEP} < 5$ Å are competent for catalysis and that the observed $k_{cat} = p_{\text{active}} k_{cat}^{\text{active}}$, where p_{active} and k_{cat}^{active} are the fractional population and turnover number of activated EIC. A threshold of $r_{K340\text{-}PEP} < 5$ Å was chosen to discriminate between active and inactive enzyme as (i) it corresponds to the $r_{K340\text{-}PEP}$ distance measured in the crystallographic model of closed, holo EIC (Figure 6c), (ii) it contains the energy minima associated with closed EIC in all our WT-MTD simulations (Figure 6e), and (iii) previous studies reported that complete loop closure is absolutely critical for efficient catalysis by TIM-barrel enzymes.[28] The populations of active and inactive enzyme were calculated by integrating the potential of mean force over $r_{K340\text{-}PEP} < 5$ Å and $r_{K340\text{-}PEP} > 5$ Å, respectively. p_{active} values of 55, 8, 68, and 6 % were obtained for *e*EIC, *t*EIC, *et*EIC, and *te*EIC, respectively (Figure 6e). The values of k_{cat}^{active} for the four EIC constructs at 40 °C and different imidazole concentrations were calculated from the corresponding values of k_{cat} using the formula: $k_{cat}^{\text{active}} = k_{cat}/p_{\text{active}}$. Interestingly, the four enzymes show very similar dependence of k_{cat}^{active} over the concentration of imidazole (Figure 6b). These results indicate that, once in the active conformation, *e*EIC, *t*EIC, *et*EIC, and *te*EIC catalyze PEP degradation

with similar turnover, and that functional regulation among these EIC constructs is achieved by modulating the thermodynamics of the open-to-closed equilibrium of holo EIC.

Conclusion

Structural biology studies reported over the past 70-years have shown that proteins can display different degrees of structural heterogeneity, going from proteins that fold into a unique rigid structure to intrinsically disordered proteins, with a continuum of possibilities in between. The degree of conformational heterogeneity can be finely tuned by external factors such as pH, ionic strength, and/or the presence of binding partners.[49] Therefore, it is no surprise that conformational disorder has been found to be implicated in regulation of important biological processes, such as signaling cascades, enzyme catalysis, and liquid-liquid phase transitions.[49] In particular, in the context of enzyme catalysis, conformational plasticity at the active site is known to control substrate specificity, with more flexible enzymes being in general more promiscuous and capable to process substrates with diverse chemical structures.[50, 51] In addition, structural heterogeneity was proposed to control turnover in several enzymatic systems.[25-28] For example, computational studies on the enzymes triosephosphate isomerase[28] and ketol-acid reductoisomerase[25] have been reported indicating that only a subset of the conformational space sampled by the enzyme-substrate complex is competent for catalysis. Based on these results, re-engineering enzymes to preferentially populate regions of the conformational space that promote reactivity was suggested as a promising strategy to increase turnover of industrially relevant reactions.[25]

In this contribution, we have provided evidence that conformational heterogeneity at the active site determines k_{cat} of the PEP degradation reaction catalyzed by EIC. Specifically, we have studied the structure/dynamics/function relationship in a mesophilic and a thermophilic variant of EIC (referred to as *e*EIC and *t*EIC, respectively) and in two additional EIC constructs

(referred to as *et*EIC and *te*EIC) engineered to have hybrid thermal stability and active site flexibility (Figure 2). By combining NMR, crystallography, computer simulations, and functional assays we have demonstrated that holo EIC exists as a mixture of states with different catalytic abilities, and that the observed turnover number is determined by the thermodynamic balance among active and inactive conformations (Figure 6). Interestingly, our data indicate that dilution of the catalytically competent states in a sea of inactive conformations is chiefly responsible for the low activity of the thermophilic enzyme at low temperature, suggesting that modulation of conformational disorder plays a role in thermal adaption of EIC. Studies are currently ongoing to investigate how temperature affects the conformational space sampled by the mesophilic and thermophilic variants of EIC, and to understand if conformational disorder is responsible for the observation that *e*EIC and *t*EIC catalyze PEP hydrolysis with similar turnover at their optimal PTS temperature (Figure 5a). Our data also indicate that conformational dynamics at the active site of holo EIC are important for activating PEP for catalysis, and that altering the conformational space sampled by the C-terminal domain is a possible avenue to achieve inhibition of the full-length EI.

EIC displays a $(\beta/\alpha)_8$-barrel fold (also known as TIM barrel fold) that, due to its ability to carry very efficient and versatile active sites, has been selected by nature to catalyze a large number of reactions.[52] All known $(\beta/\alpha)_8$-barrel enzymes have the active site residues localized at the C-terminal ends of the β-strands and within the βα-loops (which connect the C-terminal end of β-strand *n* to the N-terminal end of α-helix *n*), while residues located within the core, at the N-terminal ends of the β-strands, and within the αβ-loops (which connect the C-terminal end of α-helix *n-1* to the N-terminal end of β-strand *n*) are responsible to maintain protein stability. The presence of a catalytic face and a stability face makes the $(\beta/\alpha)_8$-barrel fold highly evolvable, and

the most common enzyme fold in nature.[52] As a consequence, $(\beta/\alpha)_8$-barrel enzymes were the subject of several protein engineering efforts aimed at designing new or improved functionalities by reshuffling of individual $(\beta\alpha)$-units or de novo design.[52] Here, we have engineered thermophilic/mesophilic constructs of EIC by merging the stability face of one enzyme with the catalytic face of the other enzyme (Figure 2). The newly designed proteins have hybrid structural properties, with the construct incorporating the scaffold of *t*EIC and the active loops of *e*EIC (*et*EIC) being highly flexible and thermostable, and the opposite hybrid (*te*EIC) displaying high rigidity and low thermal stability (Figure 2). Interestingly, functional characterization of our hybrid constructs revealed that *et*EIC catalyze PEP hydrolysis with ~20 and ~3 times higher k_{cat} than *t*EIC at 40 and 70 °C, respectively (Figure 5), suggesting that engineering hybrid mesophilic/thermophilic constructs might be a general strategy to obtain highly stable $(\beta/\alpha)_8$-barrel enzymes with significant low-temperature activity.

Methods

Protein Expression and Purification

*e*EIC and *te*EIC were cloned into a pET21a vector and were expressed and purified using protocols described previously for *e*EIC.[14] *t*EIC and *et*EIC were cloned into a vector incorporating a His-tagged EIN solubility tag at the N-terminus.[53] Expression and purification of *t*EIC and *et*EIC was performed as previously described for *t*EIC.[54] U-[^{2}H,^{15}N]/Ile(d_1)-^{13}CH$_3$/Val, Leu-(^{13}CH$_3$/^{12}C^{2}H$_3$)-labeled (for RD experiments) and U-[^{2}H,^{15}N,^{13}C]/Ile(d_1)-^{13}CH$_3$/Val, Leu-(^{13}CH$_3$/^{12}C^{2}H$_3$)-labeled (for NMR resonance assignment experiments) EIC were prepared following standard protocols for specific isotopic labeling of the methyl groups in Ile, Leu, and Val side chains.[55]

Thermal Unfolding and Circular Dichroism

The progress of temperature-induced unfolding was monitored at 222 nm in a 0.1 cm cuvette using a Jasco J-710 spectropolarimeter. The midpoint of the thermal unfolding transition (T_m) was calculated as the maximal value of the first derivative of the CD signal at 222 nm with respect to temperature. Protein concentration was ~0.4 mg/ml.

Analytical Ultracentrifugation

Sedimentation velocity experiments were conducted at 50,000 rpm and 20 ºC on a Beckman Coulter ProteomeLab XLI analytical ultracentrifuge following standard protocols.[56] Samples of the EIC were studied at concentrations ranging from ~10 to 40 µM in 100 mM NaCl, 20 mM Tris (pH 7.4), 2 mM DTT, and 4 mM $MgCl_2$. Samples were loaded in standard 12 mm, 2-channel centerpiece cells, and data collected using both the absorbance (280 nm) and Rayleigh interference (655 nm) optical detection systems. Time-corrected[57] sedimentation data were analyzed in SEDFIT 16.01c[58] in terms of a continuous c(s) distribution of sedimenting species with a resolution of 0.05 S and a maximum entropy regularization confidence level of 0.68. The solution density, solution viscosity and protein partial specific volume were calculated in SEDNTERP,[59] and sedimentation coefficients s were corrected to standard conditions $s_{20,w}$.

X-ray Crystallography

Crystals of *et*EIC grew in well C12 of the Morpheus crystallization screen (Molecular Dimensions) using a 96-well hanging-drop vapor diffusion plate setup with 100 microliter reservoirs, 210 nanoliter drop volumes, 1:1 ratio of protein to reservoir solutions and incubated at 25 °C. The Morpheus C12 well solution consisted of 0.09 M NPS salts (Sodium nitrate, Sodium phosphate dibasic, Ammonium sulfate), 0.1 M Buffer System 3 (Tris, Bicine pH 8.5) and 50 % of Precipitant Mix 4 (25 % MPD, 25 % PEG 1000, 25 % PEG 3350). *et*EIC crystals did not require cryoprotection prior to flash-freezing. Crystals of *te*EIC grew in well C10 of the PACT

premier Eco screen (Molecular Dimensions) using the same hanging drop plate setup as described above. The PACT C10 well solution consisted of 0.2 M Magnesium chloride hexahydrate, 0.1 M HEPES pH 7.0 and 20 % w/v PEG 6000. Prior to flash-freezing in liquid nitrogen, *te*EIC crystals were cryoprotected using mother liquor plus 15 % glycerol. Crystals of *e*EIC grew in well G8 of the XP Screen (MiTeGen, LLC) using a 96-well sitting-drop vapor diffusion plate setup with 50 microliter reservoirs, 210 nanoliter drop volume with a 1:1 ratio of protein to reservoir solutions and incubated at 18 °C. Well G8 of the XP Screen consisted of 1.0 M Lithium sulfate, 0.5 M Ammonium sulfate, 0.1 M tri-Sodium citrate (pH 5.6) and 1.0 mM TEW. Prior to flash-freezing in liquid nitrogen, *te*EIC crystals were cryoprotected using mother liquor plus 15 % glycerol.

X-ray diffraction data for *et*EIC and *te*EIC crystals were collected at the Advanced Light Source (Beamline 4.2.2), processed with XDS[60] and scaled with AIMLESS within the CCP4 software suite.[61] X-ray diffraction data for *e*EIC crystals were collected at the Advanced Photon Source (Beamline 23ID-D), processed with iMOSFLM[62] and scaled with AIMLESS within the CCP4 software suite. The crystal structures of *et*EIC, *te*EIC, and *e*EIC were solved by molecular replacement with the *t*EIC model (PBD ID code 2XZ7)[30] using the program Phaser.[63] Model building and refinement were carried out using PHENIX autobuild[64] and phenix.refine.[65] Coot[66] was used for graphical map inspection and manual refinement of coordinates. MolProbity was used for structural validation.[67] The atomic coordinates and structure factors have been deposited in the RCSB Protein Data Bank, with the following PDB IDs, 6VBJ (*et*EIC), 6V9K (*te*EIC), and 6VU0 (*e*EIC). Statistics on the data collection and the final models are given in Supplementary Table S3.

Molecular Dynamics Simulations

Molecular Dynamics simulations were performed using Gromacs 2018.3[68] patched with PLUMED2.5.[69] We used Charmm36 force field for the four enzyme variants. The CGENFF was used for generating the Charmm36 compatible field for the PEP ligands. The crystal structure of the tEIC-PEP complex[30] and the structural model of the tEIC-PEP complex[17] (see caption to Figure 1 for more details on the modelling) were used as the starting structure for the thermophilic and mesophilic enzyme, respectively. Starting coordinates for the PEP complexes with etEIC and teEIC were obtained by mutating the coordinate files of the tEIC-PEP and eEIC-PEP complexes, respectively, in CHARMM-GUI. All PEP bound enzyme structures were solvated with TIP3P water molecules in a rhombic dodecahedral box resulting in ~90,000 number of molecules. Appropriate number of Na^+ and Cl^- were added to neutralize the system and set the bulk salt concentration at 130 mM.

The long-range electrostatic interactions were evaluated by using the particle mesh Ewald (PME) method. All the systems were first minimized for 10,000 steps via a steepest descent method. Then system was then gradually heated up to 313 K in 200 ps subject to positional restraints. The restraint force constants were gradually decreased from 10^3 to 0 kJ nm^{-2} mol^{-1}. For the production simulations we have kept harmonic restraints on PEP with force constant 40 kJ nm^{-2} mol^{-1} to restraint the ligand in fixed orientation enabling cleaner comparisons between free energy profiles of enzymes with modified sequences.

The rest of the degrees of freedom were fully unrestrained to evolve at 313 K with Nose-Hoover thermostat and a relaxation time of 0.5 ps. The pressure was maintained at 1 bar using the Parrinello-Rahman coupling scheme. The LINCS algorithm was employed to constrain the H-atoms bonds. The time step of the production simulations was 2.0 fs. The cutoff of the non-bonded interactions was set to 10 Å.

After 50 ns equilibration in NPT ensemble well-tempered metadynamics simulations of 1 μs duration were run with the distance between the center of mass of the phosphate group of PEP and the center of mass of the ε-ammonium group of Lys[340] as the collective variable ($r_{K340-PEP}$). The parameters of 0.8 for hill height and 8 for the visa factor were used for carrying out well-tempered metadynamics runs. Plumed2.5 plugin was employed for monitoring and processing the results. To limit the comparison between near native states we have further imposed wall potentials at 1.5 nm distance with harmonic potentials and a force constant of 1,500 kJ nm^{-2} mol^{-1}.

Free energy difference between the active and inactive EIC conformations was computed by integrating potential of mean force over basins $r_{K340-PEP}$ = 2-5 Å and $r_{K340-PEP}$ = 5-15 Å, respectively. The differences are robust with respect to the choice of basins as long as they include all global and local minima. We note here that due to disordered and high entropy nature of the open state, the free energy difference is a fluctuating quantity the convergence of which has been assessed by obtaining stable mean of the free energy difference between open and closed forms which has been collected over the last 100 ns of the simulation. The Potential of Mean profiles report the averaged free energy profiles over last 100 ns showing the stable mean free energy differences between open and closed forms. See Supplementary Figure S6 for more details on simulated structures of enzymes and convergence assessment.

NMR Spectroscopy

NMR samples were prepared in 20 mM Tris buffer, pH 7.4, 100 mM NaCl, 4 mM MgCl$_2$, 1 mM ethylenediamine tetraacetic acid (EDTA), 2 mM dithiothreitol (DTT), and 90% H$_2$O/10% D$_2$O. The protein concentration (in subunits) was ~1.0 mM for all protein detected experiments (unless stated otherwise).

NMR spectra were recorded Bruker 800, 700, and 600 MHz spectrometers equipped with a z-shielded gradient triple resonance cryoprobes. Spectra were processed using NMRPipe[70] and analyzed using the program SPARKY. ^{1}H-^{15}N TROSY (transverse relaxation optimized)[71] and methyl-TROSY[72] experiments have been acquired using previously described pulse schemes. Resonance assignments of the ^{1}H-^{15}N TROSY and methyl-TROSY spectra of *e*EIC and *t*EIC were achieved by transferring the assignments obtained previously.[14, 54, 73, 74] Sequential ^{1}H/^{15}N/^{13}C backbone assignment of the *et*EIC and *te*EIC was carried out using TROSY versions [75, 76] of conventional 3D triple resonance correlation experiments (HNCO, HNCA, HNCACB, HN(CO)CA, and HN(CO)CACB).[77] Assignment of the ^{1}H-^{13}C$_{methyl}$ correlations of *et*EIC and *te*EIC was performed using out-and-back experiments.[73] Only non-stereospecific assignments are obtained for the methyl groups of the Leu and Val residues of the EIC constructs analyzed here. Therefore, methyl-TROSY correlations for such resonances were arbitrarily assigned as δ_A and δ_B (for Leu), and γ_A and γ_B (for Val). Chemical shifts assigned for *et*EIC and *te*EIC have been deposited in the BioMagResBank (accession no. 28080 and 28079, respectively).

Weighted combined ^{1}H/^{15}N ($\Delta_{H/N}$) and ^{1}H/^{13}C ($\Delta_{H/C}$) chemical shift perturbations resulting from the addition of PEP were calculated using the following equation:[78] $\Delta_{H/N,C} = ((\Delta\delta_H W_H)^2 + (\Delta\delta_{N,C} W_{N,C})^2)^{1/2}$, where W_H, W_N, and W_C are weighing factors for the ^{1}H, ^{15}N, and ^{13}C shifts, respectively ($W_H = |\gamma_H/\gamma_H| = 1$, $W_N = |\gamma_N/\gamma_H| = 0.101$, $W_C = |\gamma_C/\gamma_H| = 0.251$), $\Delta\delta_H$, $\Delta\delta_N$ and $\Delta\delta_C$ are the ^{1}H, ^{15}N and ^{13}C chemical shift differences in ppm, respectively, between free and bound states, and γ_H, γ_N and γ_C are the ^{1}H, ^{15}N and ^{13}C gyromagnetic ratios, respectively.

^{15}N and ^{13}C$_{methyl}$ RD experiments were conducted at multiple temperatures (see Results and Discussion) using a pulse sequence that measures the exchange contribution for the TROSY component of the ^{15}N magnetization,[79] or a pulse scheme for ^{13}C singe quantum CPMG (Carr-

Purcell-Meinboom-Gill) relaxation dispersion described by Kay and co-workers.[80] Off-resonance effects and pulse imperfections were minimized using a four-pulse phase scheme.[81] Experiments were performed at 600 and 800 MHz with a fixed relaxation delay but a changing number of refocusing pulses to achieve different effective CPMG fields.[82] The transverse relaxation periods were set to 60 and 30 ms for the ^{15}N and ^{13}C$_{methyl}$ experiments, respectively. The resulting relaxation dispersion curves were fit to a two-state exchange model using the Carver–Richards equation:[83]

$$R_{2,eff}\left(\frac{1}{2\delta}\right) = \text{Re}(\lambda_1) - \frac{1}{4n\delta}\ln(Q) \tag{1}$$

with

$$\lambda_1 = R_{2,\text{int}} + \frac{1}{2}\left(k_{ex} - \frac{1}{2\delta}\cosh^{-1}\left(D_+\cosh\eta_+ - D_-\cos\eta_-\right)\right) \tag{2}$$

$$D_\pm = \frac{1}{2}\left(\frac{\Psi + 2\Delta\omega_{N,C}^2}{\sqrt{\Psi^2 + \zeta^2}} \pm 1\right) \tag{3}$$

$$\eta_\pm = \sqrt{2}\delta\sqrt{\sqrt{\Psi^2 + \zeta^2} \pm \Psi} \tag{4}$$

$$\Psi = \left((p_a - p_b)k_{ex}\right)^2 - \Delta\omega_{N,C}^2 + 4p_a p_b k_{ex}^2 \tag{5}$$

$$\zeta = -2\Delta\omega_{N,C}\left((p_a - p_b)k_{ex}\right) \tag{6}$$

$$Q = \text{Re}\left(1 - m_D^2 + m_D m_Z - m_Z^2 + (1/2)(m_D + m_Z)\sqrt{p_b/p_a}\right) \tag{7}$$

$$m_D = \frac{ik_{ex}\sqrt{p_a p_b}}{d_+ z_+}\left(z_+ + 2\Delta\omega_{N,C}\frac{\sin(z_+\delta)}{\sin((d_+ + z_+)\delta)}\right) \tag{8}$$

$$\tag{9}$$

$$d_\pm = \Delta\omega_C \pm ik_{ex}, \quad z_\pm = -\Delta\omega_{N,C} \pm ik_{ex} \tag{10}$$

Here, $R_{2,eff}$ is the measured ^{15}N or ^{13}C transverse relaxation rate; $R_{2,int}$ is the ^{15}N or ^{13}C transverse relaxation rate in the absence of conformational exchange; n is the number of 180° refocusing pulses in the CPMG train; 2δ is the spacing between the centers of successive 180° pulses; k_{ex} is the exchange rate constant; p_a and p_b are the fractional populations of the major and minor species, respectively; $\Delta\omega_{N,C}$ is the nitrogen or carbon chemical shift differences between the two conformational states in rad s[-1].

The optimization was run using p_b and k_{ex} as global parameters. p_a was calculated using the formula $p_a = 1 - p_b$. $R_{2,int}$ and $\Delta\omega_{N,C}$ were fit as a peak specific parameters (i.e. each RD curve will have its own $R_{2,int}$ and $\Delta\omega_{N,C}$).

In order to reduce the number of adjustable parameters in the fit, the RD curves measured for each EIC construct at four different temperatures were fit simultaneously. In this global fitting procedure, the residue specific $\Delta\omega_{N,C}$'s were treated as temperature independent parameters, while k_{ex} and p_b at each temperature were calculated from the activation ($\Delta^{\ddagger}G$) and standard (ΔG) free energy of the conformational equilibrium using the formulas:

$$k_{ex} = \frac{k_B T}{h} e^{-\frac{\Delta^{\ddagger}G}{RT}} \tag{11}$$

and

$$p_b = \frac{K_{eq}}{1 + K_{eq}} \tag{12}$$

with

$$K_{eq} = e^{-\frac{\Delta G}{RT}} \tag{13}$$

where T is the temperature in Kelvin, R is universal gas constant, and K_{eq} is the equilibrium constant for the conformational change.

Enzymatic Assays

Enzymatic activity of eEIC, tEIC, etEIC, and teEIC for degradation of PEP to inorganic phosphate, phospho-imidazole, and pyruvate was assayed by real-time ^{1}H NMR by monitoring the disappearance of the alkene proton signals (Supplementary Figure S6). The reaction mixtures were prepared in 20 mM Tris buffer, pH 7.4, 100 mM NaCl, 4 mM $MgCl_2$, 1 mM EDTA, 2 mM DTT, 1mM Trimethylsilylpropanoic acid (TSP), and 99.99% D_2O. Enzyme concentration was 50 μM for eEIC, tEIC and etEIC, and 80 μM for teEIC. Enzymatic assays were run in duplicate, and the integrals of the signals measured for the alkene protons were converted to mM units by reference to the internal standard TSP. Integration of the NMR signals was performed using the software MNova (https://mestrelab.com/download/mnova/). Initial velocities were determined from the linear portion of the progress curves and were fit using the Michaelis-Menten equation (Supplementary Figure S6).

References

1.	Clore, G.M. & Venditti, V. Structure, dynamics and biophysics of the cytoplasmic protein-protein complexes of the bacterial phosphoenolpyruvate: sugar phosphotransferase system. *Trends Biochem Sci* **38**, 515-530 (2013).

2.	Deutscher, J. *et al.* The bacterial phosphoenolpyruvate:carbohydrate phosphotransferase system: regulation by protein phosphorylation and phosphorylation-dependent protein-protein interactions. *Microbiol Mol Biol Rev* **78**, 231-256 (2014).

3.	Postma, P.W., Lengeler, J.W. & Jacobson, G.R. Phosphoenolpyruvate:Carbohydrate Phosphotransferase Systems, in *Escherichia coli and Salmonella: Cellular and Molecular Biology*. (eds. F.C. Neidhardt, E.C. Lin & R. Curtiss) 1149-1174 (1996).

4.	Doucette, C.D., Schwab, D.J., Wingreen, N.S. & Rabinowitz, J.D. alpha-Ketoglutarate coordinates carbon and nitrogen utilization via enzyme I inhibition. *Nat Chem Biol* **7**, 894-901 (2011).

5.	Venditti, V., Ghirlando, R. & Clore, G.M. Structural basis for enzyme I inhibition by alpha-ketoglutarate. *ACS Chem Biol* **8**, 1232-1240 (2013).

6.	Hogema, B.M. *et al.* Inducer exclusion in Escherichia coli by non-PTS substrates: the role of the PEP to pyruvate ratio in determining the phosphorylation state of enzyme IIAGlc. *Mol Microbiol* **30**, 487-498 (1998).

7.	Huang, K.J. *et al.* Xanthone derivatives could be potential antibiotics: virtual screening for the inhibitors of enzyme I of bacterial phosphoenolpyruvate-dependent phosphotransferase system. *J Antibiot* **66**, 453-458 (2013).

8.	Kok, M., Bron, G., Erni, B. & Mukhija, S. Effect of enzyme I of the bacterial phosphoenolpyruvate : sugar phosphotransferase system (PTS) on virulence in a murine model. *Microbiology* **149**, 2645-2652 (2003).

9.	Edelstein, P.H., Edelstein, M.A., Higa, F. & Falkow, S. Discovery of virulence genes of Legionella pneumophila by using signature tagged mutagenesis in a guinea pig pneumonia model. *Proc Natl Acad Sci U S A* **96**, 8190-8195 (1999).

10.	Hava, D.L. & Camilli, A. Large-scale identification of serotype 4 Streptococcus pneumoniae virulence factors. *Mol Microbiol* **45**, 1389-1406 (2002).

11.	Jones, A.L., Knoll, K.M. & Rubens, C.E. Identification of Streptococcus agalactiae virulence genes in the neonatal rat sepsis model using signature-tagged mutagenesis. *Mol Microbiol* **37**, 1444-1455 (2000).

12. Lau, G.W. *et al.* A functional genomic analysis of type 3 Streptococcus pneumoniae virulence. *Mol Microbiol* **40**, 555-571 (2001).

13. Chauvin, F., Brand, L. & Roseman, S. Enzyme I: the first protein and potential regulator of the bacterial phosphoenolpyruvate: glycose phosphotransferase system. *Res Microbiol* **147**, 471-479 (1996).

14. Venditti, V. & Clore, G.M. Conformational selection and substrate binding regulate the monomer/dimer equilibrium of the C-terminal domain of Escherichia coli enzyme I. *J Biol Chem* **287**, 26989-26998 (2012).

15. Nguyen, T.T., Ghirlando, R. & Venditti, V. The oligomerization state of bacterial enzyme I (EI) determines EI's allosteric stimulation or competitive inhibition by alpha-ketoglutarate. *J Biol Chem* **293**, 2631-2639 (2018).

16. Patel, H.V., Vyas, K.A., Savtchenko, R. & Roseman, S. The monomer/dimer transition of enzyme I of the Escherichia coli phosphotransferase system. *J Biol Chem* **281**, 17570-17578 (2006).

17. Venditti, V., Tugarinov, V., Schwieters, C.D., Grishaev, A. & Clore, G.M. Large interdomain rearrangement triggered by suppression of micro- to millisecond dynamics in bacterial Enzyme I. *Nat Commun* **6**, 5960 (2015).

18. Suh, J.Y., Cai, M. & Clore, G.M. Impact of phosphorylation on structure and thermodynamics of the interaction between the N-terminal domain of enzyme I and the histidine phosphocarrier protein of the bacterial phosphotransferase system. *J Biol Chem* **283**, 18980-18989 (2008).

19. Schwieters, C.D. *et al.* 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. *J Am Chem Soc* **132**, 13026-13045 (2010).

20. Teplyakov, A. *et al.* Structure of phosphorylated enzyme I, the phosphoenolpyruvate:sugar phosphotransferase system sugar translocation signal protein. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 16218-16223 (2006).

21. Venditti, V., Schwieters, C.D., Grishaev, A. & Clore, G.M. Dynamic equilibrium between closed and partially closed states of the bacterial Enzyme I unveiled by solution NMR and X-ray scattering. *Proc Natl Acad Sci U S A* **112**, 11565-11570 (2015).

22. Palombo, M. *et al.* The relationship between folding and activity in UreG, an intrinsically disordered enzyme. *Sci Rep* **7**, 5977 (2017).

23. Larion, M., Miller, B. & Bruschweiler, R. Conformational heterogeneity and intrinsic disorder in enzyme regulation: Glucokinase as a case study. *Intrinsically Disord Proteins* **3**, e1011008 (2015).

24. DeForte, S. & Uversky, V.N. Not an exception to the rule: the functional significance of intrinsically disordered protein regions in enzymes. *Mol Biosyst* **13**, 463-469 (2017).

25. Bonk, B.M., Weis, J.W. & Tidor, B. Machine Learning Identifies Chemical Characteristics That Promote Enzyme Catalysis. *J Am Chem Soc* **141**, 4108-4118 (2019).

26. Ghose, R. Nature of the Pre-Chemistry Ensemble in Mitogen-Activated Protein Kinases. *J Mol Biol* **431**, 145-157 (2019).

27. Glowacki, D.R., Harvey, J.N. & Mulholland, A.J. Taking Ockham's razor to enzyme dynamics and catalysis. *Nat Chem* **4**, 169-176 (2012).

28. Liao, Q. *et al.* Loop Motion in Triosephosphate Isomerase Is Not a Simple Open and Shut Case. *J Am Chem Soc* **140**, 15889-15903 (2018).

29. Evangelidis, T. *et al.* Automated NMR resonance assignments and structure determination using a minimal set of 4D spectra. *Nat Commun* **9**, 384 (2018).

30. Navdaeva, V. *et al.* Phosphoenolpyruvate: sugar phosphotransferase system from the hyperthermophilic Thermoanaerobacter tengcongensis. *Biochemistry* **50**, 1184-1193 (2011).

31. Oberholzer, A.E. *et al.* Crystal structure of the phosphoenolpyruvate-binding enzyme I-domain from the Thermoanaerobacter tengcongensis PEP: sugar phosphotransferase system (PTS). *J Mol Biol* **346**, 521-532 (2005).

32. Mittermaier, A. & Kay, L.E. New tools provide new insights in NMR studies of protein dynamics. *Science* **312**, 224-228 (2006).

33. Purslow, J.A. *et al.* Active Site Breathing of Human Alkbh5 Revealed by Solution NMR and Accelerated Molecular Dynamics. *Biophys J* (2018).

34. Nguyen, V. *et al.* Evolutionary drivers of thermoadaptation in enzyme catalysis. *Science* **355**, 289-294 (2017).

35. Greenfield, N.J. Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc* **1**, 2876-2890 (2006).

36. Eisenmesser, E.Z. *et al.* Intrinsic dynamics of an enzyme underlies catalysis. *Nature* **438**, 117-121 (2005).

37. Fraser, J.S. *et al.* Hidden alternative structures of proline isomerase essential for catalysis. *Nature* **462**, 669-673 (2009).

38. Henzler-Wildman, K.A. *et al.* A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* **450**, 913-916 (2007).

39. Henzler-Wildman, K.A. *et al.* Intrinsic motions along an enzymatic reaction trajectory. *Nature* **450**, 838-844 (2007).

40. Rozovsky, S. & McDermott, A.E. The time scale of the catalytic loop motion in triosephosphate isomerase. *J Mol Biol* **310**, 259-270 (2001).

41. Whittier, S.K., Hengge, A.C. & Loria, J.P. Conformational motions regulate phosphoryl transfer in related protein tyrosine phosphatases. *Science* **341**, 899-903 (2013).

42. Frauenfelder, H., Sligar, S.G. & Wolynes, P.G. The energy landscapes and motions of proteins. *Science* **254**, 1598-1603 (1991).

43. Tokuriki, N. & Tawfik, D.S. Protein dynamism and evolvability. *Science* **324**, 203-207 (2009).

44. Lu, H.P., Xun, L. & Xie, X.S. Single-molecule enzymatic dynamics. *Science* **282**, 1877-1882 (1998).

45. Smiley, R.D. & Hammes, G.G. Single molecule studies of enzyme mechanisms. *Chem Rev* **106**, 3080-3094 (2006).

46. Elias, M., Wieczorek, G., Rosenne, S. & Tawfik, D.S. The universality of enzymatic rate-temperature dependency. *Trends Biochem Sci* **39**, 1-7 (2014).

47. Barducci, A., Bussi, G. & Parrinello, M. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Phys Rev Lett* **100**, 020603 (2008).

48. Laio, A. & Parrinello, M. Escaping free-energy minima. *Proc Natl Acad Sci U S A* **99**, 12562-12566 (2002).

49. Uversky, V.N. Protein intrinsic disorder and structure-function continuum. *Prog Mol Biol Transl Sci* **166**, 1-17 (2019).

50. Campbell, E. *et al.* The role of protein dynamics in the evolution of new enzyme function. *Nat Chem Biol* **12**, 944-950 (2016).

51. Jacquet, P. *et al.* Rational engineering of a native hyperthermostable lactonase into a broad spectrum phosphotriesterase. *Sci Rep* **7**, 16745 (2017).

52. Sterner, R. & Hocker, B. Catalytic versatility, stability, and evolution of the (betaalpha)8-barrel enzyme fold. *Chem Rev* **105**, 4038-4055 (2005).

53. Khatiwada, B., Purslow, J.A., Underbakke, E.S. & Venditti, V. N-terminal fusion of the N-terminal domain of bacterial enzyme I facilitates recombinant expression and purification of the human RNA demethylases FTO and Alkbh5. *Protein Expr Purif* **167**, 105540 (2020).

54.   Dotas, R.R. & Venditti, V. (1)H, (15)N, (13)C backbone resonance assignment of the C-terminal domain of enzyme I from Thermoanaerobacter tengcongensis. *Biomol NMR Assign* (2017).

55.   Tugarinov, V., Kanelis, V. & Kay, L.E. Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nat Protoc* **1**, 749-754 (2006).

56.   Zhao, H., Brautigam, C.A., Ghirlando, R. & Schuck, P. Overview of current methods in sedimentation velocity and sedimentation equilibrium analytical ultracentrifugation. *Curr Protoc Protein Sci* **Chapter 20**, Unit20 12 (2013).

57.   Ghirlando, R. *et al.* Improving the thermal, radial, and temporal accuracy of the analytical ultracentrifuge through external references. *Anal Biochem* **440**, 81-95 (2013).

58.   Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* **78**, 1606-1619 (2000).

59.   Cole, J.L., Lary, J.W., T, P.M. & Laue, T.M. Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. *Methods Cell Biol* **84**, 143-179 (2008).

60.   Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132 (2010).

61.   Evans, P.R. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D Biol Crystallogr* **67**, 282-292 (2011).

62.   Battye, T.G., Kontogiannis, L., Johnson, O., Powell, H.R. & Leslie, A.G. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* **67**, 271-281 (2011).

63.   McCoy, A.J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007).

64.   Terwilliger, T.C. *et al.* Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* **64**, 61-69 (2008).

65.   Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75**, 861-877 (2019).

66.   Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).

67.   Chen, V.B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21 (2010).

68. Abraham, M.J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19-25 (2015).

69. Tribello, G.A., Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G. PLUMED 2: New feathers for an old bird. *Computer Physics Communications* **185**, 604-613 (2014).

70. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293 (1995).

71. Pervushin, K., Riek, R., Wider, G. & Wuthrich, K. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci U S A* **94**, 12366-12371 (1997).

72. Tugarinov, V., Hwang, P.M., Ollerenshaw, J.E. & Kay, L.E. Cross-correlated relaxation enhanced 1H[bond]13C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J Am Chem Soc* **125**, 10420-10428 (2003).

73. Tugarinov, V., Venditti, V. & Marius Clore, G. A NMR experiment for simultaneous correlations of valine and leucine/isoleucine methyls with carbonyl chemical shifts in proteins. *J Biomol NMR* **58**, 1-8 (2014).

74. Dotas, R.R. & Venditti, V. Resonance assignment of the 128 kDa enzyme I dimer from Thermoanaerobacter tengcongensis. *Biomol NMR Assign* (2019).

75. Pervushin, K., Riek, R., Wider, G. & Wuthrich, K. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 12366-12371 (1997).

76. Tugarinov, V., Muhandiram, R., Ayed, A. & Kay, L.E. Four-dimensional NMR spectroscopy of a 723-residue protein: chemical shift assignments and secondary structure of malate synthase g. *J Am Chem Soc* **124**, 10025-10035 (2002).

77. Clore, G.M. & Gronenborn, A.M. Determining the structures of large proteins and protein complexes by NMR. *Trends Biotechnol.* **16**, 22-34 (1998).

78. Mulder, F.A., Schipper, D., Bott, R. & Boelens, R. Altered flexibility in the substrate-binding site of related native and engineered high-alkaline Bacillus subtilisins. *J Mol Biol* **292**, 111-123 (1999).

79. Loria, J.P., Rance, M. & Palmer, A.G., 3rd A TROSY CPMG sequence for characterizing chemical exchange in large proteins. *J Biomol NMR* **15**, 151-155 (1999).

80. Lundstrom, P., Vallurupalli, P., Religa, T.L., Dahlquist, F.W. & Kay, L.E. A single-quantum methyl 13C-relaxation dispersion experiment with improved sensitivity. *J Biomol NMR* **38**, 79-88 (2007).

81. Yip, G.N. & Zuiderweg, E.R. A phase cycle scheme that significantly suppresses offset-dependent artifacts in the R2-CPMG 15N relaxation experiment. *J Magn Reson* **171**, 25-36 (2004).

82. Mulder, F.A., Skrynnikov, N.R., Hon, B., Dahlquist, F.W. & Kay, L.E. Measurement of slow (micros-ms) time scale dynamics in protein side chains by (15)N relaxation dispersion NMR spectroscopy: application to Asn and Gln residues in a cavity mutant of T4 lysozyme. *J Am Chem Soc* **123**, 967-975 (2001).

83. Carver, J.P. & Richards, R.E. A general two-site solution for the chemical exchange produced dependence of T_2 upon the Carr-Purcell pulse separation. *J Magn Reson* **6**, 89-105 (1972).

84. Purslow, J.A., Khatiwada, B., Bayro, M.J. & Venditti, V. NMR Methods for Structural Characterization of Protein-Protein Complexes. *Front. Mol. Biosci.* **7**, 9 (2020).

Figures
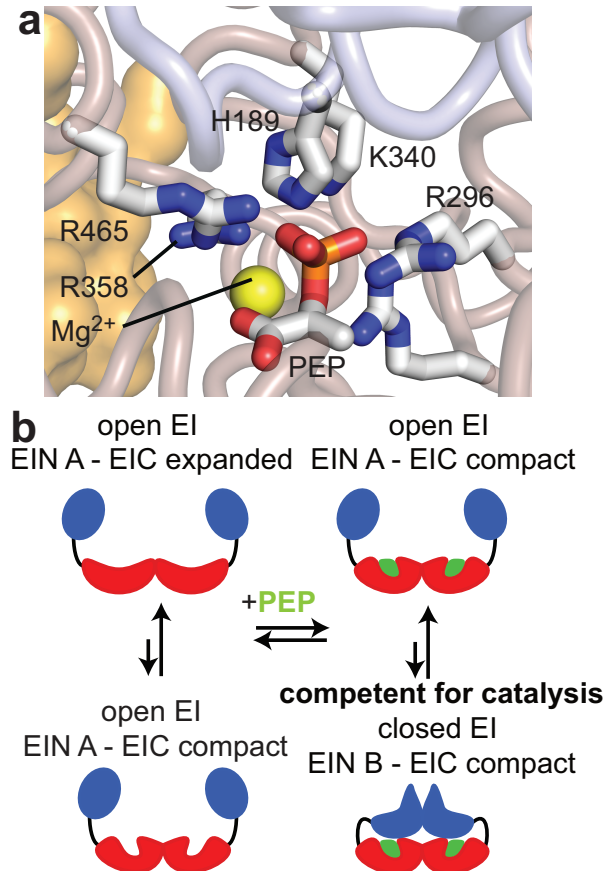
Figure 1. EI conformational rearrangements during catalysis. (**a**) Structural model of PEP bound in the active site of closed EI. Modelling was performed based on the crystal structure of the *t*EIC-PEP complex[30] and on the X-ray structure of a phosphorylated *e*EI intermediate captured immediately after the autophosphorylation reaction.[20] Details on the modelling are provided elsewhere.[17] The positioning of PEP and of the His[189] side-chain allows in-line phosphoryl transfer from PEP bound to the EIC domain to the EIN domain. PEP, the side-chain of His[189], and side-chains of residues from the EIC domain that interact with the phosphate group of PEP are shown as solid sticks. The magnesium ion is displayed as a yellow sphere. The backbone of the EIN and EIC domains are shown as transparent blue and salmon tubes, respectively. The second subunit is shown as a transparent orange surface. (**b**) Schematic view of EI conformational equilibria. The EIN and EIC domains are colored blue and red, respectively. The PEP molecule is colored green.
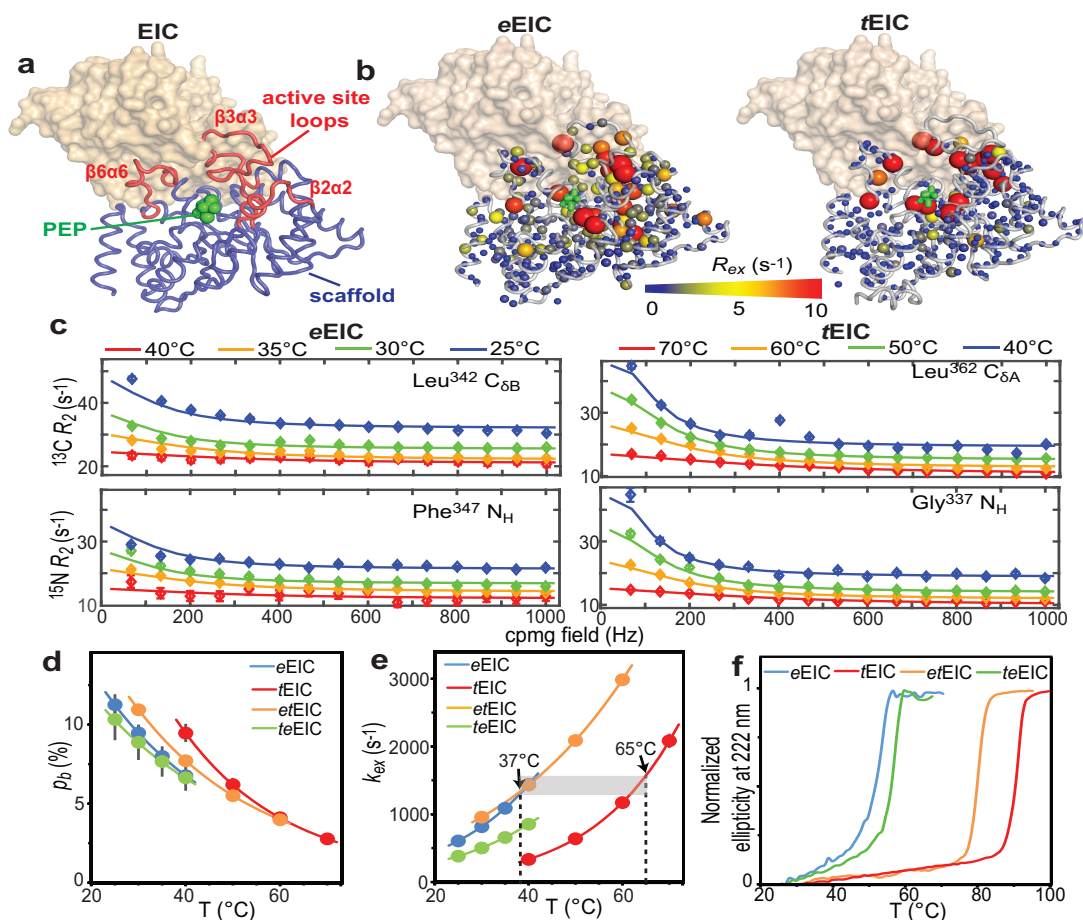
Figure 2. Design of thermophilic/mesophilic EIC hybrids. (**a**) Crystal structure of tEIC-PEP complex. Active site loops are shown as red tubes. Protein scaffold is shown as blue tubes. PEP is displayed as green spheres. The second subunit is shown as a transparent orange surface. (**b**) 800 MHz exchange contribution to the transverse relaxation rates (R_{ex}) measured for eEIC and tEIC and 25 and 40 °C, respectively, are plotted on the EIC structure as spheres with the relationship between size and color of each sphere and R_{ex} value depicted by the color bar. R_{ex} values are displayed on one subunit of the EIC dimer. The second subunit is shown as a transparent surface. (**c**) Examples of typical 800 MHz relaxation dispersion data measured for eEIC (left panel) and tEIC (right panel) at four different temperatures. Data are shown for the Leu[342]-δ methyl group (top) and Phe[347] amide group (bottom) of eEIC, and for the Leu[362]-δ methyl group (top) and Gly[337] amide group (bottom) of tEIC. Experimental data are represented by filled-in circles. Best-fit curves are shown as solid lines. Different colors correspond to different temperatures (blue, green, orange, and red correspond to 25, 30, 35, and 40 °C, respectively, for eEIC, and to 40, 50, 60, and 70 °C, respectively, for tEIC). Similar plots for all the analyzed NMR signals are shown in Supplementary Figures S2-S5. Fitted p_b (**d**) and k_{ex} (**e**) for the expanded-to-compact equilibrium in apo eEIC (blue), tEIC (red), etEIC (orange), and teEIC (green) are plotted versus the experimental temperature. 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 eEIC (37 °C) and tEIC (65 °C). (**f**) Temperature-induced unfolding of eEIC (blue), tEIC (red), etEIC (orange), and teEIC (green).

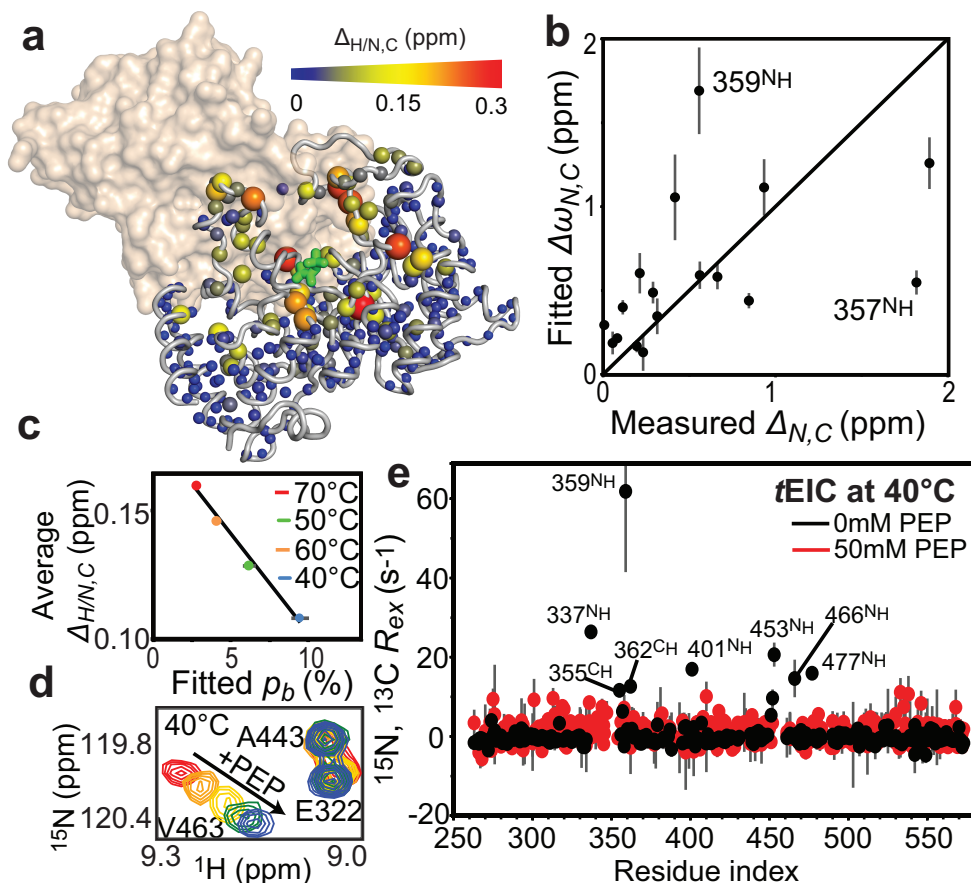**Figure 3. NMR analysis of holo *t*EIC.** (**a**) Weighted combined chemical shift perturbations ($\Delta_{H/N,C}$) induced by 10 mM PEP on the ^{1}H–^{15}N and ^{1}H-^{13}C$_{methyl}$ TROSY correlation spectra of *t*EIC displayed on the structure of the *t*EIC-PEP complex as spheres with the relationship between size and color of each sphere and chemical shift perturbation depicted by the color bar. $\Delta_{H/N,C}$ values are displayed on one subunit of the EIC dimer. The second subunit is shown as a transparent surface. (**b**) The $\Delta\omega_{N,C}$ parameters obtained from the fits of the RD data on apo *t*EIC are plotted versus the ^{15}N and ^{13}C chemical shift perturbations ($\Delta_{N,C}$) induced by addition of 10 mM PEP to *t*EIC at 70 °C. Tyr^{357} and Ala^{359} show a poor agreement between $\Delta\omega_N$ and Δ_N. These residues are adjacent to Arg^{358} that establishes salt-bridges with the phosphate group of PEP. Therefore, the presence of PEP is likely to induce additional perturbations to the ^{15}N chemical shift of Tyr^{357} and Ala^{359} that are not reflected in the corresponding $\Delta\omega_N$ (which is fitted from RD data on the apo protein). (**c**) $\Delta_{H/N,C}$ measured at different temperatures for *t*EIC are averaged over all the NMR peaks and plotted versus the population of compact state obtained by fitting the RD data on the apo protein. Data measured at 40, 50, 60, and 70 °C are shown as blue, green, orange, and red circles, respectively. Linear regression of the data is shown as a solid black line. (**d**) Close-up view of the ^{1}H-^{15}N TROSY spectrum of *t*EIC showing the effect of increasing concentrations of PEP on selected cross-peaks at 40 °C. The color code is as follows: red, 0 mM PEP; orange, 0.2 mM PEP; yellow, 0.5 mM PEP; green, 1.0 mM PEP; blue, 2.0 mM PEP. The fact that peak position follows a linear path upon increasing PEP concentration indicates that the free and bound forms of *t*EIC are in fast exchange.[84] (**e**) Exchange contribution to the transverse relaxation rates (R_{ex}) at 800 MHz and 40 °C measured for samples of *t*EIC in the absence (black) and in the presence (red) of 50 mM PEP.
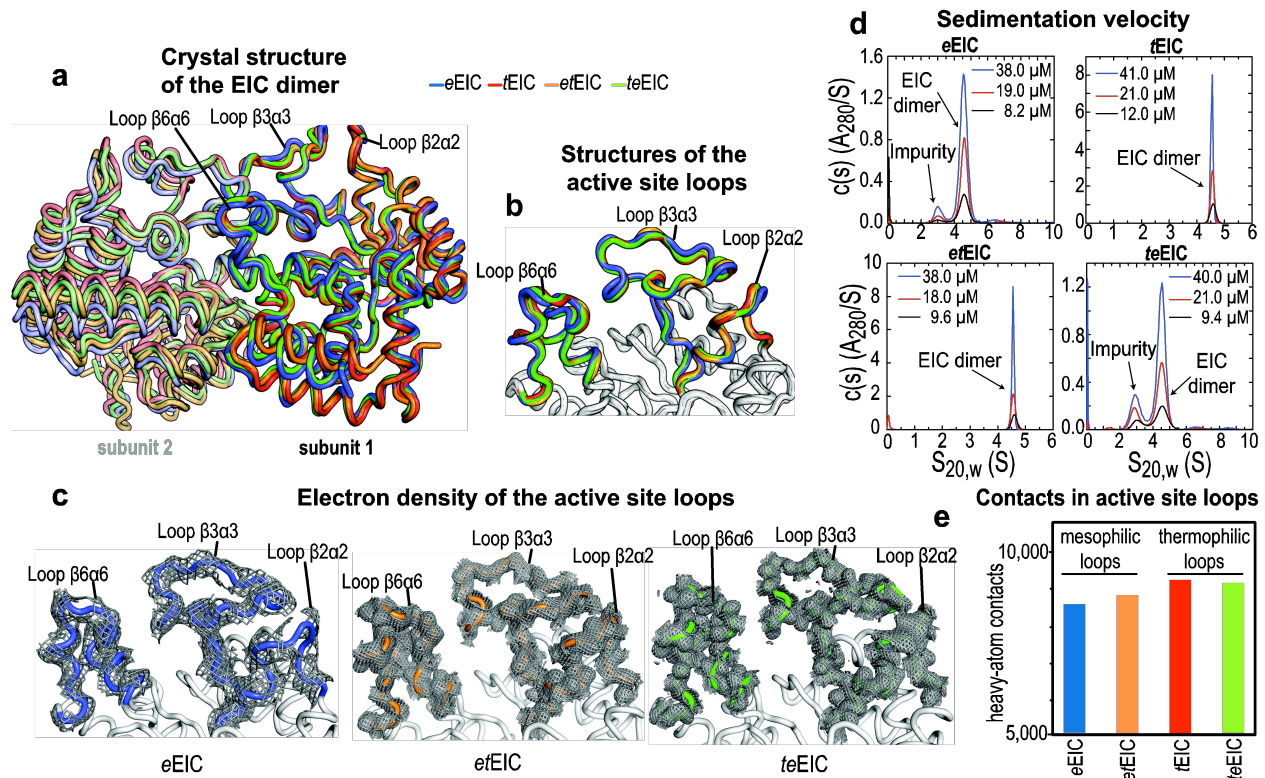
Figure 4. Crystal structures of EIC. (**a**) Overlay of the crystal structure of *e*EIC (blue), *t*EIC (red), *et*EIC (orange), and *te*EIC (green). Subunit 1 and 2 are indicated with bright and pale colors, respectively. (**b**) Close up view of the active site loops in *e*EIC (blue), *t*EIC (red), *et*EIC (orange), and *te*EIC (green). Statistics for crystal structure of *e*EIC, *et*EIC, and *te*EIC (obtained in this work) are shown in Supplementary Table S3. (**c**) Electron density maps (2fo-fc) of the active site loops in *e*EIC (left), *et*EIC (center), and *te*EIC (right) are shown as a gray mesh contoured at 0.5σ. (**d**) *c(s)* distributions for *e*EIC (top-left), *t*EIC (top-right), *et*EIC (bottom-left), and *te*EIC (bottom-right) obtained at different loading concentrations (ranging from ~50 to 10 μM) based on sedimentation velocity absorbance data collected at 50 kilo-revolutions per minute and 20 °C (see Methods). The sedimentation experiments indicate that EIC is dimeric for all tested constructs within the tested concentration range. Peaks at $S_{20,W} < 4$ S that do not show concentration dependent *c(s)* absorbance profiles (i.e. they do not report on the monomer–dimer equilibrium) are attributed to small amounts of contaminants in the AUC sample. (**e**) Number of interatomic contacts among heavy atoms in the active site loops of *e*EIC (blue), *t*EIC (red), *et*EIC (orange), and *te*EIC (green). Contacts for atom *i* were calculated as the number of heavy atoms within a distance of 5 Å from atom *i*. Data displayed in bar graph are the sum of intrasubunit and intersubunit contacts.
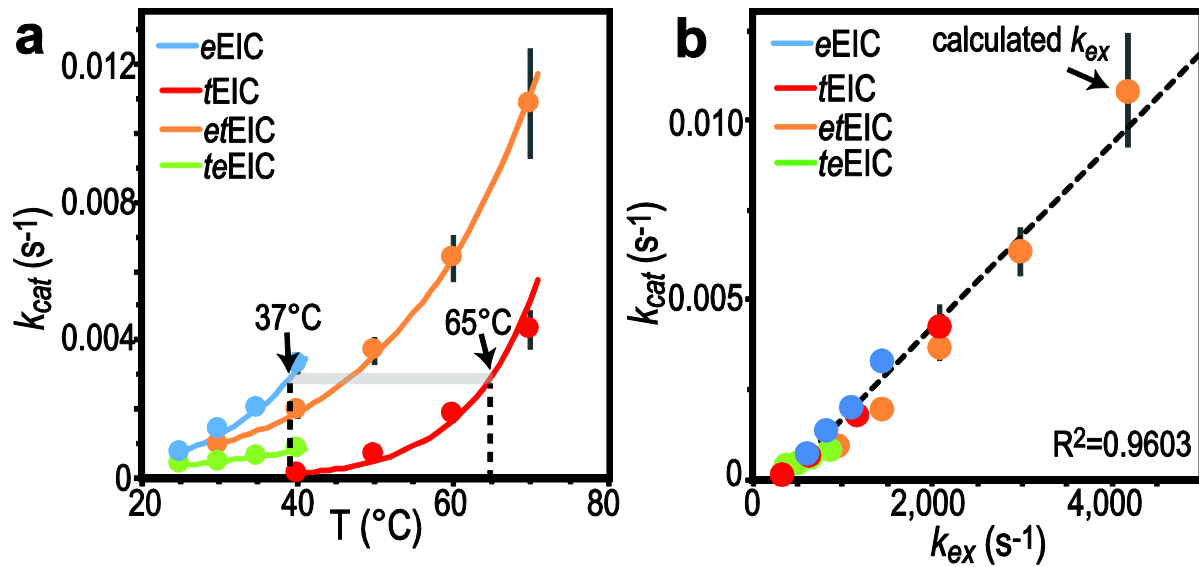
Figure 5. Temperature dependence of PEP hydrolysis. (**a**) Turnover number (k_{cat}) versus temperature for the PEP hydrolysis reaction catalyzed by eEIC (blue), tEIC (red), etEIC (orange), and teEIC (green). Experimental data are shown as filled-in circles. Modeling of the experimental data using the Eyring equation is shown as solid lines. Vertical dashed lines are at the optimal PTS temperatures for eEIC (37 °C) and tEIC (65 °C). (**b**) k_{cat} values are plotted versus the corresponding k_{ex} values (i.e. k_{cat} measured for eEIC at 40 °C is plotted versus k_{ex} measured for eEIC at 40 °C). Blue, red, orange, and green circles are for eEIC, tEIC, etEIC, and teEIC, respectively. The value of k_{ex} at 70 °C for etEIC was calculated using the Eyring equation and the $\Delta^{\ddagger}H$ and $\Delta^{\ddagger}S$ values reported in Supplementary Table S1.
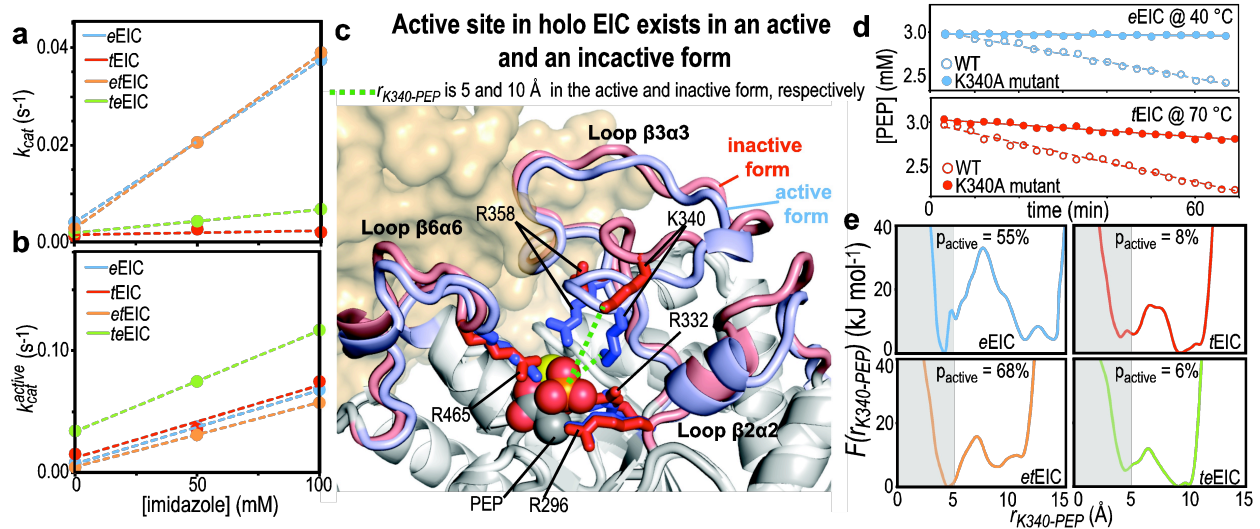
Figure 6. Conformational heterogeneity regulates EIC activity. (**a**) Turnover number (k_{cat}) versus imidazole concentration for PEP degradation (i.e. the sum of PEP hydrolysis and imidazole phosphorylation, see Supplementary Figure S6c) catalyzed by *e*EIC (blue), *t*EIC (red), *et*EIC (orange), and *te*EIC (green). Experimental data are shown as filled-in circles. Linear regressions are shown as solid lines. (**b**) k_{cat} values reported in (a) are rescaled according to $k_{cat}^{active} = k_{cat}/p_{active}$, were p_{active} is the fractional population of activated EIC and k_{cat}^{active} is the turnover number associated with EIC in the active conformation. (**c**) The coordinates of *e*EIC from the crystal structure of phosphorylated, closed *e*EI (that was used to generate the atomic-resolution model of the *e*EIC-PEP complex, see caption to Figure 1a) are superimposed onto the crystal structure of the *t*EIC-PEP complex. The active site loops are colored blue and red for *e*EIC and *t*EIC, respectively. Positively charged side-chains coordinating PEP are shown as solid sticks. PEP is shown as spheres. The magnesium ion is shown as a yellow sphere. The second subunit is shown as transparent surface. The distance between the phosphate group of PEP and the ε-ammonium group of Lys[340] ($r_{K340-PEP}$) is shown as a dashed green line. (**d**) PEP concentration versus time measured on samples containing 3 mM PEP and 50 µM enzyme. Data in the top panel are for the wild-type *e*EIC (filled-in blue circles) and the *e*EIC Lys[340]Ala mutant (open blue circles) at 40 °C. Data on the bottom panel are for the wild-type *t*EIC (filled-in red circles) and the *t*EIC Lys[340]Ala mutant (open red circles) at 70 °C. Linear regressions of the data are shown as solid or dashed lines for the wild-type and mutant enzymes, respectively. (**e**) Potential of mean force versus $r_{K340-PEP}$ obtained from 1-µs MTD simulations run on the complexes formed by PEP with *e*EIC (blue), *t*EIC (red), *et*EIC (orange), and *te*EIC (green). Values of p_{active} were obtained by integrating $F(r_{K340-PEP})$ over $r_{K340-PEP} < 5$ Å (gray box).

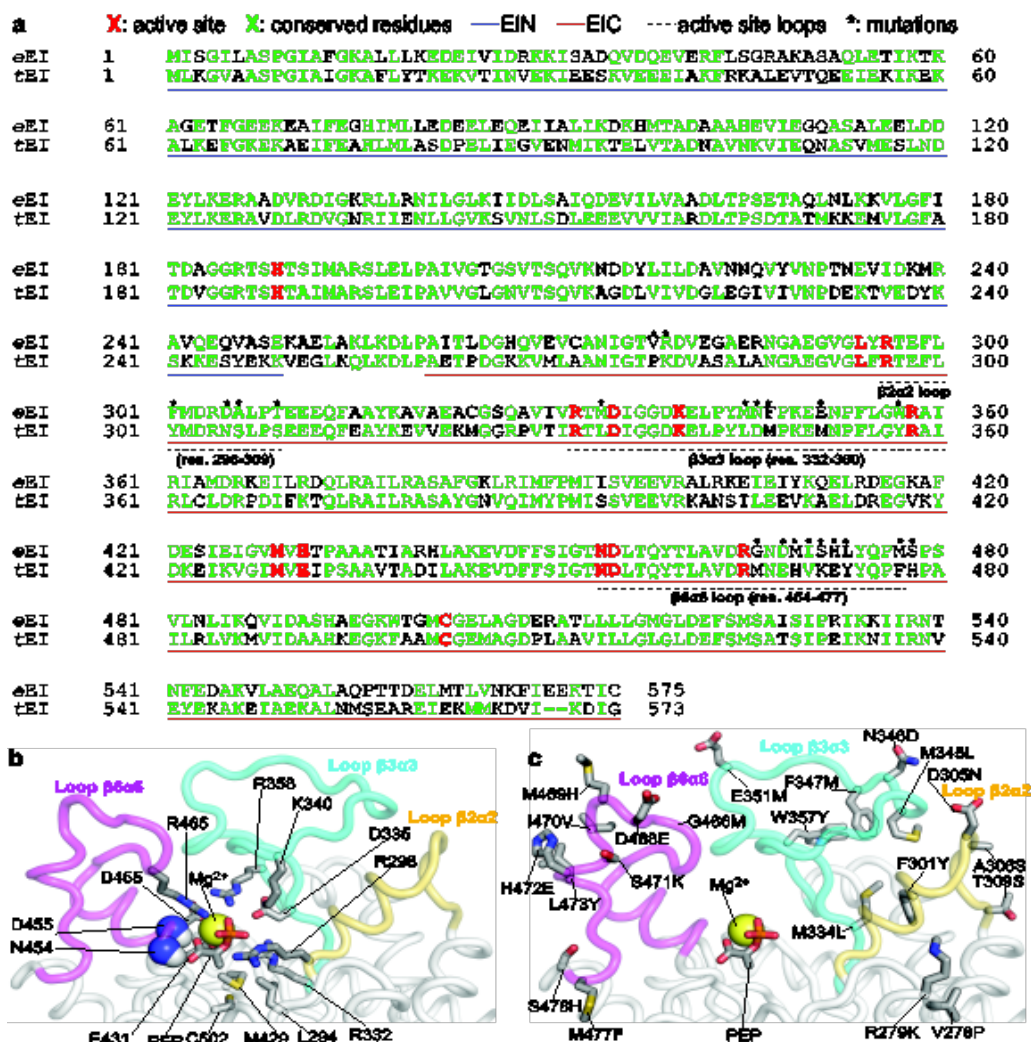# Supplemental Information

Supplemental Figures

Figure S1. Sequence alignment of *e*EI and *t*EI. (**a**) 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. Close-up view of the active site loops of *e*EIC highlighting the position of (**b**) residues involved in binding of the substrate and metal cofactor and (**c**) the 21 single point mutations performed to engineer the thermophilic/mesophilic EIC hybrids. PEP and relevant side-chains are shown as solid sticks. Backbone amides in contact with PEP are shown as spheres. The magnesium ion is displayed as a yellow sphere. The β2α2, β3α3, and β6α6 loops are colored yellow, light blue, and purple, respectively. The second subunit is not shown. 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.

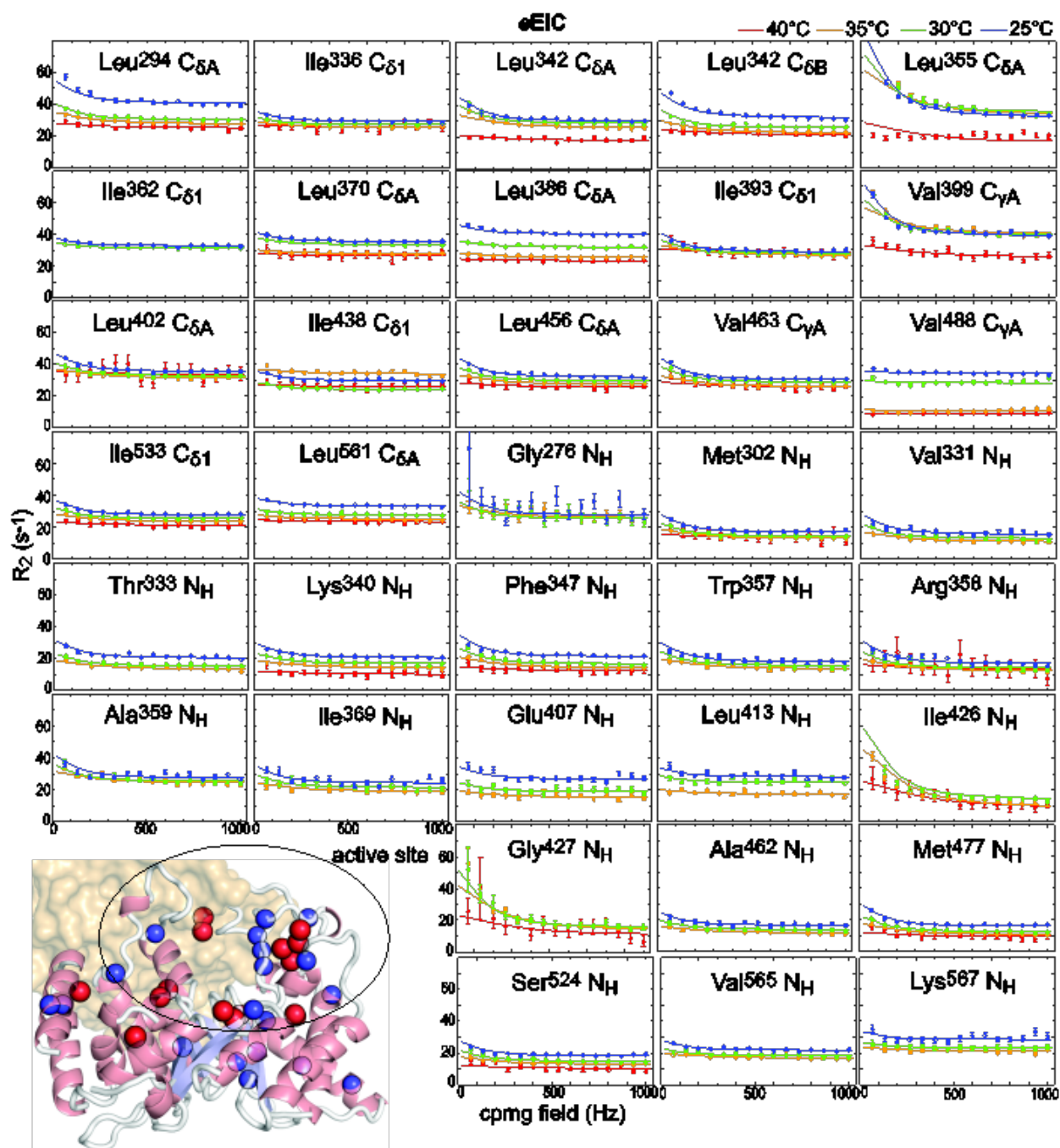**Figure S2. RD data on apo _e_EIC.** Global fitting of the 36 ^{15}N and [13]C_methyl relaxation dispersion curves at 25 (blue), 30 (green) 35 (orange), and 40 (red) °C that describe μs-ms dynamics in apo _e_EIC. Experimental data are reported as circles. Results of the global fit are shown as solid 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.
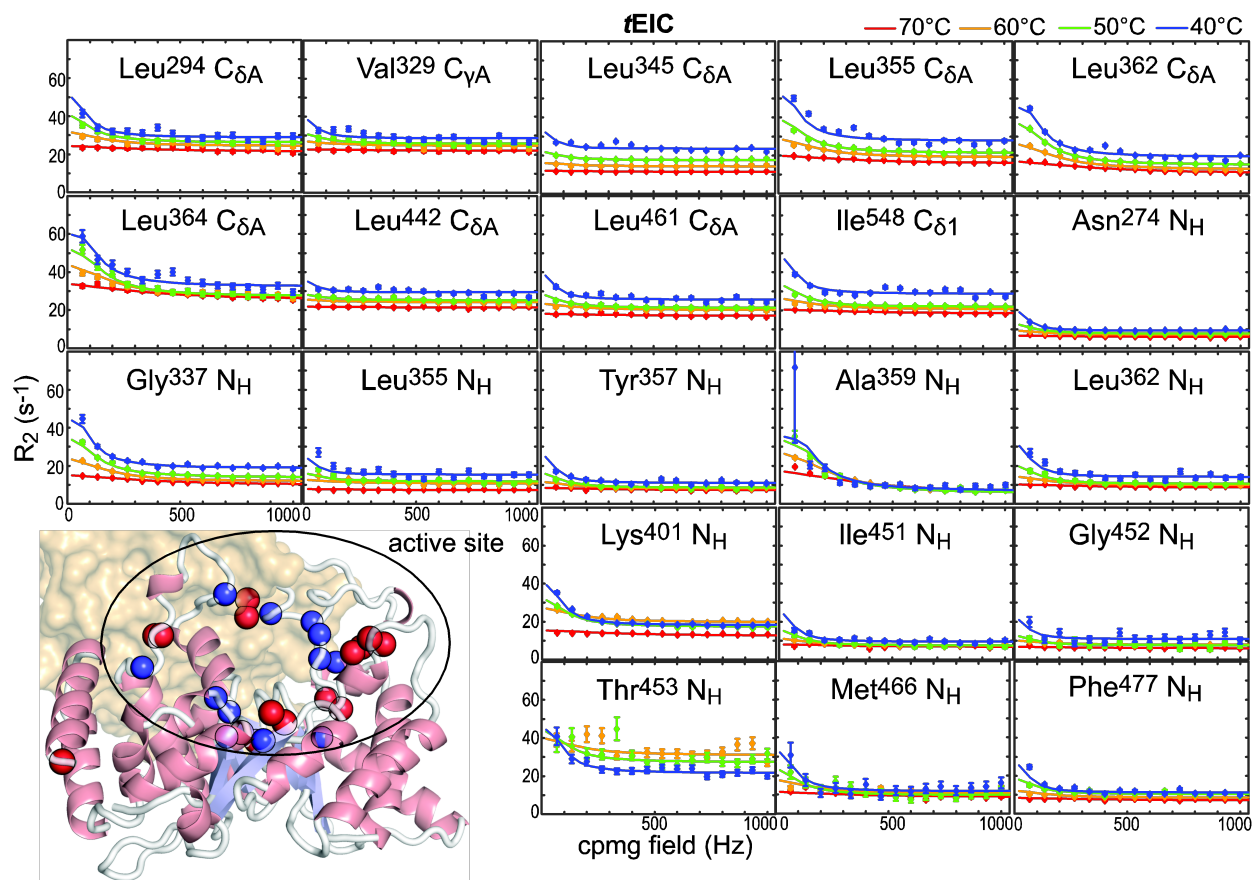
Figure S3. RD data on apo *t*EIC. Global fitting of the 21 ^{15}N and [13]C_methyl relaxation dispersion curves at 40 (blue), 50 (green) 60 (orange), and 70 (red) °C that describe μs-ms dynamics in apo *t*EIC. Experimental data are reported as circles. Results of the global fit are shown as solid lines. 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.
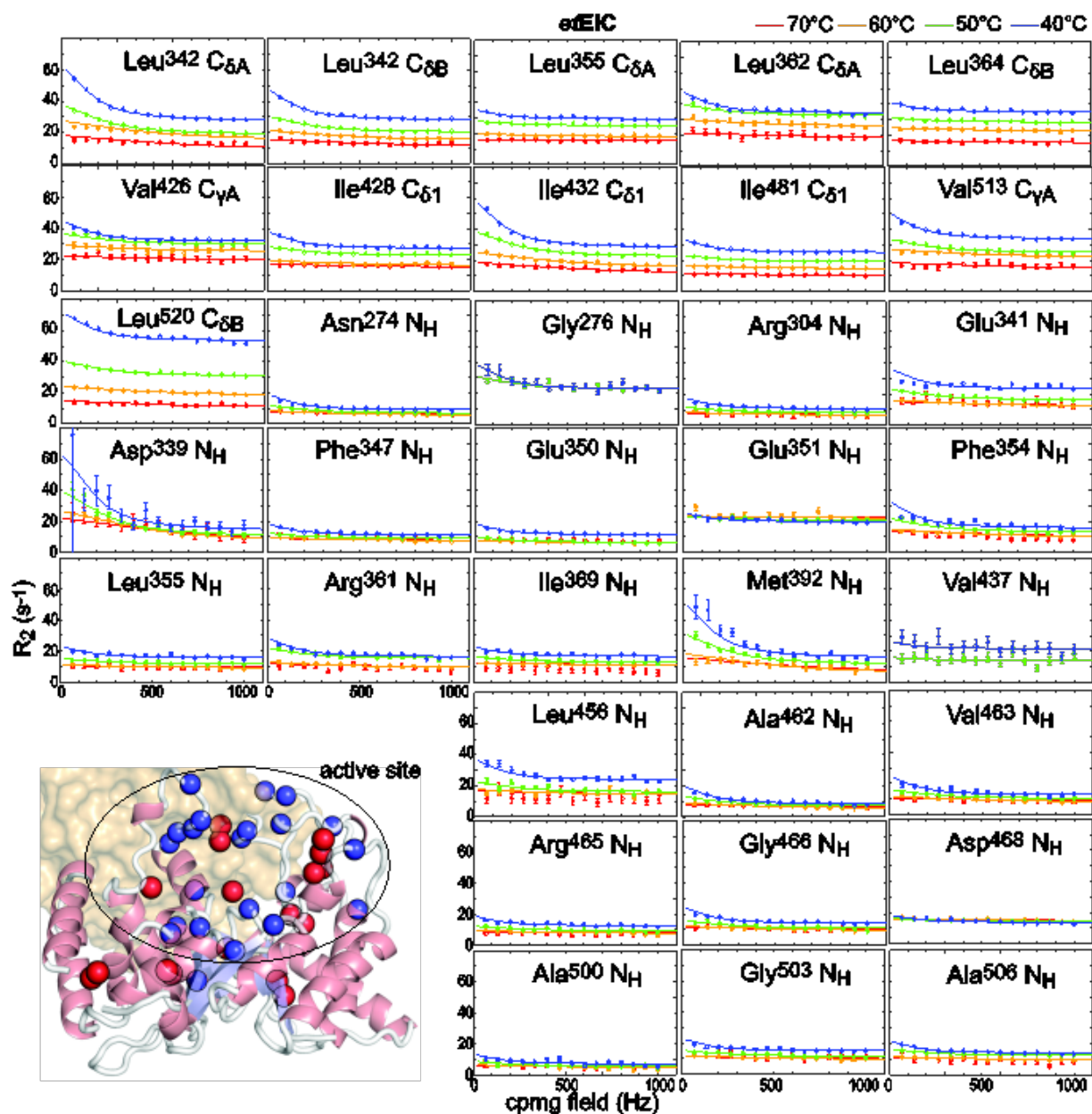
Figure S4. **RD data on apo *et*EIC.** Global fitting of the 34 ^{15}N and ^{13}C$_{methyl}$ relaxation dispersion curves at 40 (blue), 50 (green) 60 (orange), and 70 (red) °C that describe μs-ms dynamics in apo *et*EIC. Experimental data are reported as circles. Results of the global fit are shown as 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.
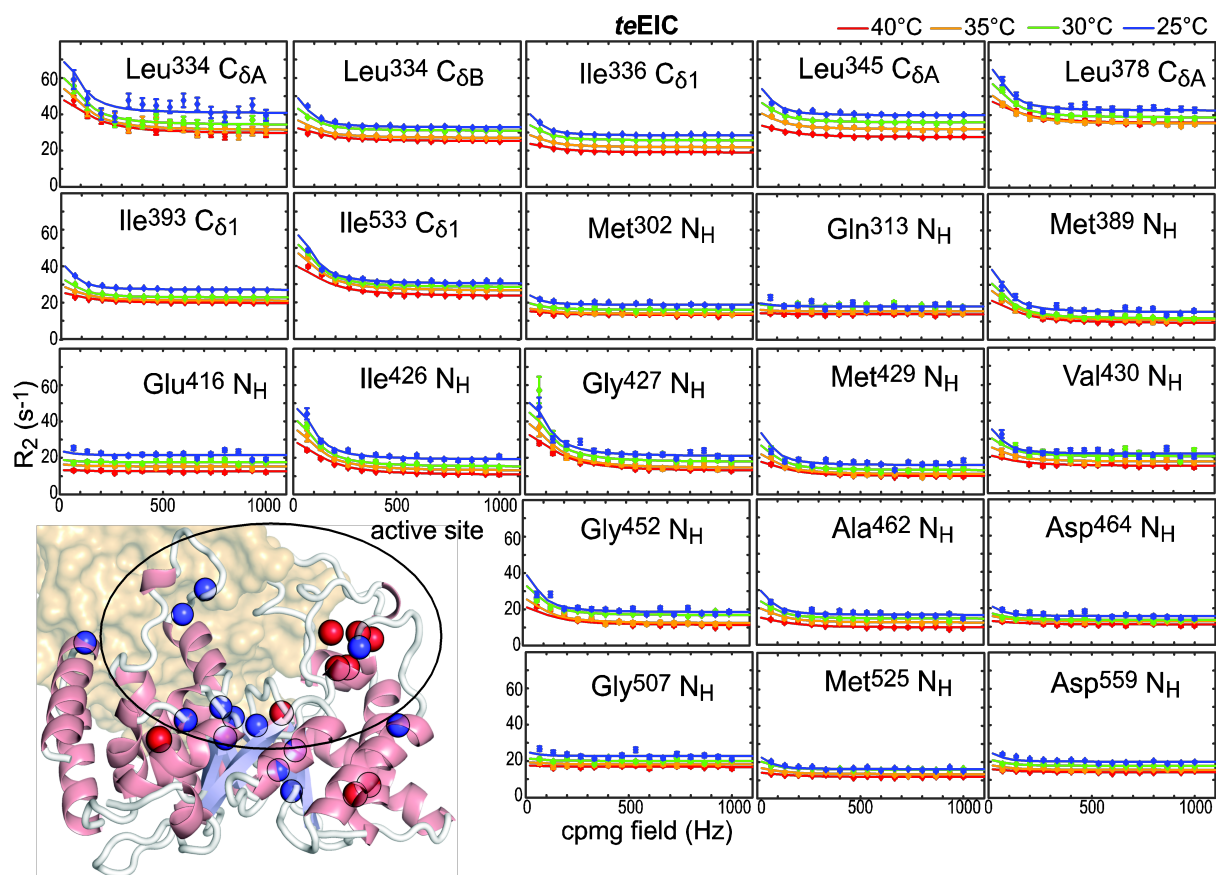
Figure S5. RD data on apo *te*EIC. Global fitting of the 21 ^{15}N and [13]C$_{methyl}$ relaxation dispersion curves at 25 (blue), 30 (green) 35 (orange), and 40 (red) °C that describe μs-ms dynamics in apo *te*EIC. Experimental data are reported as circles. Results of the global fit are shown as solid 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.
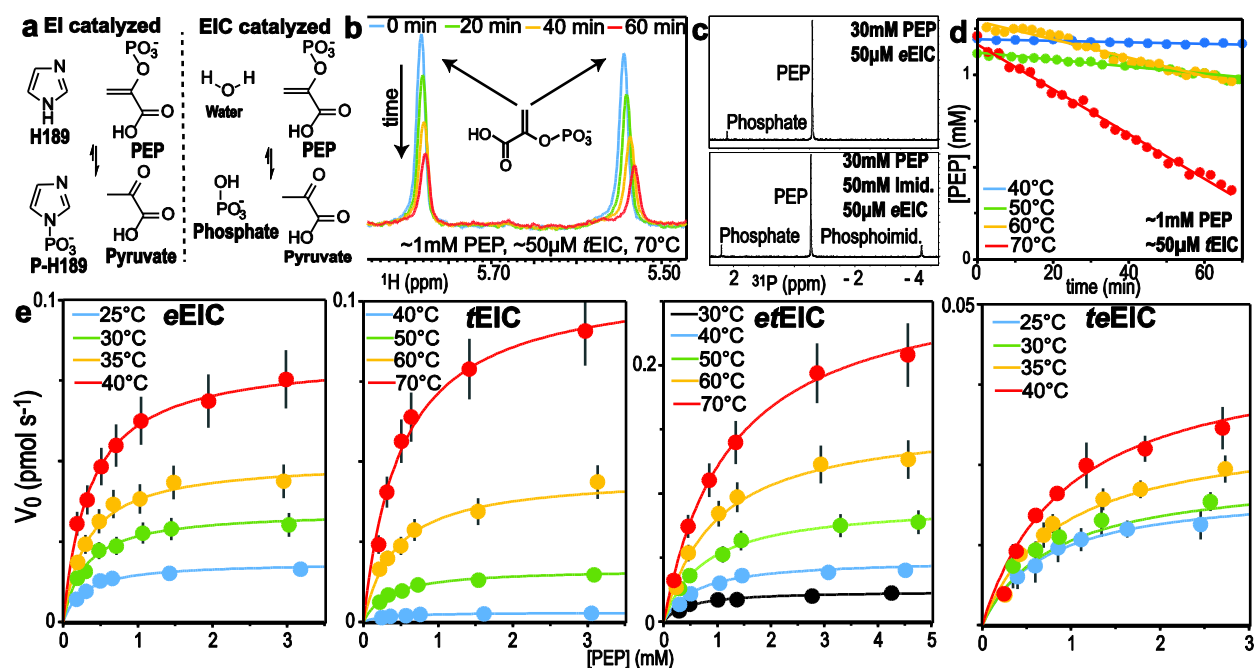
Figure S6. Enzymatic assay for EIC. (**a**) Enzymatic reactions catalyzed by full-length EI (left) and isolated EIC (right). (**b**) Region of the ^{1}H NMR spectrum of PEP displaying the signals of the two alkene protons. Spectra were measured at 70 °C in the presence of 1 mM PEP and 50 μM tEIC. Blue, 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 eEIC (top) or 30 mM PEP, 50 mM imidazole, and 50 μM eEIC (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 tEIC 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 eEIC, tEIC, etEIC, and teEIC measured at different temperatures. Experiential data are shown as filled-in circles. Fitted curves are shown as solid lines.
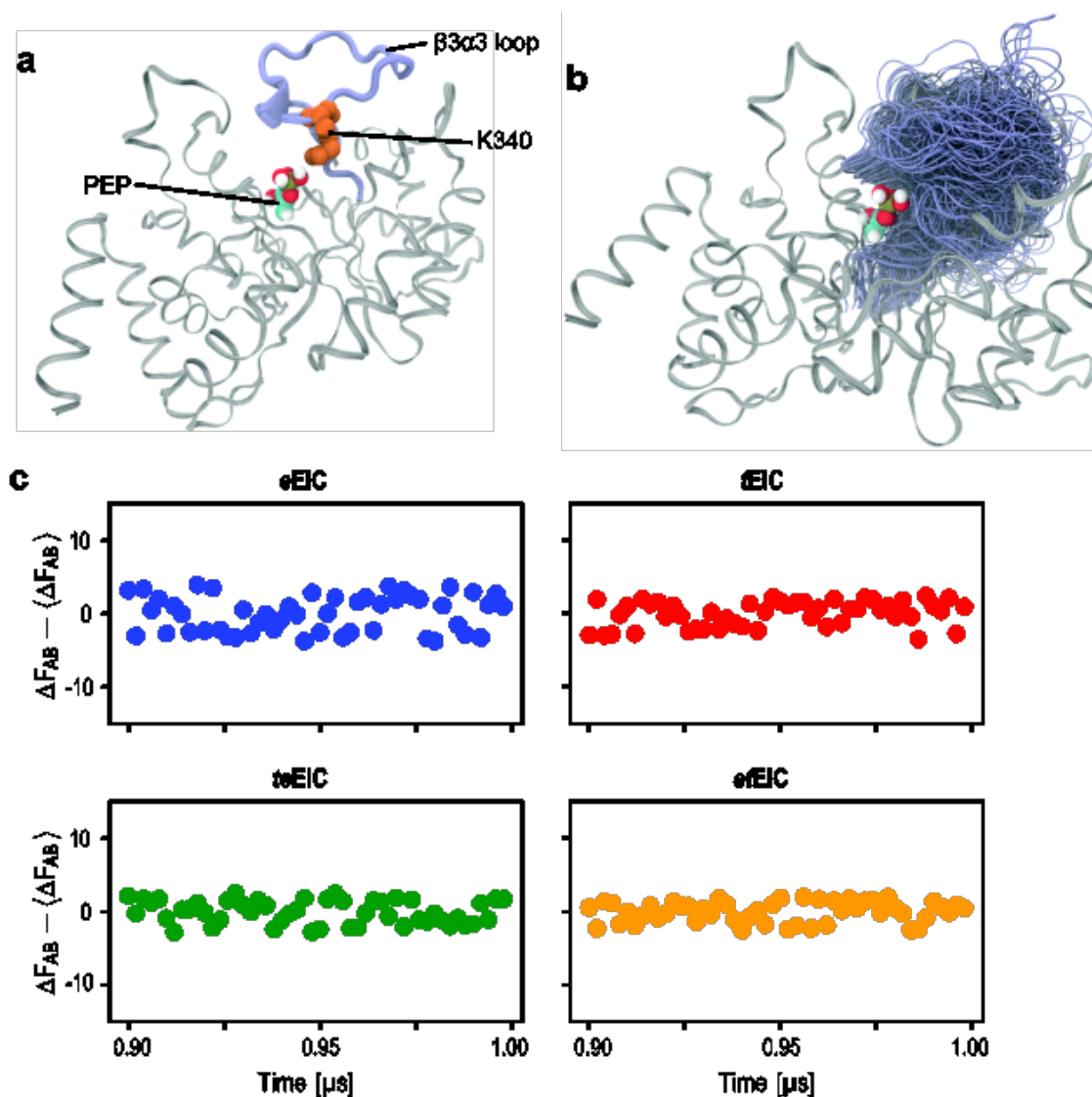
Figure S7. Convergence assessment for WT-MTD simulations. (**a**) Shown is the structure of *e*EIC highlighting the position of PEP and Lys[340]. (**b**) Shown are the loop conformations sampled by *e*EIC during the 1 µs WT-MTD simulation. Structures were extracted at 10-ns intervals. (**c**) Shown are 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.

Supplemental 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	77	87	97±7	-	-	-	16±2	-156±6	11	9	8	7	-	-	-	-29±2	-114±6
	539	740	1,002	1,342	-	-	-	45±2	-43±5									
*t*EIC	-	-	-	32	40	48	58	15±2	-168±4	-	-	-	9	6	4	3	-39±2	-142±5
	-	-	-	303	599	1,124	2,026	54±2	-26±4									
*et*EIC	-	105	-	111	115	119	-	1±0.2	-203±5	-	11	-	8	6	4	-	-30±2	-118±7
	-	853	-	1,321	1,975	2,867	-	31±2	-85±5									
*te*EIC	39	45	50	57	-	-	-	16±3	-160±10	10	9	8	7	-	-	-	-25±2	-102±5
	342	459	608	797	-	-	-	41±3	-58±10									
*te*EIC*	37	44	51	60	-	-	-	22±3	-141±10	10	9	8	7	-	-	-	-25±2	-102±5
	352	475	635	840	-	-	-	42±3	-54±10									

[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. X-ray 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.

CHAPTER 4.   RESONANCE ASSIGNEMNT OF THE 128 kDa ENZYME I DIMER FROM *THERMOANAEROBACTER TENCONGENSIS*

Rochelle Dotas[1] and Vincenzo Venditti[1, 2]

Author Affiliations

[1] Department of Chemistry, Iowa State University

[2] Roy J. Carver Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State

University

Abstract

Enzyme I (EI) of the bacterial phosphotransferase system (PTS) utilizes

phosphoenolpyruvate (PEP) as a source of energy in order to transport sugars across the cellular

membrane. PEP binding to EI initiates a phosphorylation cascade that regulates a variety of

essential pathways in the metabolism of bacterial cells. Given its central role in controlling

bacterial metabolism, EI has been often suggested as a good target for antimicrobial research.

Here, we report the ^1H_N, ^{15}N, ^{13}C', $^1H_{methyl}$, and $^{13}C_{methyl}$ chemical shifts of the 128 kDa

homodimer EI from the thermophile *Thermoanaerobacter tengcongensis*. In total 79% of the

expected backbone amide correlations and 80% of the expected methyl TROSY peaks from U-

$[^2H, ^{13}C, ^{15}N]$, Ile$^{\delta 1}$-$[^{13}CH_3]$, Val-Leu-$[^{13}CH_3/^{12}CD_3]$ labeled EI were assigned. The reported

assignments will enable future structural studies aimed at illuminating the fundamental

mechanisms governing long-range interdomain communication in EI and at indicating new

therapeutic strategies to combat bacterial infections.

Biological Context

Enzyme I is the first enzyme of the phosphotransferase system (PTS) (Clore and Venditti 2013), a signal transduction pathway that controls multiple bacterial functions, including (i) transport and phosphorylation of carbohydrates, (ii) chemotaxis, (iii) biofilm formation, (iv) interactions between carbon and nitrogen metabolisms, (v) catabolic gene expression, (vi) potassium transport, and (vii) inducer exclusion (Deutscher et al. 2014; Postma et al. 1996). The multiple regulatory functions of PTS are controlled by the phosphorylation state of EI that, in turn, determines the phosphorylation state of all other PTS components (Clore and Venditti 2013; Deutscher et al. 2014). Therefore, EI is a master regulator of bacterial metabolism and a promising target for antimicrobial design.

Several structural studies have been reported on EI from *Escherichia coli*, revealing that the functional form of the enzyme is a 128 kDa dimer comprising two structurally and functionally distinct domains separated by a helical linker (Schwieters et al. 2010; Teplyakov et al. 2006). The N-terminal domain (EIN, residues 1-248) contains the phosphorylation site (His189) and the binding site for the phosphocarrier protein HPr (the second PTS protein) (Chauvin et al. 1996). The C-terminal domain (EIC, residues 261-575) is responsible for protein dimerization and contains the binding site for the substrate phosphoenolpyruvate (PEP) and for the small-molecule metabolite α-ketoglutarate (αKG) (Chauvin et al. 1996; Venditti et al. 2013), which can act as competitive inhibitor or an allosteric activator of EI depending on the oligomerization state of the enzyme (Nguyen et al. 2018). The three dimensional fold of EI is modulated by an intricate series of intra- and inter-domain conformational equilibria including: (i) a monomer/dimer equilibrium via EIC (Patel et al. 2006), (ii) a compact/expanded equilibrium within EIC (Venditti and Clore 2012; Venditti et al. 2015b), (iii) a state A/state B equilibrium within EIN (Schwieters et al. 2010; Teplyakov et al. 2006; Venditti et al. 2015a),

and (iv) an open/closed equilibrium that involves reorientation of EIN relative to EIC (Schwieters et al. 2010; Teplyakov et al. 2006; Venditti et al. 2015a; Venditti et al. 2015b). However, the contribution of each equilibrium in functional regulation of EI remains largely unsolved.

Here we present resonance assignment for the ^1H_N-^{15}N and ^1H-$^{13}C_{methyl}$ TROSY spectra (Pervushin et al. 1997; Tugarinov et al. 2003) of EI from the thermophilic bacterium *Thermoanaerobacter tengcongensis*. The high thermal stability of the thermophilic enzyme (melting temperature = 90 °C) (Navdaeva et al. 2011; Oberholzer et al. 2005) will enable NMR studies on the structure, dynamics, and thermodynamics of the EI conformational equilibria on an extended temperature range (40 – 80 °C). We expect this analysis to reveal fundamental mechanistic insights into the synergetic coupling among intradomain, interdomain, and intersubunit conformational equilibria that regulates EI and, more generally, the activity of complex multidomain enzymes.

Materials and Methods

Protein Expression and Purification

Full length EI and the isolated EIN domain from the thermophile *T. tengcongensis* were cloned into a pet21A vector (Novagen). For expression of *T. tengcongensis* EIC, a construct including EIC, a N-terminal Maltose Binding Protein (MBP)-tag, and a N-terminal His-tag (His$_6$-MBP-EIC) was cloned into a pET11a vector (Novagen). All proteins were expressed in *E. coli* BL21(DE3) (Invitrogen) cell line. Cells were grown at 37°C in a D$_2$O-based minimal medium. Composition of the medium was adjusted in order to obtain U-[^2H, ^{13}C, ^{15}N], Ile$^{\delta1}$-[$^{13}CH_3$], Val-Leu-[$^{13}CH_3$/$^{12}CD_3$] labeled proteins, as described by Tugarinov et al (Tugarinov et al. 2006). At OD$_{600}$ ~ 0.8 the temperature was decreased to 20°C and expression was induced with 1 mM isopropyl-D-thiogalactopyranoside (IPTG). Cells were harvested by centrifugation (40 min at

4,000 g) after 16 h of induction. EI and EIN were purified following the same procedure employed for the *E. coli* homologues (Venditti et al. 2011; Venditti et al. 2015b). EIC was purified as described previously (Dotas and Venditti 2017). All proteins were incubated at 80 °C for two hours in order to precipitate any remaining protein contaminants.

NMR Spectroscopy

Samples were prepared in 20 mM Tris-HCl pH 7.4, 100 mM NaCl, 4 mM MgCl$_2$, 1 mM EDTA, 2 mM DTT, and 90% H$_2$O/10% D$_2$O (v/v). The protein concentration was ~ 1 mM. All NMR experiments were measured at 70°C using 600, 700, and 800 MHz spectrometers equipped with triple resonance z-shielded gradient cryoprobe. Sequential ^{1}H/^{15}N/^{13}C backbone assignments were carried out using transverse relaxation optimized (TROSY) versions (Pervushin et al. 1997; Tugarinov et al. 2003) of conventional 3D triple resonance correlation experiments (HNCO, HNCA, HN(CO)CA, HNCACB and HN(CO)CACB) (Clore and Gronenborn 1998). 2D ^{1}H-^{13}C$_{methyl}$ TROSY experiments were assigned using 3D HMCM[CG]CBCA and HMCM(CBCA)CO experiments (Tugarinov and Kay 2003; Tugarinov et al. 2014). All data were processed using NMRPipe (Delaglio et al. 1995) and analyzed using the program Sparky (http://www.cgl.ucsf.edu/home/sparky).

Weighted combined ^{1}H$_N$/^{15}N ($\Delta_{H/N}$) and ^{1}H/^{13}C ($\Delta_{H/C}$) chemical shift perturbations resulting from comparison of the NMR spectra acquired for the isolated domains and the full-length EI were calculated using the following equation: $\Delta_{H/N,C}=((\Delta\delta_H W_H)^2+(\Delta\delta_{N,C} W_{N,C})^2)^{1/2}$, where W_H, W_N and W_C are weighing factors for the ^{1}H, ^{15}N and ^{13}C shifts, respectively ($W_H=|\gamma_H/\gamma_H|=1$, $W_N=|\gamma_N/\gamma_H|=0.101$, $W_C=|\gamma_C/\gamma_H|=0.251$); $\Delta\delta_H$, $\Delta\delta_N$, and $\Delta\delta_C$ are the ^{1}H, ^{15}N, and ^{13}C chemical shift differences in ppm, respectively; and γ_H, γ_N, and γ_C are the ^{1}H, ^{15}N, and ^{13}C gyromagnetic ratios, respectively

Assignment and Data Deposition

Given the large molecular size of EI, the NMR signals of the enzyme relax too quickly to allow acquisition of the conventional set of triple resonance experiments (HNCO, HNCA, HNCACB, HN(CO)CA, and HN(CO)CACB) required to achieve assignment of the backbone ^1H_N, ^{15}N, $^{13}C_\alpha$, $^{13}C_\beta$ and $^{13}C'$ chemical shifts. Here, the assignment of the ^1H_N-^{15}N and ^1H-$^{13}C_{methyl}$ TROSY spectra of EI were obtained by transferring the assignments of the isolated EIN and EIC domains on the spectra of the full-length protein. This approach is feasible because the amide and methyl TROSY spectra of the isolated domains overlay well with the spectra of the full-length protein (Fig. 1). Ambiguities in the assignment were sorted out by comparing 3D HNCO spectra acquired for EI, EIN, and EIC.

Backbone resonance assignment for EIC at pH 7.4 and 70 °C was reported recently by us (BMRB code: 27183) (Dotas and Venditti 2017). Here, ^1H-$^{13}C_{methyl}$ resonances of U-[^2H, ^{13}C, ^{15}N], Ile$^{\delta 1}$-[$^{13}CH_3$], Val-Leu-[$^{13}CH_3$/$^{12}CD_3$] labeled EIC were assigned using out-and-back experiments (Tugarinov and Kay 2003; Tugarinov et al. 2014). 78% of the expected ^1H-$^{13}C_{methyl}$ resonances were assigned (94 out of 121), and the chemical shift information was deposited in the BioMagResBank (http://www.bmrb.wisc.edu/) (Ulrich et al. 2008) under the BMRB accession code 27832.

Backbone and side chain assignments for EIN were recently obtained at pH 6.5 and 37 °C (BMRB code: 30326) (Evangelidis et al. 2018) and used to solve the solution structure of the N-terminal domain. Here, HNCO, HNCA, HNCACB, HN(CO)CA, HN(CO)CACB, HMCM(CG)CBCA, and HMCM(CGCBCA)CO experiments were used to obtain ^1H_N, ^{15}N, $^{13}C_\alpha$, $^{13}C_\beta$, $^{13}C'$, $^1H_{methyl}$, and $^{13}C_{methyl}$ chemical shifts of U-[^2H, ^{13}C, ^{15}N], Ile$^{\delta 1}$-[$^{13}CH_3$], Val-Leu-[$^{13}CH_3$/$^{12}CD_3$] labeled EIN. In total, 95% of all backbone resonances (94% of ^1H_N, 92% of ^{15}N, 98% of $^{13}C_\alpha$, 99% of $^{13}C_\beta$ and 92% of 13C′ nuclei) and 93% of the expected ^1H-$^{13}C_{methyl}$ peaks

were assigned (109 out of 117). Interestingly, the C_α and C_β secondary chemical shifts measured at pH 6.5 and 37 °C are in good agreement with the ones measured at pH 7.4 and 70 °C (Fig. 2), indicating that the change in pH and temperature does not perturb the fold of EIN, and providing confidence on the obtained assignments. Chemical shifts measured for EIN at pH 7.4 and 70 °C were deposited on the BioMagResBank under the BMRB accession code 27833.

Using the chemical shift assignments of the isolated domains as a reference, 439 ^1H_N-^{15}N and 196 ^1H-$^{13}C_{methyl}$ correlations were assigned for the full-length EI (Figs. 3 and 4), corresponding to 79% and 80% of the expected amide and methyl TROSY peaks, respectively. The unassigned peaks predominantly cluster at the loop preceding the active site His[189] on EIN (Gly[184] – Thr[187]) and within the EIC domain (Fig. 5a). These areas are expected to experience conformational dynamics on the μs-ms timescale, suggesting that exchange broadening might cause attenuation of these resonances beyond the detection limit. An additional stretch of contiguous unassigned residues is observed at the linker connecting the N- and C-terminal domains of EI (Lys[243] – Pro[264]). These residues either were not included in the constructs of the isolated EIN and EIC domains, or their assignments could not be transferred unambiguously from the spectra of the isolated domains to the spectrum of full-length EI (i.e. they experience a combined ^1H/^{15}N or ^1H/^{13}C chemical shift change > 0.4 ppm going from the isolated domains to the full-length enzyme). It is also interesting to note that combined chemical shift differences larger than 0.1 ppm are observed exclusively in the vicinity of the linker helix (Fig. 5b,c), which confirms the absence of significant interdomain interactions in the apo EI. In this regard, it is worth mentioning that we do not expect these chemical shift differences to originate from unfolding of the linker helix, as the C_α and C_β secondary chemical shifts measured for the C-terminal residues of the isolated EIN domain are consistent with formation of an α-helix

(Evangelidis et al. 2018). However, as C_α and C_β chemical shifts could not be measured for the full-length enzyme, we cannot prove unambiguously the presence of the helical linker in solution.

^1H_N, ^{15}N, $^{13}C'$, $^1H_{methyl}$, and $^{13}C_{methyl}$ chemical shifts assigned for EI have been deposited in the BioMagResBank under the accession number 27762.

Acknowledgments

References

Chauvin F, Brand L, Roseman S (1996) Enzyme I: the first protein and potential regulator of the bacterial phosphoenolpyruvate: glycose phosphotransferase system. Res Microbiol 147:471-479

Clore GM, Gronenborn AM (1998) Determining the structures of large proteins and protein complexes by NMR. Trends in biotechnology 16:22-34

Clore GM, Venditti V (2013) Structure, dynamics and biophysics of the cytoplasmic protein-protein complexes of the bacterial phosphoenolpyruvate: sugar phosphotransferase system. Trends Biochem Sci 38:515-530 doi:10.1016/j.tibs.2013.08.003

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277-293

Deutscher J et al. (2014) The bacterial phosphoenolpyruvate:carbohydrate phosphotransferase system: regulation by protein phosphorylation and phosphorylation-dependent protein-protein interactions. Microbiol Mol Biol Rev 78:231-256 doi:10.1128/MMBR.00001-14

Dotas RR, Venditti V (2017) (1)H, (15)N, (13)C backbone resonance assignment of the C-terminal domain of enzyme I from Thermoanaerobacter tengcongensis. Biomol NMR Assign doi:10.1007/s12104-017-9788-x

Evangelidis T et al. (2018) Automated NMR resonance assignments and structure determination using a minimal set of 4D spectra. Nat Commun 9:384 doi:10.1038/s41467-017-02592-z

Navdaeva V et al. (2011) Phosphoenolpyruvate: sugar phosphotransferase system from the hyperthermophilic Thermoanaerobacter tengcongensis. Biochemistry 50:1184-1193

Nguyen TT, Ghirlando R, Venditti V (2018) The oligomerization state of bacterial enzyme I (EI) determines EI's allosteric stimulation or competitive inhibition by alpha-ketoglutarate. J Biol Chem doi:10.1074/jbc.RA117.001466

Oberholzer AE, Bumann M, Schneider P, Bachler C, Siebold C, Baumann U, Erni B (2005) Crystal structure of the phosphoenolpyruvate-binding enzyme I-domain from the Thermoanaerobacter tengcongensis PEP: sugar phosphotransferase system (PTS). Journal of molecular biology 346:521-532

Patel HV, Vyas KA, Mattoo RL, Southworth M, Perler FB, Comb D, Roseman S (2006) Properties of the C-terminal domain of enzyme I of the Escherichia coli phosphotransferase system. J Biol Chem 281:17579-17587 doi:10.1074/jbc.M508966200

Pervushin K, Riek R, Wider G, Wuthrich K (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci U S A 94:12366-12371

Postma PW, Lengeler JW, Jacobson GR (1996) Phosphoenolpyruvate:Carbohydrate Phosphotransferase Systems. In: Neidhardt FC, Lin EC, Curtiss R (eds) Escherichia coli and Salmonella: Cellular and Molecular Biology. pp 1149-1174

Schwieters CD, Suh JY, Grishaev A, Ghirlando R, Takayama Y, Clore GM (2010) 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. Journal of the American Chemical Society 132:13026-13045

Teplyakov A et al. (2006) Structure of phosphorylated enzyme I, the phosphoenolpyruvate:sugar phosphotransferase system sugar translocation signal protein. Proceedings of the National Academy of Sciences of the United States of America 103:16218-16223

Tugarinov V, Hwang PM, Ollerenshaw JE, Kay LE (2003) Cross-correlated relaxation enhanced 1H[bond]13C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. Journal of the American Chemical Society 125:10420-10428 doi:10.1021/ja030153x

Tugarinov V, Kanelis V, Kay LE (2006) Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. Nat Protoc 1:749-754 doi:10.1038/nprot.2006.101

Tugarinov V, Kay LE (2003) Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. Journal of the American Chemical Society 125:13868-13878 doi:10.1021/ja030345s

Tugarinov V, Venditti V, Marius Clore G (2014) A NMR experiment for simultaneous correlations of valine and leucine/isoleucine methyls with carbonyl chemical shifts in proteins. J Biomol NMR 58:1-8 doi:10.1007/s10858-013-9803-1

Ulrich EL et al. (2008) BioMagResBank. Nucleic Acids Res 36:D402-408 doi:10.1093/nar/gkm957

Venditti V, Clore GM (2012) Conformational selection and substrate binding regulate the monomer/dimer equilibrium of the C-terminal domain of Escherichia coli enzyme I. J Biol Chem 287:26989-26998 doi:10.1074/jbc.M112.382291

Venditti V, Fawzi NL, Clore GM (2011) Automated sequence- and stereo-specific assignment of methyl-labeled proteins by paramagnetic relaxation and methyl-methyl nuclear Overhauser enhancement spectroscopy. J Biomol NMR 51:319-328 doi:10.1007/s10858-011-9559-4

Venditti V, Ghirlando R, Clore GM (2013) Structural basis for enzyme I inhibition by alpha-ketoglutarate. ACS Chem Biol 8:1232-1240 doi:10.1021/cb400027q

Venditti V, Schwieters CD, Grishaev A, Clore GM (2015a) Dynamic equilibrium between closed and partially closed states of the bacterial Enzyme I unveiled by solution NMR and X-ray scattering. Proc Natl Acad Sci U S A 112:11565-11570 doi:10.1073/pnas.1515366112

Venditti V, Tugarinov V, Schwieters CD, Grishaev A, Clore GM (2015b) Large interdomain rearrangement triggered by suppression of micro- to millisecond dynamics in bacterial Enzyme I. Nat Commun 6:5960 doi:10.1038/ncomms6960
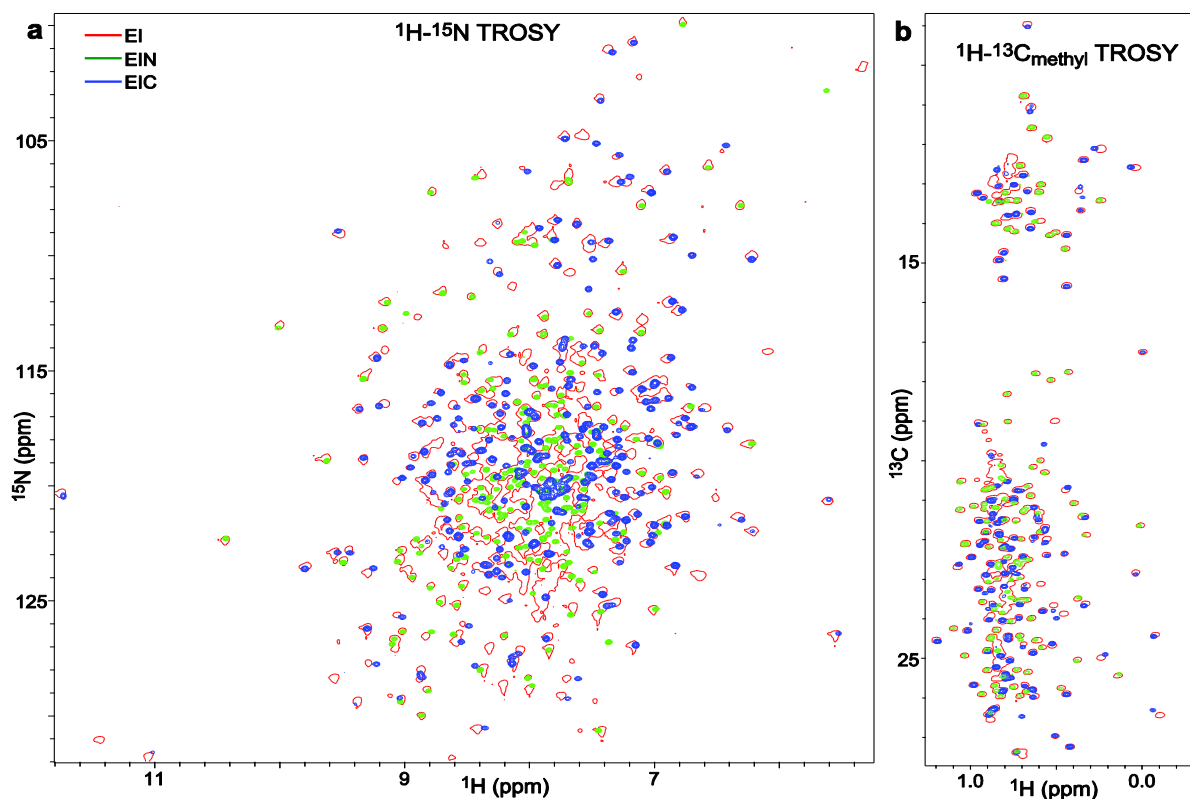
Figures

Fig. 1 Overlay of (**a**) the ^1H_N–^{15}N TROSY and (**b**) ^1H–$^{13}C_{methyl}$ TROSY spectra acquired at 800 MHz, pH 7.4, and 70 °C for the 128 kDa full-length EI dimer (red), the 27 kDa EIN domain (green), and the 70 kDa EIC dimer (blue). Only the lowest contour is shown for the spectrum of full-length EI. All proteins were U-[^2H, ^{13}C, ^{15}N], Ile$^{\delta 1}$-[$^{13}CH_3$], Val-Leu-[$^{13}CH_3$/$^{12}CD_3$] labeled
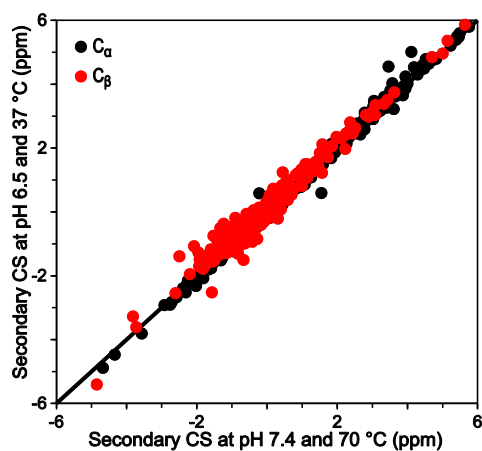
Fig. 2 C_α and C_β secondary chemical shifts measured for EIN at pH 6.5 and 37 °C (BMRB code: 30326) (Evangelidis et al. 2018) versus the data measured at pH 7.4 and 70 °C (this work)
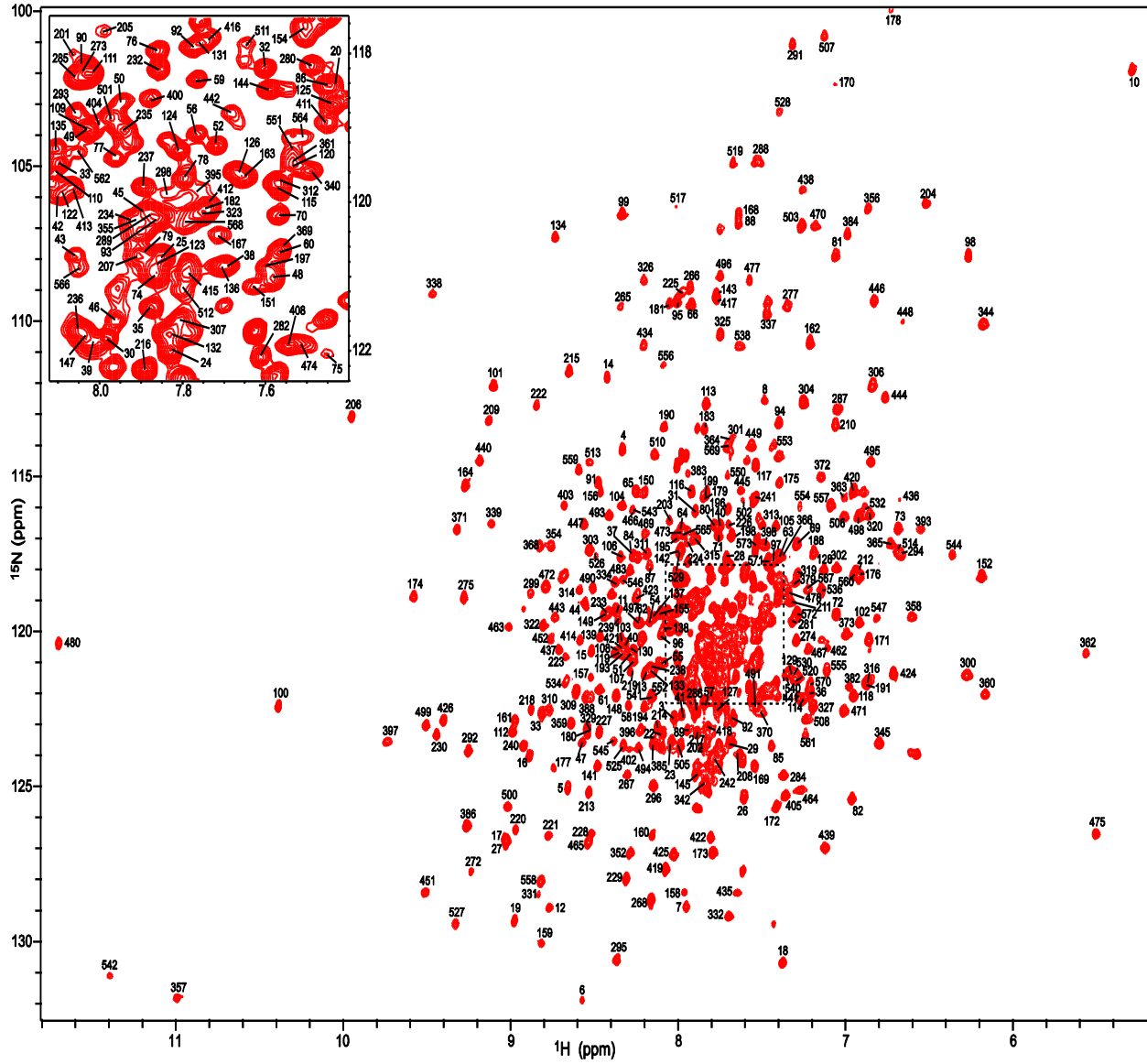
Fig. 3 800 MHz $^{1}H_N$-^{15}N-TROSY spectrum of EI from *T. tengcongensis* at pH 7.4 and 70 °C. Assigned cross-peaks are labeled with the corresponding sequence numbers. The central region of the spectrum (central box) is expanded (×2) in the top left corner
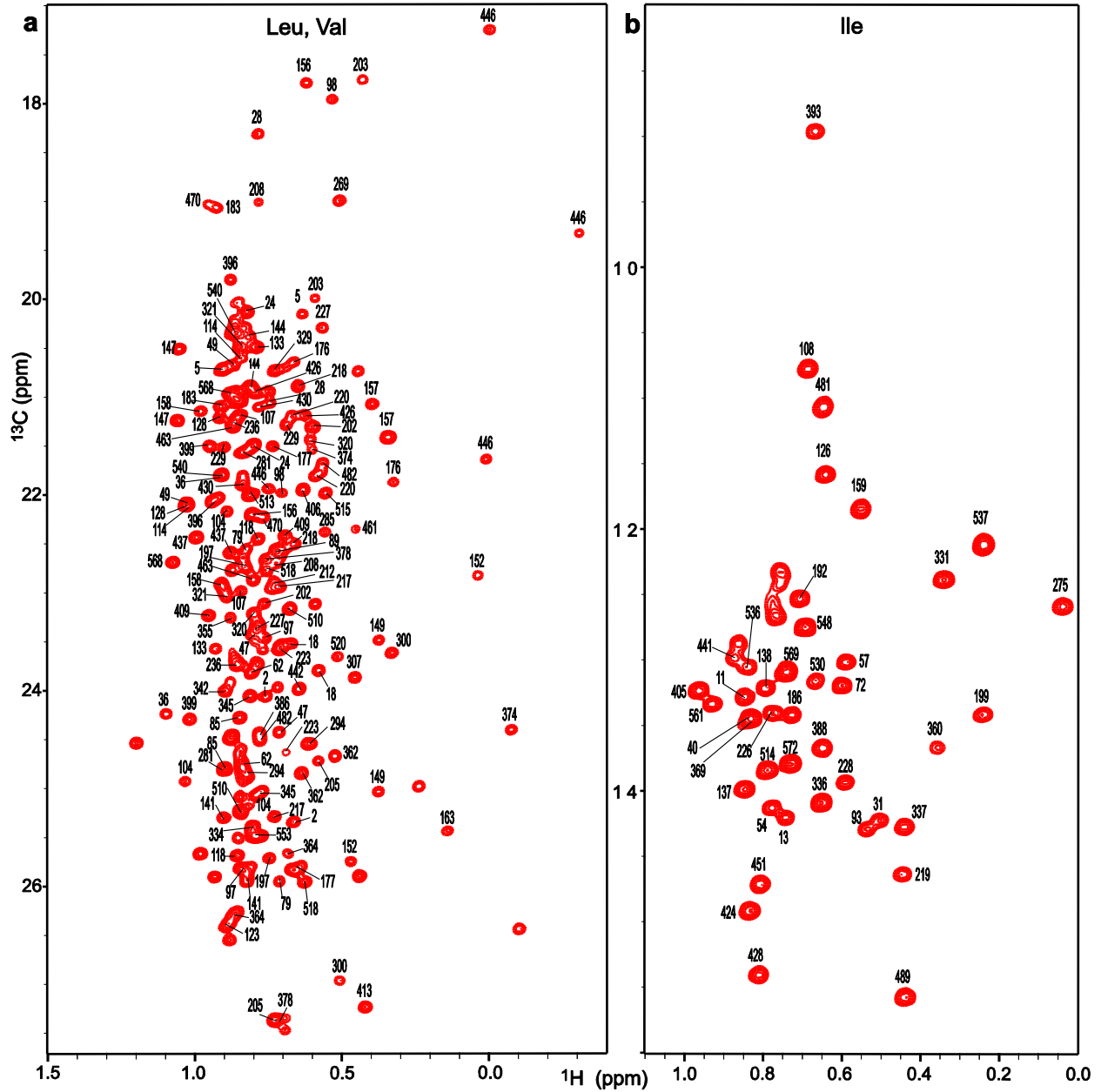
Fig. 4 800 MHz ^{1}H-^{13}C$_{methyl}$ TROSY spectrum of EI from *T. tengcongensis* at pH 7.4 and 70 °C. Val/Leu and Ile portions of the spectrum are shown in (**a**) and (**b**), respectively. Assigned cross-peaks are labeled with the corresponding sequence number. Since methyl groups were not labeled stereospecifically, two cross-peaks are expected for Val (γ1 and γ2 positions) and Leu (δ1 and δ2 positions) residues. No stereospecific assignment was performed in this work
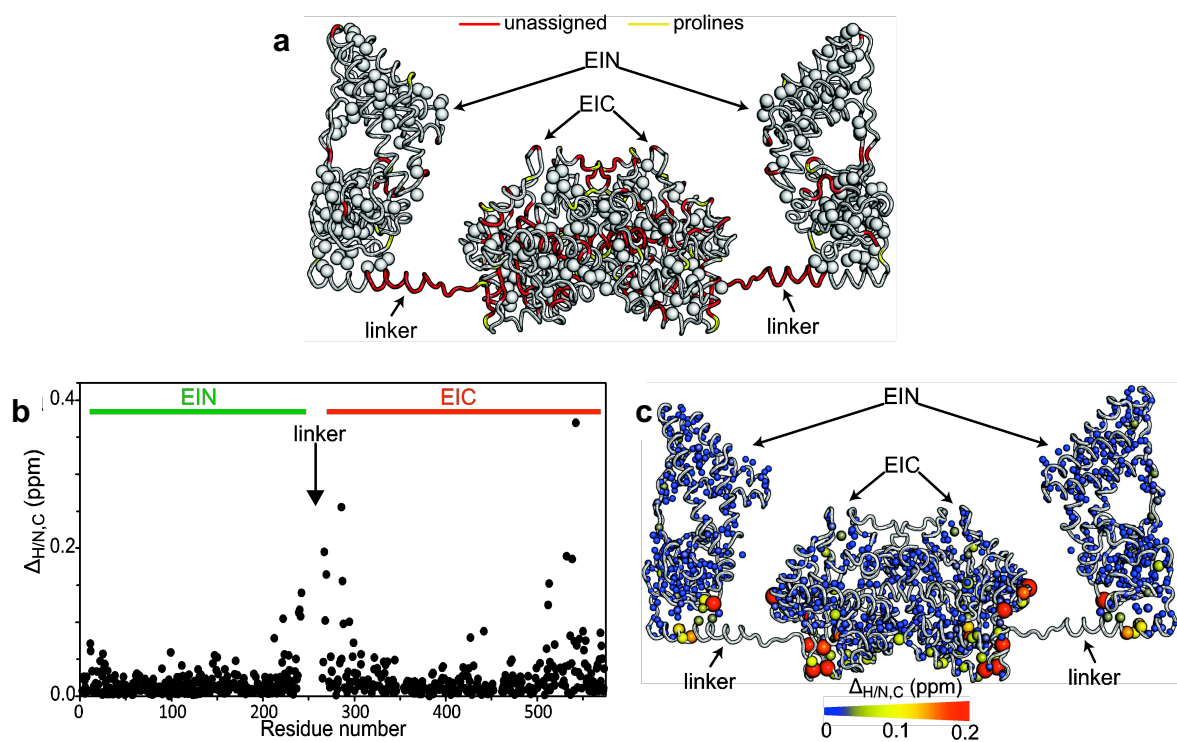
Fig. 5 (**a**) Structure of the *T. tengcongensis* EI dimer created by least-square fit of the experimental structures of the isolated EIC dimer (PDB code: 2BG5) (Navdaeva et al. 2011) and the isolated EIN domain from *T. tengcongensis* (PDB code: 5WOY) (Evangelidis et al. 2018) onto the structure of the full-length *E. coli* EI (PDB code: 2KX9) (Schwieters et al. 2010). Assigned methyl groups are shown as spheres. Residues with unassigned HNCO correlations are shown red. Proline residues are colored yellow. (**b**) Weighted combined $^1H_N/^{15}N$ ($\Delta_{H/N}$) and $^1H/^{13}C$ ($\Delta_{H/C}$) chemical shift perturbations measured by comparing the ^1H_N–^{15}N and ^1H–$^{13}C_{methyl}$ TROSY spectra of the full-length EI to the spectra of the isolated EIN and EIC domains. (**c**) $\Delta_{H/N,C}$ are displayed on the structure of EI as spheres with the relationship between size and color of each sphere and chemical shift perturbation depicted by the color bar

CHAPTER 5.   GENERAL CONCLUSION

The importance of obtaining a comprehensive understanding of the connection between protein structure, conformational dynamics, and activity in regulating protein function is well established, however deconvoluting their role proves non-trivial particularly in large multi-domain proteins. Rational protein design has emerged as a tool to aid in the clarity of this relationship. Additionally, the concept of conformational plasticity and structural heterogeneity have been suggested as justification for the link between structure/dynamics/function in proteins. Here we implement mesophilic/thermophilic hybrids allowing us to selectively target conformational equilibria to exploit this link.

In this research contribution we investigate the structure/dynamics/function relationship in isolated EIC using four constructs (described previously), temperature, and ligand binding to probe the expanded-to-compact equilibrium. We provide evidence that EIC exists as a mixture of states with different catalytic capabilities and degrees of plasticity. It is unclear how the populations of the active and inactive conformations of EIC are affected in regard to temperature and what trend they may follow. This would shed light on whether conformational disorder is responsible for similar catalytic rates observed for eEIC and tEIC at their optimal PTS temperature. This computational work is currently being investigated.

In continuation of this project, implications of intradomain motion determined for isolated EIC in the regulation of interdomain dynamics and activity (phosphoryl transfer) of EI should be investigated. A similar rational followed in this work can be adapted to study the structure/dynamics/function relationship in the mesophilic/thermophilic homologue pair eEI/tEI. Mutants of EI can be engineered implementing EIC domain hybrids (teEIC and etEIC) established and extensively characterized in the work presented. These proposed constructs will

allow modulation of the intradomain motion and the effects on interdomain rearrangement and activity of EI, which in turn will shed light on molecular mechanisms triggering interdomain rearrangements in EI. The structure of full-length EI hybrid constructs can be investigated as outlined in the manuscript reported in Chapter 4. Research strides have already begun on detailing the self-regulatory mechanisms in EI which would provide an unrivaled atomic resolution description for a system exhibiting such complexity.