Universidad de Chile
Facultad de Ciencias
Departamento de Biología

SINGLE MOLECULE STUDIES BY OPTICAL TWEEZERS: FOLDING AND UNFOLDING OF GLUCOKINASE FROM *Thermococcus litoralis*

Christian Andrés Marcelo Wilson Moya

Directores de Tesis:

Dr. Jorge Babul, Universidad de Chile, Santiago.

Dr. Carlos Bustamante, Universidad de California, Berkeley.

Santiago, Chile

Marzo 2011

SINGLE MOLECULE STUDIES BY OPTICAL TWEEZERS: FOLDING AND UNFOLDING OF GLUCOKINASE FROM *Thermococcus litoralis*

Tesis
entregada a la

Universidad De Chile

en cumplimiento parcial de los requisitos

para optar al grado de

Doctor en Ciencias con Mención en Biología Molecular, Celular y Neurociencias

Facultad de Ciencias

por

CHRISTIAN ANDRES MARCELO WILSON MOYA

Marzo 2011

Directores de Tesis:

Dr. Jorge Babul Cattán

Dr. Carlos Bustamante Monteverde

FACULTAD DE CIENCIAS

UNIVERSIDAD DE CHILE

INFORME DE APROBACION

TESIS DE DOCTORADO

Se informa a la Escuela de Postgrado de la Facultad de Ciencias que la Tesis de Doctorado presentada por el candidato

CHRISTIAN ANDRES MARCELO WILSON MOYA

ha sido aprobada por la comisión de Evaluación de la Tesis como requisito para optar al grado de Doctor en Ciencias con mención en Biología Molecular Celular y Neurociencias, en el examen de Defensa Privada de Tesis rendido el día 9 de marzo de 2011.

Directores de Tesis:

Dr. Jorge Babul Cattán .....

Dr. Carlos Bustamante Monteverde .....

Comisión de Evaluación de la Tesis

Dra. Cecilia Vergara Montecinos (Presidenta) .....

Dra. María Cecilia Rojas Garrido .....

Dr. Jaime Eyzaguirre Philippi .....

Dr. Juan Brunet Polanco .....

Dedication.
To my family…

Acknowledgements.

"Our development depends on the professionals that we bring up: individuals with critical thinking that integrate into society and transform it for the good of all. Our universities should have the capacity for research and for generating knowledge, and the Government must make sure that they succeed" (Jorge Babul, Investigation in the Universities, La Tercera 21/01/2010)

Thanks to the help of Professor Tito Ureta (my undergraduate thesis advisor, who has been a mentor throughout my academic career) and also with the help of Professor Octavio Monasterio, I had a short-term research experience at the University of Texas Medical Branch. Once I was there, Rodrigo Maillard, who has been a lab friend ever since, invited Professor Carlos Bustamante to give a lecture. Despite not knowing much English, I found the topic extremely innovative, but I wasn't absolutely sure that what I understood was correct: the manipulation of a single molecule! Something I didn't believe it was possible until now. The next day, during lunch with the South-American "mafia", we had the chance to talk with Professor Bustamante and I was able to ask him about his work, so I learned that it was really possible to work at the single molecule level. Another important subject we discussed was the origins of our Pisco. A couple of months earlier I had been to Valle del Elqui (Chile) with my family, where I learned that the city of Pisco Elqui was founded on 1936, but Professor Bustamante told me that the city Pisco of Perú was founded on 1640, which started an interesting debate about the origins of the alcoholic beverage. Afterwards the debate was settled by Professor Octavio, who said: "*These are two different liquors with the same name*".

Fascinated by Professor Bustamante's lecture about single molecules and his clarity to explain, and with the intent of delving deeper into this subject, María José Gallardo (Cote) and Jaime Andrés Rivas and me, invited him to give a course in Chile. It wasn't easy, because we had to compete for the MECESUP funding with other applicants in the same year. Finally we got the funding. At that time, since I was interested in the biophysics of proteins, I joined the lab of Professor Jorge Babul, who gave me support and help along the way to become a doctor. I always thank him for his wise and spicy advices and everything that he has taught me, and also for our thesis revision marathons! Besides that I admire him for his work in defense of the rights of the scientists. Great scientists work at his lab, like Mauricio Baez with whom I worked at the lab and who helped me with my doctoral thesis; we've had great moments together! Also Profe Victoria Guixé, Ana Preller and Ricardo Cabrera and lots of fellow scientists have helped me, like Rodrigo Guerrero, who taught me the art of the activity assay with formazan. Felipe Merino, Fernando Olate, César Ramírez, Rapa and specially, Andrés Rivas who have guided me with the glucokinase enzyme. I also thank the entire gang from the Laboratory of Biochemistry and Molecular Biology of the Faculty of Sciences, University of Chile, and my friends from other labs like Cote, with whom I have had great times of friendship and scientific exchange, and Carolina Sepúlveda, who helped us with the EPR measurements and has become a very good friend. In this way I

stumbled upon a very interesting problem that could be directly solved by doing studies at the single molecule level. And one thing led to the other and I ended up working with Professor Bustamante at his lab. He has become a wise and generous guide in the path of science for me. I visited Berkeley thanks to a CONICYT scholarship for graduate students and I still work at this place. I had a great time and learned a lot, besides getting to know excellent people who have helped me so much. Like Susan Marqusee, who let me work at her lab, and had taught me about the importance of efficiently conducting the experiments. Steve Smith, Errol Watson, Marta Kopaczynska, Jesse Dill, Phillip Elms, David King, Kambiz Hamadani, Katie Tripp, Katie Hart, Rodrigo Maillard and the entire Bustamante and Marqusee labs. I want to thank all the members of my committee for giving me the best advices and ideas. To my aunt Becky and Juan for help me with my English and to my uncle Raúl Bellemans, a mechanical engineer who inspired me to learn about mechanical processes. And last but not least, to my family, because none of this would have been possible without their enduring and unconditional support, to them I dedicate this thesis. To my parents for having taught me that in life one has to work for vocation, with effort, perseverance, dedication and love. They have supported me always, just like my brothers and sisters. Catita, my beloved daughter, with her crazy little things and ideas has always supported me and understood my passion for science, giving me the space and freedom to stay abroad for long periods of learning and, when we get together, having the best time a father could ever have with his daughter!

Thanks to everyone!

"Is not enough asking our scholars to come back to Chile, we must have places for them in the academy and in the companies in order to achieve the development that we hope for". (Jorge Babul, Changes in Becas Chile, La Tercera 22/06/2010)

Agradecimientos.

"Nuestro desarrollo depende de los profesionales que formemos: individuos críticos, que se incorporen a la sociedad y la transformen para el bien de todos. Nuestras universidades debieran tener la capacidad de investigar y generar conocimiento y el Estado debe velar porque lo logren." (Jorge Babul, Investigación en las universidades, La Tercera 21/01/2010)

Gracias a la ayuda del profesor Tito Ureta (tutor de mi Tesis de pregrado, quien ha sido un mentor hasta hoy) y del profesor Octavio Monasterio, realicé una pasantía en la University of Texas Medical Branch. Una vez allí, Rodrigo Maillard (compañero y amigo de laboratorio hasta el día de hoy) invitó al profesor Carlos Bustamante a dar una charla. Pese a mi inglés, encontré el tema expuesto extremadamente novedoso, pero me quedé con la duda de si lo que había escuchado y entendido era realmente a lo que se refería: ¡lograr la manipulación de una molécula a la vez!; algo que no había pensando que fuese posible aún. Al día siguiente, durante un almuerzo con la "mafia" sudamericana, tuvimos más tiempo de conversar con el profe Carlos y preguntarle del trabajo y comprender que realmente era posible manipular moléculas individuales. Otro tema central del almuerzo fue el origen de nuestro Pisco. Sucede que un par de meses antes había visitado con mi familia el Valle del Elqui, donde aprendí que la ciudad Pisco Elqui había sido fundada en 1936, hecho que relaté a los comensales. El profe Carlos me contó que la ciudad Pisco de Perú fue fundada en 1640, comenzando así una entretenida disputa sobre el origen de la bebida alcohólica. Posteriormente, la disputa fue zanjada por el Profesor Octavio...diciendo, "*¡son 2 licores diferentes con el mismo nombre!*"
Fascinado por el tópico de moléculas individuales y la claridad de exposición del profe Carlos, con la idea de profundizar más, invitamos, junto María José Gallardo (Cote) y Jaime Andrés Rivas (compañeros y amigos del doctorado), al profe a impartir un curso de Moléculas Individuales en Chile. No fue fácil, pues tuvimos que competir por el financiamiento de MECESUP con otros postulantes de ese año. Finalmente el profe ganó la competencia. En ese tiempo, como ya estaba interesado en la biofísica de proteínas, me uní al laboratorio del Profesor Jorge Babul, quien desde un principio me dio su apoyo, ayuda y guía en el camino a ser Doctor. Siempre le agradezco sus sabios y condimentados consejos, todo lo que me ha ido enseñando y nuestras ¡maratónicas revisiones de la Tesis! Además, lo admiro por su labor en la defensa de los derechos de los científicos. En su laboratorio trabajan grandes científicos, por ejemplo Mauricio Baez, con quien comencé a trabajar directamente en el laboratorio y me ayudó en el desarrollo de mi tesis de doctorado: ¡hemos tenido grandes momentos juntos! Además de los profes Victoria Guixé, Ana Preller y Ricardo Cabrera, muchos compañeros me han ayudado, como Rodrigo Guerrero, enseñándome el arte de medición de actividad enzimática utilizando formazán, Felipe Merino, Fernando Olate, César Ramírez, Rapa y en especial, Andrés Rivas, que me ha guiado con la enzima glucoquinasa. Agradezco a todos mis compañeros del Laboratorio de Bioquímica y Biología Molecular de la Facultad de Ciencias de la Universidad de Chile y a mis amigos de otros laboratorios

como la Cote, con quien he pasado muy buenos momentos de intercambio científico y de amistad y Carolina Sepúlveda, quien nos ayudó en un comienzo con las medidas de EPR y se ha convertido en una muy buena amiga. Así, encontré un problema interesante de Tesis, cuya mejor y más directa forma de resolver era por medio del estudio de moléculas individuales, y el profe Carlos me invitó a su laboratorio a trabajar, guiándome en el camino de la ciencia hasta hoy. Visité Berkeley gracias a una beca de pasantías de CONICYT y aún trabajo en ese lugar; lo he pasado muy bien y he aprendido montones, además de conocer a gente que me ha ayudado mucho. Por ejemplo Susan Marqusee, quien me ha permitido trabajar en su laboratorio, me ha guiado y enseñado la importancia del desarrollo de experimentos en forma eficiente y concreta. Steve Smith, Errol Watson, Marta Kopaczynska, Jesse Dill, Phillip Elms David King, Kambiz Hamadani, Katie Tripp, Katie Hart, Rodrigo Maillard y todos los miembros del laboratorio del profe Bustamante y la Profe Marqusee. Quisiera agradecer a los miembros de mi comisión, por aportarme muy buenos consejos e ideas. A mi tía Becky y Juan por su apoyo en el inglés. A mi tío Raúl Bellemans quien como ingeniero mecánico, me guió por el gusto de entender los procesos mecánicos. Además nada de esto hubiera sido posible sin el apoyo incondicional de mi familia, a quienes ha sido dedicada esta Tesis. A mis padres por haberme enseñado que las cosas en la vida se hacen por vocación, con empeño, esfuerzo, perseverancia, dedicación, y mucho amor. Ellos, junto a mis hermanos, me han apoyado siempre. La Catita, mi hija, que con sus *locuritas* siempre me ha apoyado y comprendido, permitiendo que pueda estar en el extranjero aprendiendo y ¡aprovechando al máximo nuestro tiempo juntos en cualquier parte del mundo!

¡Muchas gracias a todos!

"No basta con pedirles a los becarios que regresen, debemos tener lugar para ellos en la academia y en las empresas para lograr el desarrollo que pretendemos." Jorge Babul, Cambios en Becas Chile, La Tercera 22/06/2010)

Participation in events and recognitions in the period of the Thesis

Invited Lectures.

Departamental seminar Universidad Andrés Bello. Estudio de Moléculas Individuales mediante Pinzas Opticas: El caso de la Glucoquinasa. Host: Dr. Jaime Eyzaguirre. January 13, 2010, Chile

Departamental seminar Universidad de Chile. Estudio de Moléculas Individuales mediante Pinzas Opticas: el caso de la Glucoquinasa. Host: Dr. Jorge Babul. March 24, 2010, Chile

Seminar Universidad de Buenos Aires. Estudio de Moléculas Individuales mediante Pinzas Opticas: el caso de la Glucoquinasa. Host: Dr. Sergio Kaufman. October 7, 2010, Argentina

Meeting Presentations

3rd Latin American Protein Society. Mechanical stability studies of glucokinase from *Thermococcus litoralis* at the single molecule level. Wilson, C.A.M, Smith, S.B., Babul, J., Marqusee, S., and Bustamante, C. Salta, October 13-16, 2010, Argentina

3rd Latin American Protein Society. Single-Molecule Studies on Allosteric Regulation of cAMP Receptor Protein (CRP) from *Escherichia coli.* Maillard, R.A., Wilson, C.A.M, Barton, S., Lee, J.C. and Bustamante, C. Salta, October 13-16, 2010, Argentina

XXXIII Annual Meeting of the Chilean Society of Biochemistry and Molecular Biology. Single Molecules Studies by optical tweezers: unfolding of glucokinase from *Thermoccocus litoralis.* Wilson, C.A.M, Smith, S.B., Babul, J., Marqusee, S., and Bustamante, C. Chillan September 28-October 1, 2010, Chile.

23rd Symposium of The Protein Society. Extended monomeric intermediate upon cold denaturation of the dimeric phophofructokinase-2 from *E. coli.* Wilson, C., Baez, M. Babul, J. Boston, July 24-29, 2009, USA

23rd Symposium of The Protein Society. Kinetic and EPR studies of the role of divalent metal cations in *E. coli* Pfk-2. Rivas-Pardo, Wilson, C.A.M., Caniuguir, A., Guixe, V. Boston, July 24-29, 2009, USA

XXXI Annual Meeting of the Chilean Society of Biochemistry and Molecular Biology. Thermodynamic analysis of the thermal unfolding of phosphofructokinase-2 at low and

high temperatures. Baez, M., Wilson, C.A., and Babul, J. Chillán, September 23-26, 2008, Chile

XXXI Annual Meeting of the Chilean Society of Biochemistry and Molecular Biology. Binding of divalent metals in phosphofructokinase-2 from *E. coli*. Kinetic and EPR studies. Rivas-Pardo, J.A., Wilson, C.A.M., Caniuguir, A., Guixé, V. Chillán, September 23-26, 2008, Chile

XXV Biochemistry students National Asociation Congress (ANEB). Aplicación del análisis del control metabólico al estudio de la regulación de la síntesis de glicógeno *in vivo* en oocitos de anfibio. Wilson, C.A.M., Preller, A., Quiroga, D. and Ureta, T. Concepción, August 6-9, 2008, Chile

22nd Annual Symposium of the Protein Society. Unfolded monomeric intermediates upon cold and heat denaturation of the dimeric phosphofructokinase-2 from *E. coli*. Babul, J., Baez, M., and Wilson, C.A.M. San Diego, July 19-23, 2008, USA

33rd FEBS Congress and 11th IUBMB Conference. *In vivo* flux control coefficient of glycogen synthase in frog oocytes. Ureta, T. Preller, A. and Wilson, C. Atenas, 28 June – 3 July, 2008, Greece.

XXXVII Annual Meeting of the Brazilian Society of Biochemistry and Molecular Biology - SBBq and XI Congress of the PABMB. Cold denaturation of phosphofructokinase-2 of *E. coli*. Wilson, C.A.M., Baez, M. and Babul, J. Agua de Lindoia, May 17-20, 2008, Brazil.

Awards

Best Lecturer at the the Chilean Society XXXIII Annual Meeting of the Biochemistry and Molecular Biology Society of Chile 2010. Single Molecules Studies by optical tweezers: unfolding of glucokinase from *Thermoccocus litoralis*. Wilson, C.A.M., Smith, S.B., Babul, J., Marqusee, S. and Bustamante, C.

Publications during the period of the Thesis

Divalent metal cation requirements of phosphofructokinase-2 from *E. coli*. Evidence for a high affinity binding site for Mn^{2+}. Rivas-Pardo, J.A., Caniuguir, A., Wilson, C.A.M., Babul. J., Guixé, V. Arch. Biochem. Biophys., 505, 60-66, 2011.

Folding kinetic pathway of phosphofructokinase-2 from *Escherichia coli*: a homodimeric enzyme with a complex domain organization. Baez, M., Wilson, C.A.M., Babul, J. Submitted to FEBS Letters.

Sources of funding.

FONDECYT 1090336, Dr. Jorge Babul

NIH, Dr. Susan Marqusee

HHMI, Dr. Carlos Bustamante

Scholarship CONICYT (Comisión Nacional de Investigación Científica y Tecnológica)

Assistantship abroad by CONICYT and MECESUP at the University of California, Berkeley under Dr. Carlos Bustamante's tutorship, co-adviser of the Doctorate Thesis, 2008.

Mini Grant (Beca de Apoyo a Tesis Doctoral, CONICYT) 2009-2010

TABLE OF CONTENTS

LIST OF TABLÆS

LIST OF FIGURES

LIST OF ABBREVIATIONS

aa	Amino acid
ADN	Ácido desoxiribonucleico
ADP	Adenosine 5' diphosphate
AFM	Atomic Force Microcopy
ATP	Adenosine 5' triphosphate
B-DNA	B form DNA
BN	Blue native
bp	Base pairs
BSA	Bovine serum albumin
CCD	Charge-coupled device
CD	Circular Dichroism
ΔG	Free energy of unfolding
DNA	Deoxyribonucleic acid
DTDP	2,2'-Dithiodipyridine
DTT	Dithiotreitrol
Fx	Force applied perpendicular to the pipette
Fy	Force applied in the direction of the pipette
Fz	Axial force
G6PDH	Glucose-6-phosphate dehydrogenase
HHMI	Howard Hughes Medical Institute

IPTG	Isopropyl β-D-1-thiogalactopyranoside
kDa	kilo Dalton
Keq	Equilibrium constant
LB	Luria-Bertani
NAD^+	Nicotinamide
NADH	Nicotinamide reduced form
NBT	Nitro blue tetrazolium
nm	nanometer
PAGE	Polyacrylamide gel electrophoresis
PMS	Phenazine methosulfate
PCR	Polymerase chain reaction
pN	picoNewton
PSD	Position sensitive detector
SAXS	Small-angle X-ray scattering
s-DNA	Stretched form DNA
SDS	Sodium dodecyl sulfate
tlGK	Glucokinase from *Thermococcus litoralis*
ULVWF	Ultra large form of Von Willebrand factor
UTP	Uridine 5' triphosphate
WLC	Worm like chain
$x_u^{\mp}$	Unfolding distance to the transition state
$x_f^{\mp}$	Folding distance to the transition state

ABSTRACT

Single-molecule manipulation has increasingly become a useful methodology for the study of macromolecular dynamics including unfolding-refolding transitions. In this study we use the optical tweezers to investigate the effect of substrate on the mechanical stability of glucokinase from the hiperthermophilic archeon *Thermococcus litoralis*. To this end, we derivatize the enzyme with DNA handles, via reactive cysteines and attach them to polystyrene beads. The enzyme with different handle attachments was separated using a native gel electrophoresis assay, and the activity of the enzyme in each gel band was monitored *in situ* by coupling the formation of glucose-6-phosphate to the formation of formazan and the DNA-protein-DNA construct was purified. Different constructs were pulled to investigate the effect of the ADP substrate on the stability of the protein. At 28 pN, the construct S4C/T57C showed hopping between two conformations with a difference in extension of 5.3 nm. For the unfolding reaction, the distances to the transition state were $x^{\ddagger} = 3.09 \pm 0.35$ nm in the absence and 2.81 ± 0.62 nm in the presence of ADP, respectively; the corresponding values during refolding were 2.24 ± 0.24 nm and 2.46 ± 0.43 respectively. The rate of unfolding extrapolated to zero force was an order of magnitude smaller in the presence than in absence of ADP, with no changes in the extrapolated refolding rates. These results suggest that the folded state of the protein is stabilized by the substrate with no effect on the distance to the transition state for the unfolding reaction.

EXTENDED ABSTRACT

The effect of force on protein structure and associated changes of protein function is a subject of current intensive research. Mechanical forces are generated inside the cell during such diverse molecular processes as transcription, replication, translation, chromosomal segregation, protein unfolding, translocation of proteins across membranes, cell locomotion. Recent technological advances now allow the application and measurement of forces on biomolecules with extreme precision. In particular, the so-called "analytical optical and magnetic tweezers" instruments can manipulate single molecules, such as proteins and nucleic acids, while measuring their internal stress forces generated in the course of biological processes.

Single molecule level studies, called *in singulo* studies, have become very relevant during the last years, because of their advantages in comparison to molecular assembly studies (from many molecules, also called *in multiplo* studies) when it comes to obtain specific information about biological phenomena. With *in multiplo* studies it is possible to get robust population measurements (average measurements), which can be interpreted given the premise that populations should be homogeneous. On the other hand, *in singulo* studies follow the behavior of an individual molecule in time, thus making it possible to obtain not just the average behavior of many molecules, but rather the whole distribution and individual behaviors of a population that may not be homogeneous. *In singulo* studies have also another important characteristic: it is a very direct approach. When using the *in multiplo* studies, the behavior of the molecules must

be inferred instead of directly gauged, and therefore a model of the system must be developed.

In this study we use optical tweezers to investigate the effect of substrate on the mechanical stability of glucokinase from the hyperthermophilic archeon *Thermococcus litoralis*, which lives at temperatures over 80 °C.

The optical tweezers machine catches particles in a potential well formed by near-infrared light. This near-infrared laser beam is tightly focused by a high-numerical aperture microscope objective to create the large spatial gradient in light intensity necessary to form a stable trap. In an optical trap, the trapped particle can experience a force if it is displaced from the center of the trap. When the optical trap behaves as a Hookean spring, this force (F) can be calculated using the spring or Hook's constant (k) of the trap and the displacement (Δx) so that: $F = k \Delta x$.

Glucokinase is a monomeric enzyme that has several properties that make it a good subject for optical tweezers studies. It is a monomer, which allows one to study the single protein without complications due to dissociation. It has no cysteines residues. This property is important, since the approach requires binding the protein molecule to DNA fragments that are used as "molecular handles" to attach and manipulate it by means of polystyrene beads. These DNA handles are attached by means of cysteines to the protein. The absence of cysteines allows us to do cysteine engineering in the appropriate places in the protein without affecting the protein's structure.

Currently there are few studies that show the effect of substrate on the

mechanical properties of enzymes and none of these studies have been conducted in thermodynamic equilibrium conditions. To this end, we derivatize the enzyme with DNA handles, via reactive cysteines and attached them to polystyrene beads. The sites where the cysteines were inserted are zones that do not have a high degree of amino acid conservation and are relatively exposed to the solvent so that the DNA handles can react. Also, in these zones the binding sites are perpendicular to the hinge of the protein.

Due to the re-entrant topology (or discontinuous subdomains) of glucokinase, besides targeting the N- and C-terminal ends of the protein, other pulling axes perpendicular to the hinge of the protein should be looked for in order to study the effect of the substrate in the mechanical stability of the enzyme under equilibrium conditions.

After the DNA handle attachment reaction, five species may be obtained: DNA alone, DNA-DNA, DNA-protein, DNA-protein-DNA and protein alone. These species were separated using a native gel electrophoresis assay and the activity of the enzyme in each gel band was monitored *in situ* by coupling the formation of glucose-6-phosphate to the formation of formazan. Thus, it was shown, for the first time that proteins are active even when attached to two 558 bp DNA handles.

The DNA-protein-DNA construct was purified by means of electroelution. Before purification, the sample contained only a small fraction of the desired construct (~5% DNA-protein-DNA). The large amount of DNA-DNA in the sample would bind to the beads and interfere with the fishing step in the optical tweezers. Purification by electroelution resulted in a sample preparation with > 70% of the desired product.

Similar results were obtained with all the glucokinase mutants studied. This procedure significantly improved the fishing step in the optical tweezers.

Different constructs were pulled to investigate the effect of the ADP substrate on the stability of the protein. When the enzyme was pulled from the N and C-terminal ends (S4C/R467C mutant), and was unfolded under force, the presence of intermediates and hysteresis were observed. The total theoretically expected length for global unfolding was never obtained, indicating that we could be starting from some intermediate species that the enzyme had under force.

The enzyme mutated at positions perpendicular to the hinge (T57C/T418C) showed three intermediate species when pulled: one with 80% of the total structure, a second corresponding to 40% (that seems to be same as the first but broken into two species) and a third one corresponding to 20% of the total structure. This mutant shows the total expected unfolding when adjusted to the *worm like* chain model. When the experiment was done in the presence of ADP, the unfolded intermediate with 20% of the total structure was observed much less often.

At 28 pN, the S4C/T57C construct showed hopping (exploring of two or more structures at the same force) between two conformations with a difference in extension of around 5.5 nm. The ΔG of unfolding in the absence and in the presence of ADP was 11.2 and 12.0 kcal/mol, respectively. A very high value considering that around 10% of the total structure is being pulled between the DNA handles, which suggests coupling folding of this region to the rest of the structure because of the re-entrant topology of the protein. For the unfolding reaction, the distances to the transition state were $x^{\ddagger}$ =

3.09 ± 0.35 nm in the absence and 2.81 ± 0.62 nm in the presence of ADP, respectively. The corresponding values during refolding were 2.24 ± 0.24 nm and 2.46 ± 0.43 respectively. The rate of the unfolding, extrapolated to zero force, was an order of magnitude smaller in the presence than in absence of ADP ($5x10^{-8}$ and $3x10^{-7}$ s^{-1}, respectively), with no changes in the extrapolated refolding rates. These results suggest that the folded state of the protein is stabilized by the substrate with no effect on the distance to the transition state for the unfolding reaction.

RESUMEN EXTENDIDO

El efecto de la fuerza sobre la estructura de proteínas y los cambios en su función asociados a la fuerza representan un área de investigación de gran interés en los últimos años. Dentro de la célula se generan fuerzas mecánicas en procesos moleculares tan diversos como transcripción, replicación, traducción, segregación cromosomal, desplegamiento de proteínas, translocación de proteínas intermembranas y movimiento celular. Los recientes avances tecnológicos han permitido la aplicación y medición de fuerzas sobre biomoléculas con extrema precisión. En particular, las técnicas de manipulación analítica tan ampliamente conocidas como pinzas magnéticas y pinzas ópticas, permiten la manipulación de moléculas individuales, como proteínas y ácidos nucleicos, además de obtener información acerca de la fuerza interna de estiramiento generada en el curso de procesos biológicos.

Los estudios realizados a nivel de molécula individual, conocidos con el término *in singulo*, han cobrado gran relevancia en los últimos años, principalmente por sus ventajas comparativas sobre aquellos estudios realizados en asambleas estadísticas moleculares, también llamados *in multiplo,* al momento de obtener información específica sobre algún fenómeno biológico. Los estudios *in multiplo* permiten obtener medidas de población promedio, consideradas robustas, partiendo de la premisa que se trabaja sobre una población homogénea. Por otro lado, los estudios *in singulo* cuentan con características relevantes, que corresponden a un enfoque directo, pues cuando se utilizan estudios *in multiplo*, el comportamiento de las moléculas debe ser inferido en

vez de ser directamente medido. Además es necesario desarrollar previamente un modelo del sistema en estudio.

En esta tesis utilizamos la técnica de pinzas ópticas para investigar el efecto que ejerce el sustrato sobre la estabilidad mecánica de la enzima glucoquinasa de la arquea hipertermófila *Thermococcus litoralis*, que crece a temperaturas superiores a 80 °C.

La instrumentación del equipo de pinzas ópticas se basa en el atrapamiento de partículas bajo un pozo potencial homogéneo formado por luz compuesta de longitudes de onda en el infrarrojo cercano. El haz de luz está dirigido por un objetivo de microscopio de índice de apertura numérico alto, que permite la creación de un gradiente espacial de intensidades de luz, necesario para formar una trampa estable. En una trampa óptica, la partícula atrapada experimenta una fuerza si es desplazada del centro de la trampa. Cuando la trampa se comporta como un resorte de Hook, esta fuerza (F) se puede calcular usando la constante de Hook (k) y el desplazamiento (Δx) de la partícula, $F=k\,\Delta x$.

La glucoquinasa es una enzima monomérica que posee varias cualidades para ser considerada un buen modelo de estudio en las pinzas ópticas. El hecho que la proteína sea monomérica, es ventajoso, pues cuando se trabaja sobre moléculas individuales existe una alta dilución de las especies proteicas lo que promueve la disociación de las subunidades. Además, la ausencia de cisteínas en esta glucoquinasa permite, a través de manipulación genética, agregar cisteínas en regiones apropiadas de la proteína sin provocar alteraciones estructurales. Estas cisteínas pueden ser entonces usadas para insertar mangas de ADN para manipular mecánicamente la proteína.

Hasta el momento, no existen muchos estudios que muestren el efecto del sustrato en las propiedades mecánicas de las proteínas y ninguno de ellos se ha desarrollado bajo condiciones de equilibrio termodinámico. Para este fin, derivatizamos la glucoquinasa de *T. litoralis* a través de cisteínas con mangas de ADN, que a su vez se unen a bolitas de poliestireno. Los sitios en que fueron insertadas las cisteínas presentan un bajo grado de conservación de aminoácidos y se encuentran relativamente expuestos al solvente permitiendo la reactividad del ADN sobre ellos. Los sitios de unión a las mangas fueron además elegidos por estar en zonas que se encuentran perpendiculares a la bisagra de la proteína.

Debido a la topología re-entrante (o subdominios discontinuos) de la glucoquinasa, además de estirar de la proteína de los extremos N y C-terminal, es conveniente imponer otros ejes de estiramiento, perpendiculares a la bisagra de la proteína, los que deberían buscarse con el propósito de estudiar el efecto de los sustratos en la estabilidad mecánica de la enzima bajo condiciones de equilibrio.

El tratamiento de la molécula con las mangas de ADN origina cinco especies en solución: ADN sólo ADN-ADN, ADN-proteína-ADN, ADN-proteína, ADN y proteína sola. Las especies generadas se separaron por medio de un gel nativo de electroforesis y la actividad de la enzima en cada banda del gel fue monitoreada *in situ* por acoplamiento de la formación de glucosa-6-fosfato a la formación de formazán. De esta manera se demostró, por primera vez, que la proteína es activa con las mangas de ADN de 558 bp cada una.

El constructo ADN-proteína-ADN fue purificado mediante electroelución. La

muestra previa a la electroelución contenía sólo un 5% de ADN-proteína-ADN y cantidades sustanciales de ADN-ADN que interfieren con el proceso de atrapamiento por las pinzas ópticas. Después de la electrolución, el constructo ADN-proteína-ADN fue enriquecido en un 70%. Resultados similares fueron obtenidos con todas las mutantes de la glucoquinasa utilizadas. Este procedimiento mejoró notablemente el proceso de atrapamiento por las pinzas.

Se construyeron diferentes mutantes de cisteína para investigar el efecto del ADP sobre la estabilidad de la proteína. Cuando la enzima fue estirada por sus extremos N y C (mutante S4C/R467C) desplegándola bajo fuerza, se observó la presencia de histéresis y la aparición de numerosos intermediarios de desplegamiento. La longitud teórica esperada (por el modelo de cadena vermiforme) no se observó, lo que podría indicar que el proceso comienza desde alguna especie intermediaria parcialmente desplegada que la enzima experimenta bajo fuerza.

La enzima mutada en posiciones perpendiculares a la bisagra (T57C/T418C) reveló la existencia de tres especies intermediarias: una con 80% de la estructura total, una segunda correspondiente al 40% (que parece ser idéntica a la primera especie pero dividida en dos sub-intermediarios) y una tercera correspondiente a 20% de la estructura total. Esta mutante mostró un desplegamiento total cuando se ajustó a un modelo de cadena vermiforme (WLC). Cuando el experimento fue llevado a cabo en presencia de ADP, el intermediario desplegado en un 20% se observó en mucha menor proporción.

A 28 pN el constructo S4C/T57C, mostró un salto de un estado a otro (exploración de dos o más estructuras a la misma fuerza) entre dos conformaciones, con

una diferencia en extensión de alrededor de 5,5 nm. El ΔG de desplegamiento en ausencia y presencia de ADP fue de 11,2 y 12,0 kcal/mol, respectivamente. Estos valores fueron altos considerando que alrededor de un 10% del total de la estructura de la proteína forma parte de los sitios de estiramiento entre las mangas de ADN, lo que sugiere un acoplamiento energético del desplazamiento entre la región manipulada y el resto de la proteína probablemente debido a la estructura re-entrante de la proteína. Para la reacción de desplegamiento, la distancias al estado de transición fueron $x^{\ddagger} = 3,09 \pm 0,35$ nm en ausencia y 2,81 ± 0,62 nm en presencia de ADP. Los valores correspondientes al proceso de replegamiento fueron de 2,24 ± 0,24 nm y 2,46 ± 0,43 respectivamente. La velocidad de desplegamiento fue extrapolada a cero fuerza, obteniéndose un valor en presencia de ADP de un orden de magnitud menor que en ausencia de ADP ($5x10^{-8}$ y $3x10^{-7}s^{-1}$, respectivamente), sin variaciones en las velocidades de plegamiento obtenidas por extrapolación. Los resultados sugieren que el estado plegado de la proteína es estabilizado por el sustrato sin ejercer efectos en la distancia entre el estado nativo y el estado de transición de la reacción de desplegamiento.

INTRODUCTION

The importance of force in biological studies

Mechanical processes are a key component of many biological events. The relevance of the role played by mechanical force is evident in such diverse phenomena as the transportation by molecular motors, the segregation of chromosomes, the formation and liberation of vesicles, the packing of a virus. But force also plays an important role in more subtle processes such as signal transmission and protein mechanosensation via protein unfolding (Bustamante et al., 2004; Cecconi et al., 2005; Zhang et al., 2009).

To better understand the role of mechanical force as a mechanosensor, there is a recent example in the activating mechanism of the coagulation factor ULVWF (ultra large form of Von Willebrand factor). It has been shown that as one sector of this protein unfolds mechanically, an amino acid sequence gets exposed thus allowing a protease to cut it, which then activates ULVWF (Zhang et al., 2009). These studies show the force generated by blood flow is necessary for the unfolding of the A2 domain of the ULVWF protein. As a result of this force, the cutting site for the metalloprotease ADAMTS13 is exposed, and the hemostatic cascade is initiated. With the help of the optical tweezers technique, it was possible to directly observe and measure the force needed to unfold this domain.

Force may be used as another variable to study chemical or biological reactions. Usually reactions are studied by means of manipulating several parameters such as temperature, chemical agents, and pressure. All of these variables may change the

thermodynamic and kinetic parameters of the reactions. The application of force can alter those parameters too.

The advantage of studying the reactions with a focus on force is that the reaction coordinate is an easily quantifiable physical parameter, for example, the end-to-end distance of a molecule as it is being stretched (Tinoco, 2004; Bustamante et al., 2004). In thermodynamics, the energy change of a system is defined by the heat exchange (q) and the work done (w) by the system. Then we can write that the internal energy of a system (E) is given by:

$dE = dq + dw$

This expression can be re-written in terms of entropy (S), pressure (P)- volume (V) and mechanical work, as:

$dE = (TdS) + (-PdV + \int Fdx)$

A more practical approach is to rewrite this equation using temperature and pressure as independent variables with the Gibbs free energy:

$dG = -SdT + VdT + VdP + Fdx$

At constant temperature and pressure, the work required to extend a system by Δx is:

$W = \int Fdx$

By measuring the applied force we can determine the thermodynamic and kinetic parameters of the studied reactions. Bell (1978) was the first to describe -in a phenomenological way- the force dependence of the kinetics of a reaction. Currently, new transitions state theories stemming from the application of Kramers' (1940) theory have emerged (Dudko et al., 2008).

Many studies have been conducted in order to unveil the mechanical properties of biological materials. These studies include the unfolding of biopolymers like DNA, RNA, and proteins, molecular motors, analysis de DNA-polymerase interactions and mechanosensors (Bustamante et al., 2011).

In singulo versus in multiplo studies

Single molecule level studies, called *in singulo* studies (Bustamante, 2008), have become very relevant during the recent years, because of their advantages compared to standard ensemble-based studies (from many molecules, also called *in multiplo* studies) when it comes to obtain specific information about biological phenomena. *In multiplo* studies allow to obtain population measures (average measures) and have been robust in this regard, starting with the premise that populations should be homogeneous. On the other hand, *in singulo* studies follow the behavior of an individual molecule in time, thus making it possible to obtain not just the average behavior of many molecules, but rather the whole distribution and individual behaviors of a population that may not be homogeneous. Besides, *in singulo* methods also permit the examination of molecules which behavior falls very far from the average population behavior. *In singulo* studies have also another important characteristic: it is a very direct approach. When using the *in multiplo* studies, the behavior of the molecules must be inferred instead of directly gauged, and therefore a model of the system must be developed, for example if in the solution we have a monodal distribution or higher number of configurations. In contrast, the outcome of the *in singulo* studies is a direct measurement of the protein behavior.

The data obtained from these experiments are necessarily noisy because they reflect the stochastic behavior of the very same molecules. Therefore, the fluctuations dominate the experimental signals (Bustamante, 2008). Besides, many cellular processes are dependent on the participation of only a small sub-set of molecules and, thus, *in multiplo* studies, that generate average measures of a huge number of molecules, do not portray a reliable picture of the natural noise and fluctuations that accompany these cellular processes. For example, processes such as the replication of the main chain of DNA are carried by one single molecule. Besides, considering that one *E. coli* bacterium has a volume of ~ 1 μm^3, we could easily show that one single molecule inside that cell (i.e. a protein of 50 kDa), is already at a concentration of approximately 1.6 nM, and therefore it is capable of performing most, if not all, the second order reaction processes in which it participates.

Another advantage *in singulo* studies have versus the *in multiplo* ones is that using the fluctuations theorems derived by Jarzynski and Crooks (Jarzynski, 1997; Crooks, 1999) we can correlate values obtained in non-equilibrium conditions with values obtained in equilibrium conditions, which has been regarded as fundamental advance in the field of thermodynamics (Collin et al., 2005).

Another important advantage of *in singulo* studies, particularly those that involve single molecule manipulations, is that it is possible to quantify the force that certain cellular processes generate.

Finally, *in singulo* experiments need much less quantity of experimental materials than *in multiplo* studies. Frequently, certain materials such as proteins or some

DNAs are difficult to obtain in large quantities, but, advantageously, there is no need for large quantities in single molecule experiments.

An example that shows the power of single molecule manipulation methods is the study of the packing of the virus Phi29 (Moffitt et al., 2009). The Phi29 virus DNA-packaging motor uses chemical energy and transforms it in mechanical energy to generate DNA internal pressures of about 60 atm. This pressure allows the virus to inject the DNA into the cells. Using a single molecule approach, the pauses that are generated in the motor operation may be observed and it was establish that the binding of ATP to the motor is highly cooperative, leading to a complete segregation between ATP binding and translocation of the motor, requiring all the subunits of the motor to have the nucleotide bound in order to pack the DNA. This information would be unattainable *in multiplo* due to the temporal and ensemble average of these studies.

Measurement of force at the single molecule level

At present some techniques have been developed, such as atomic force microscopy, optical and magnetic tweezers, that allow to directly measure the mechanical forces that certain biological processes produce (for a complete review about individual molecule manipulation techniques see Deniz et al., 2008 and Bustamante et al., 2000). There are other techniques that also allow measurement of force at the single molecule level, as micro-needle manipulation, biomembrane force probe and flow-induced stretching. Currently, the first 3 ones are the most used and also the ones that have demonstrated better versatility to measure force at the single

molecule level. Magnetic tweezers are based on the use of a magnetic field to act on a ferromagnetic object, typically a magnetic bead. Force on the bead can be generated in this way. These instruments use a CCD camera to follow the movement of the bead. Out of the three techniques, magnetic tweezers have the worst spatial resolution (5-10 nm), but they can reach very low forces (0.01 pN). Among their advantages is worth mentioning that problems like radiation heating of the sample or photodamage, are absent. Besides, magnetic tweezers allow one to apply torque on the bead by rotating the external magnets, and have the advantage that they generate very stable force fields that can be applied simultaneously to many individual molecules in the field of view of the microscope (Figure 1).

Atomic force microscopy has been the most used technique for protein study at an individual molecule level until now. It was first used for visualizing biological material through the movement of a cantilever. The cantilever can also be used for the application of forces by moving it away and toward the sample, thus helping to quantify the mechanical properties of certain materials including individual molecules. It has many qualities, namely, a broad force range can be applied using the cantilever (10-10,000 pN) and presents very good spatial resolution (0.5-1 nm) (Figure 1).

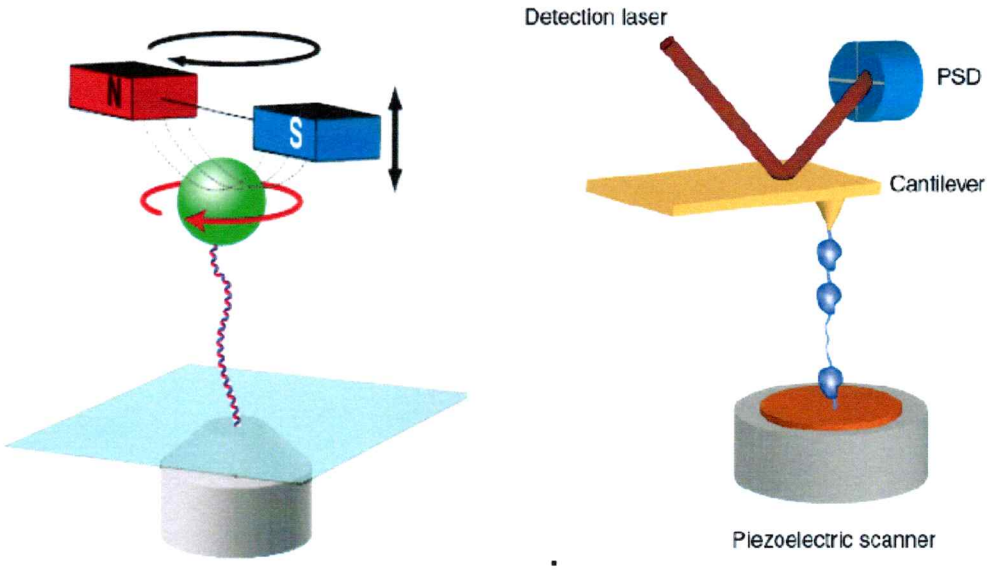


Figure 1. Schematic representation of Magnetic Tweezers and Atomic Force Microscope. Left: Magnetic Tweezers. A molecule of DNA is observed attached to a glass surface and to a magnetic bead. Through a magnetic field applied by some rare earth magnets, force can be applied to the molecule for stretching it, besides rotating it. Right: Atomic Force Microscope. A polyprotein is attached to the surface of a piezoelectric scanner and to a cantilever. The piezoelectric scanner allows a perpendicular movement to the protein, generating force. The force is measured through the reflection of a laser from a cantilever to a PSD detector. Figure taken from Neuman and Nagy (2008).

Optical tweezers

Another important technique for single molecule manipulation is optical trapping (mentioned above) which allows us to apply force and manipulate individual molecules such as proteins, and to measure the forces developed during the course of certain biological processes. This technique, whereby microscopic particles are captured in a potential well formed by focused light, was developed by Ashkin et al. (1970) at Bell Laboratories (see also Ashkin et al., 1986). They found that by focusing a laser beam through a microscope objective, any particle of high refractive index, such as glass or plastic, can be caught by a light beam and held at the focus. The catching mechanism is due to the balance of two types of optical forces: scattering forces, which push the object in the direction of the light beam, and gradient forces, which attract the particle to the focal point. When the gradient forces exceed the scattering forces, particle is stably trapped. Generally, a near-infrared laser beam is used, whose wavelength is not absorbed and thus remains innocuous to most biological materials.

Such near-infrared laser beam is tightly focused by a high-numerical-aperture microscope objective lens to create the large spatial gradient in light intensity necessary to form a stable trap. To a first approximation, the trap behaves as a Hookean spring; forces can be generated on an object when it is displaced from the center of the trap and the force can be calculated from a product of the spring constant of the trap k, and the object displacement, Δx. In practice this method has been replaced in our instrument by direct measurement of light momentum (see Smith et al., 2003). Such methodology has been applied to measure the stability of proteins and to observe events of folding and

unfolding of proteins at a single-molecule level (Cecconi et al., 2008). Also, it has been possible to determine the characteristics of several molecular motors such as kinesin, myosin, RNA and DNA polymerases, several translocases of rings of the superfamily AAA+ ATPases, follow the dynamic of ribosomes during its elongation phase, etc. (Bustamante et al., 2011). Moreover, as mentioned above, this method allows to experiment with processes that occur in non-equilibrium conditions (Liphardt et al., 2002).

Among the most interesting applications of these techniques are those experiments that follow *in singulo* the individual behavior of a molecule and try to determine, for example, the mechanochemical aspects of its functioning. Particularly, in this work we wish to understand the changes that occur in a protein when the substrate binds, and understand which molecular parameters are affected by the binding of the substrate.

Glucokinase from *Thermococcus litoralis*

The protein used in this study is glucokinase from *Thermococcus litoralis* (tlGK), a thermophile archae which lives at temperatures above 80 °C. Since this enzyme is a monomer, it has several properties that make it a good model for optical tweezers studies. A monomeric protein has an advantage over an oligomeric one, because when working on individual molecules there is a very large dilution of the species that may induce subunit dissociation. It has no cysteines: this is important, since the DNA handles are attached by means of cysteines to the protein. The absence of

cysteines allows us to carry out cysteine engineering in specific places without affecting the protein´s structure.

This enzyme is also interesting since, unlike the kinases from most organisms, it uses ADP instead of ATP as a phosphoryl donor, and could participate in the modified Embden-Mayerhof pathway, since it has been proposed that these thermophilic organisms use ADP instead of ATP in this pathway (ADP shows higher thermostability than ATP, especially in the presence of divalent cations; Ito et al., 2001).

It has been demonstrated that the site of nucleotides binding for the tlGK is in a similar structural position to the proteins of the ribokinase family, independently of whether they bind ADP (as in the case of tlGK) or ATP (as is the case for other family members; Ito et al., 2001).

A remarkable feature of these proteins is the strict conservation of the binding site for the α- and β-phosphates of ADP and for the β- and γ-phosphates of ATP (Ito et al., 2001)

Comparison of the crystallographic structures of tlGK with kinases that use ATP as phosphoryl donor, shows that there are conserved structural and functional motifs, and that the recognition sites of the α- and β-phosphates of ADP in tlGK are found in the same position as those that recognize the β- and γ-phosphates of the ATP in the ATP dependent kinases (Ito et al., 2003; Koga et al., 2000). However, when comparing kinases depending on ATP with kinases depending on ADP, the recognition of the adenine and ribose varies due to the different number of phosphate groups in these nucleotides (Figure 2).

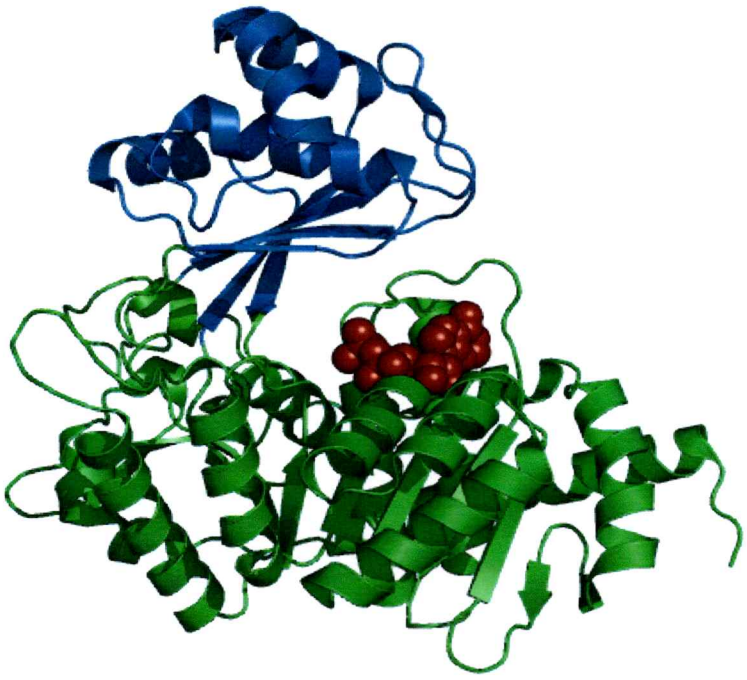


Figure 2. Three-dimensional structure of glucokinase from *Thermococcus litoralis* (1GC5). Above. Ribbon diagram of tlGK including an ADP molecule. The tlGK protein is composed of 467 amino acid residues and 2 domains, with a reentrant topology. The big domain (from aa 1 to 39, 119 to 177 and 210 to 467) is in green and the small domain (from aa 40 to 118 and 178 to 209) in blue. In brown is the ADP molecule. Below. Schematic representation of tlGK reentrant topology colored as above. Structure determined by Ito et al. (2001).

The kinetic mechanism of the tlGK is not completely understood. Nevertheless, some studies have provided some insights into the mechanism of binding of the substrates. In the crystal structure of tlGK with ADP, a conformation where the active site is semi-closed has been identified (Ito et al., 2001). The crystallographic structure of glucokinase from *P. furiosus* in complex with glucose and a non hydrolysable ADP suggests a closure of the active site induced by the substrates; similar to what is observed with ribokinase and adenosine kinase (Ito et al., 2003). These observations suggest that ADP binding induces a small closure of the enzyme and that upon glucose binding, the active site closes further. This interpretation further implies the order in which the substrates bind; first ADP and then glucose.

The effect of substrate in single molecule studies of proteins

Up to now there are not many studies that show the effect of substrate in the mechanical properties of proteins and none of them have been conducted in conditions of thermodynamic equilibrium. The work of Bertz and Rief (2009) show the importance of the pulling axis to observe the effect of the substrate. They have worked with the MBP protein and when the protein was stretched from the N and C termini, no substrate effect was seen, but when the protein was pulled by an axis perpendicular to the hinge zone, the presence of the substrate resulted in a mechanical stabilization of the protein in the hinge zone (Figure 3). Other works have shown the stabilization effect of the substrate but have been restricted to forces applied at the N and C termini only in non equilibrium conditions. For example, the stabilization of mouse dihydrofolate reductase

and Im9 has been shown in this manner (Junker et al., 2005; Hann et al., 2007). A recent work close to thermodynamic equilibrium conditions has revealed the interaction of calmodulin with calcium and several target peptides in detail (Junker et al., 2009).

In the present work, our main purpose was to find an appropriate pulling axis in order to study the effect of the substrate ADP on the properties of tlGK under equilibrium conditions. The constructs S4C/T418C and S4C/T57C were pulled to investigate the effect of the ADP substrate on the stability of the protein. At 28 pN, the construct S4C/T57C showed hopping between two conformations with a difference in extension of around 5.3 nm. The ΔG of unfolding in absence and in presence of ADP was 11.2 and 12.0 kcal/mol, respectively. A very high value considering that around 10% of the total structure is being pulled between the DNA handles, which suggests energetic coupling of the unfolding between various regions of the protein due to its the re-entrant topology of the protein. For the unfolding reaction, the distances from the native to the transition state were $x^{\ddagger} = 3.09 \pm 0.35$ nm in the absence and 2.81 ± 0.62 nm in the presence of ADP, respectively; the corresponding values during refolding were 2.24 ± 0.24 nm and 2.46 ± 0.43 respectively. The rate of the unfolding extrapolated to zero force was an order of magnitude smaller in the presence than in the absence of ADP, with no changes in the extrapolated refolding rates. These results suggest that the folded state of the protein is stabilized by the substrate with no effect on the distance to the transition state for the unfolding reaction.

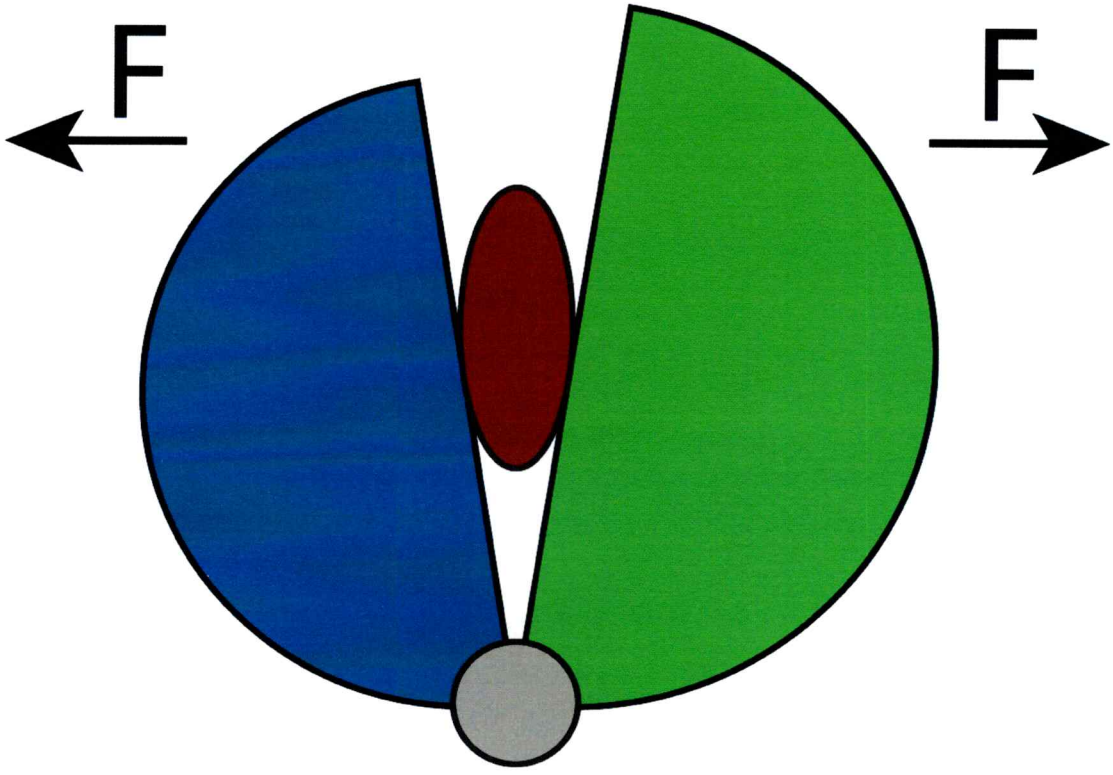


Figure 3. Direction of the applied force in the maltose binding protein. The force is applied perpendicular to the hinge of the protein. In these conditions it was possible to observe mechanical effects on the protein when adding the substrate. The same direction of force application will be utilized in the tlGK studies. Colors as in Figure 2. Figure modified from Bertz and Rief (2009).

OBJECTIVES

Main objective

The main purpose of this work was to find an appropriate pulling axis in order to study the effect of the ADP substrate on the thermodynamic and mechanical stability of glucokinase from *Thermococcus litoralis* under equilibrium conditions at the single molecule level using the optical tweezers methodology.

Specific objectives

1) Design of cysteine mutants in places that do not have a high degree of conservation, are exposed to the solvent (so the DNA handles can react) and in zones where the binding sites are left perpendicular to the hinge of the protein.

2) Purification and characterization of the glucokinase cysteine mutants.

3) Modification of the glucokinase and mutant enzymes to be studied at the single molecule level

a. Activation of cysteines with DTDP.

b. Attachment of the DNA handles to the enzyme

c. Attachment of the DNA-protein-DNA to the optical tweezers.

d. Kinetic studies in bulk of the enzyme attached to the DNA handles.

e. Enzyme activity measurements of the DNA-protein-DNA construct

f. Purification of the DNA-protein-DNA species

4) Search of a mutant glucokinase that shows equilibrium reversible unfolding under mechanical manipulation

a. Determination of the force at which mechanical unfolding occurs with the different glucokinase mutants constructed.

b. Determination of the kinetic parameters for the folding and unfolding transitions *in singulo*

5) Determination of the effect of the ADP substrate on the mechanical stability of the protein. Determination of thermodynamic (ΔG) and kinetic parameters (distance to the transition state and folding/unfolding rate constants at zero force) in the presence and absence of ADP

MATERIALS AND METHODS

MATERIALS

Bacterial strains

The *E. coli* strain BL21 (DE3) was used. It has a copy of T7 RNA polymerase gene in its genome under the control of the *UV5* promoter, inducible by IPTG, transformed with plasmid pET17b (obtained from Dr. Takayoshi Wakagi of the Biotechnology Department, University of Tokio) that contains the tlGK gene for overexpression of the enzyme. The *E. coli* strain XL10 GOLD was used for the transformation and extraction of the mutant plasmids. Both strains were chemically competent by the Inoue method (Inoue et al., 1990). The bacteria recombinants were kept in aliquots of approximately 1 ml in 50% glycerol at -80 °C.

Primers

The primers used to introduce mutations in the tlGK gene were designed according to the instruction of the QuickChange ® Site-Directed Mutagenesis Kit from Stratagene and were synthesized by IDT (Integrated DNA Technologies, US).

Culture media

The culture media used for bacterial strain growth and for protein overexpression was LB (Luria Bertani) from Fisher Scientific Inc. Culture media were sterilized by autoclaving.

Reagents

The reagents NAD^+, ADP, NBT, DTDP, PMS, and the enzyme G6PDH were obtained from Sigma Chemical Co. Gradient gels (4-20% Ready gel) and Micro Bio-Spin® 6 columns were obtained from Bio-Rad Laboratories Inc. Qiaquick® Gel extraction kit, Qiaprep® Spin Miniprep kit, HiSpeed® Plasmid midi kit were from Qiagen, and sybr green from Invitrogen

METHODS

Overexpression of tlGK and of the mutant proteins

In order to over express tlGK, 20 ml of the LB media were inoculated with a colony of the BL21 strain (DE3) transformed with the molded pET17b which contains the gene for the enzyme and it was incubated over night at 37 °C. The following day, 6 liters of LB media, complemented with 200 μg/ml ampicillin and 34 μg/ml chloramphenicol, were inoculated with the previous culture and incubated at 37°C with agitation until the culture reached an absorbance of 0.5 at 600 nm. To induce the expression of the tlGK, IPTG was added to a final concentration of 0.8 mM and the incubation was continued at 37°C for 4 h. The induced bacteria were collected through centrifugation for 20 min at 6,000 rpm at 4°C (Sorvall Evolution RC, rotor SLC6000, Thermo Electron Co.) in order to get a bacterial sediment.

Glucokinase purification

Purification we started with the bacterial sediment (mention above) from 6 l of culture resuspended in 100 ml of buffer A: 100 mM Tris-HCl pH 7.8, 10 mM $MgCl_2$. The bacteria suspension was lysed by sonication (Sonicator 3000, Misonix, Inc.); 8 pulses of 30 s (60 watt) and 30 s of cooling. The lysed bacteria were centrifuged for 20 min at 12,000 rpm (refrigerated Sorvall RC-5B, rotor SS34, Du Pont Instruments). The mutant enzymes and wild type tlGK were purified with the procedure of Koga et al. (2000) with some modifications. This method has four purification steps after breaking the cells by sonication. The first step is a thermal shock at 90 °C for 30 min. Denatured proteins were discarded through centrifugation for 20 min at 12,000 rpm (refrigerated Sorvall RC-5B, rotor SS34, Du Pont Instruments). Afterwards the supernatant was made 60% in ammonium sulfate, incubated for 1 h and loaded into a Toyopearl butyl-650M (TOSOH) hydrophobic column (11 x 1.8 cm) equilibrated with buffer A with 60% of ammonium sulfate. After washing the column with the equilibration buffer the protein was eluted with a gradient from 60% to 0% ammonium sulfate and the fractions with activity were pooled. The pooled fractions were dialyzed twice in buffer A and loaded into an ionic exchange column (Hitrap-Q HP, 1.6 x 2.5 cm, 5 ml pre-packed, General Electric, formerly Amersham Biosciences). After washing the column with buffer A, a NaCl gradient from 0 to 1M was applied; the fractions with activity were pooled (the enzymatic assay is mentioned below) and concentrated in an Amicon (Millipore) membrane to 2 ml. The sample was then loaded to a size exclusion column, (Hi load^{TM} 16/60 Superdex 75, 2.6 x 60 cm, 120 ml pre-packed, General Electric,

formerly Amersham Biosciences) equilibrated with buffer A. During the purification of mutants that have cysteines, 2 mM DTT was added in all purification steps.

Normally, 4-5 mg of enzyme were obtained per liter of culture at the end of the purification procedure.

Site directed mutagenesis

All sited directed mutagenesis were performed with the QuickChange ® Site-Directed Mutagenesis Kit (Stratagene). The primers were made with the specification given by the company. All plasmid the mutations were verified by sequencing (ELIM Biopharmaceuticals, Inc. California)

Enzymatic assay

The enzymatic activity was measured spectrophotometrically following absorbance at 340 nm at 40°C by coupling the production of glucose-6-phosphate to the reduction of NAD^+, using the auxiliary enzyme glucose-6-phosphate dehydrogenase (G6PDH) modified from Tsuge et al. (2002).

$$\text{ADP + glucose} \xrightarrow{\text{tlGK}} \text{glucose-6-P + AMP}$$

$$\text{glucose-6-P + NAD}^+ \xrightarrow{\text{G6PDH}} \text{6-phosphoglucono-}\delta\text{-lactone + NADH}$$

The reaction was performed in 100 mM Tris HCl pH 7.8, 0.5 mM NAD^+, 10 mM $MgCl_2$, 10 mM glucose, 2 mM ADP, with 1.72 µg of G6PDH and the sample tlGK in a final volume of 0.7 ml. A unit of enzyme (U) is defined as the quantity of enzyme

that catalyzes the formation of 1 μmol of product for min in the defined experimental conditions.

Protein determination

Protein concentration was determined by the Bradford (1976) method using BSA as a standard.

Generation of DNA handles

The handles were made as described by Cecconi et al., (2008). The DNA handles were synthesized by means of PCR, using as template plasmid pGEMEX from Promega. For the generation of DNA handles the reaction volume was 9.6 ml and modified DNA oligos (biotin, digoxigenin and SH in the 5' end) were used as primers. Specifically, the primers were: 5' thiol-GCT-ACC-GTA-ATT-GAG-ACC-AC, 5' biotin or digoxigenin-CAA-AAA-ACCCCT-CAA-GAC-CC. The DNA handles were purified with HiSpeed® Plasmad Midi Kit from Qiagen using 3 mM DTT in each step.

Cysteine activation and DNA handles attachment

The activation of cysteines was made by the method of Cecconi et al. (2008) with slight modifications. The protein, at a concentration of about 10 μM, was incubated with 10 mM DTT for two hours at room temperature. Afterwards, the protein was passed through a Micro Bio-Spin 6 column (Bio-Rad), previously equilibrated with 100 μM DTDP. The DTDP reacts with a sulfhydryl-containing molecule, such as the

cysteine residues in proteins, to form an activated thiol-pyridine molecule and pyridine-2-thione. The thiol pyridine-2-thione is released into solution and can be detected as an increase in absorbance at 343 nm (Shank, 2004).

Once the protein sample was passed through the column, DTDP was added to the sample solution so as to reach a concentration 25 fold higher than that of the protein (sometimes a white DTDP precipitate was formed). The optimum reaction time for tlGK with DTDP was two hours. After the reaction was completed, the excess DTDP was removed from the enzyme using two Micro Bio-Spin 6 columns (Bio Rad) and the reaction with DNA was started using a molar ratio of DNA to protein of 4:1. Meanwhile, the DNA handles were incubated for 2 h with 10 mM DTT and DTT was removed using the same columns of the DTDP removal procedure. The columns were equilibrated with a 100 mM Tris-HCl pH 7.8, 10 mM $MgCl_2$ buffer to proceed with the DNA handle reaction.

Degree of solvent exposure of the cysteines

VADAR program version 1.7 was used for determining the percentage of exposure to the solvent of different residues. This program uses the three-dimensional structure of tlGK to carry out the determinations (Willard et al., 2003; http://redpoll.pharmacy.ualberta.ca/vadar/).

In-gel activity assay

We used a coupled assay that led to the formation of a blue precipitate called

formazan in order to visualize glucokinase activity. With this assay the blue formazan precipitate is generated only in the zone of the gel (PAGE) where glucokinase is present. This assay is based on coupling the reaction of the lactate dehydrogenase with the formation of formazan (Guerrero, 2008).

$$NAD^+ + \text{lactate} \rightleftharpoons NADH + H^+ + \text{pyruvate}$$

$$NADH^+ + PMS_{(ox)} \rightleftharpoons NAD^+ + PMS_{(reduced)}$$

$$PMS_{(reduced)} + NBT \rightleftharpoons PMS_{(ox)} + \text{formazan}_{(pp)}$$

For measuring tlGK, its reaction is coupled to the reaction of G6PDH so as to form NADH.

$$\textbf{ADP + glucose} \xrightarrow[\textit{tlGK}]{Mg^{+2}} \textbf{AMP + glucose-6-P}$$

$$NAD^+ + \text{glucose-6-P} \xrightleftharpoons{\text{glucose-6-P DH}} NADH + H^+ + \text{6-phosphogluconolactone}$$

$$NADH^+ + PMS_{(ox)} \rightleftharpoons NAD^+ + PMS_{(red)}$$

$$PMS_{(red)} + NBT \rightleftharpoons PMS_{(ox)} + \text{formazan}_{(pp)}$$

To conduct a gel activity assay, first a native PAGE is run. The blue native gel consists of a running a gel under native conditions, using Coomassie Blue instead of SDS. The samples are treated in the following way. The protein sample, or DNA-protein-DNA, is mixed with loading buffer (this is 4X buffer): 0.418 g of BisTris, 0.107 ml of HCl 6 M, 4 ml of glycerol, 0.117 g of NaCl, 0.4 g of ponceau S, in a final volume of 10 ml, pH 7.2. When setting up the electrophoresis system, the buffer that corresponds to the cathode buffer is 0.02 % in Coomassie-G250 and the running buffer (20X) is made with 209.2 g of BisTris and 179.2 g of Tricine in 1 liter, pH 6.8.

The gels used to run the samples are native gels in polyacrylamide gradient (4-20% Ready Gels, Bio-Rad or Invitrogen). The gel is run at 150V for 1.5 h at room temperature. Then the gel is soaked with running buffer for a couple of hours with continuous buffer changes to discolor it and is incubated with the following solution at 40°C for 1 h: 100 mM Hepes pH 8.0, 1 mM NAD, 2 mM ADP, 9 mM glucose, 7 mM $MgCl_2$, 0.39 mM phenazine methosulfate, 0.09 mM NBT and 25 U of G6PDH in a total volume of 30 ml. To stop the reaction the gel is soaked in water.

DNA-protein-DNA purification

The native gel described above is used for purification by means of electroelution. Gels are run with the sample of interest in duplicate (in each gel). One half of the gel will be used to detect the glucokinase activity of the different species and the other will be used for the electroelution of the selected species. Once one of the halves has been treated to detect glucokinase activity, the selected bands are cut from

the corresponding spots of the untreated half with a razor blade. Then the gel pieces are placed in the Whatman electroelution device (Figure 4). The elutrap device consists of a 1.2 x 10 cm sample chamber limited at each end by a BT1 membrane which has a cut off of 3-5 kDa; molecules with lower molecular weights can pass through it under the influence of an electric field. A microporous membrane BT2 acts as a prefilter that prevents agarose, acrylamide and other big particles to enter to the purified sample. Together, the BT2 and BT1 membranes form a trap into which the sample migrates. The electric field acts as the driving force filtering the molecules through the membranes. When the voltage is switched off, the membranes seal the trap, preventing diffusion of the sample out of the trap (Figure 4). The minimum volume that one can use in the camera is 200 µl. The sample is left over night at 100 V and the next day the polarity is inverted for 20 sec so as to loosen the sample stuck to the membranes.

Once the electroelution is finished, the sample is taken with a plastic pipette and concentrated by means of Microcom (Millipore). The pure DNA-protein-DNA was kept in aliquots of approximately 50 µl in 50% glycerol at -20 °C.

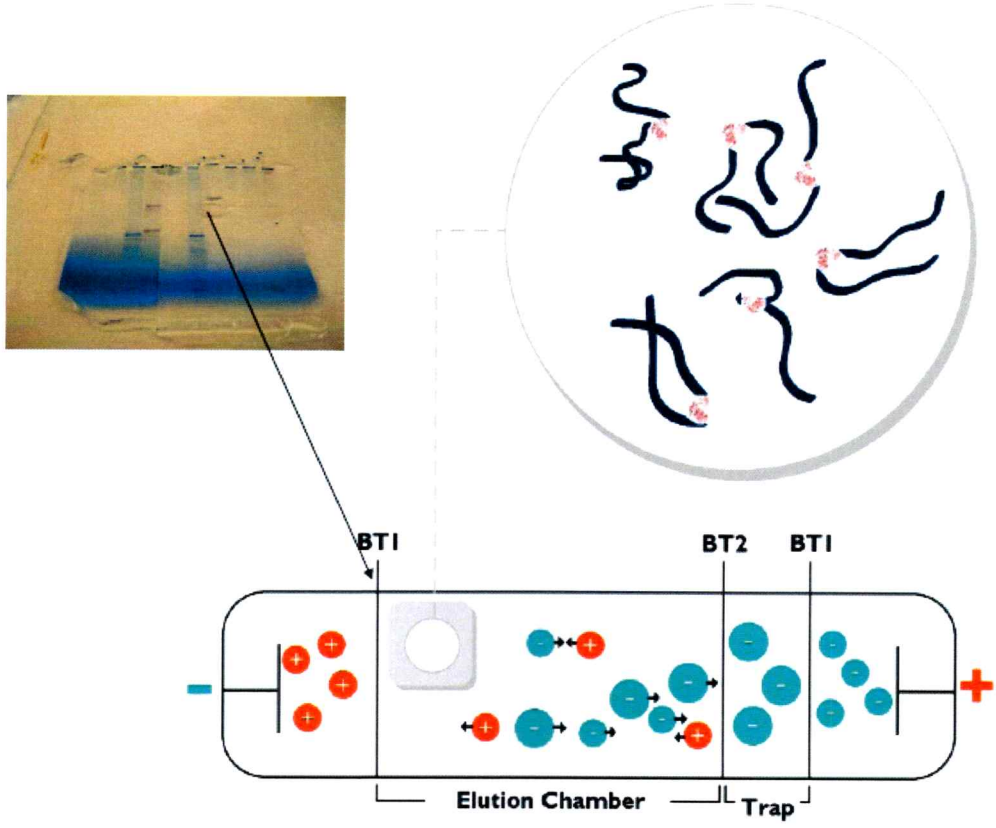


Figure 4. Schematic representation of the electroelution device. The electroelution system consists of a standard electrophoresis camera with 2 membranes (BT1 and BT2) that generate a trap for the electroeluted molecules. In the elution chamber the piece of gel is placed with the molecules to be extracted. Figure adapted from Whatman (Elutrap Electroelution System Manual)

Attachment of DNA-protein-DNA to polystyrene beads

Pure protein-DNA constructs were first reacted with polystyrene beads coated with anti-digoxigenin antibodies (dig-beads) for 15 min at room temperature (RT). The best dilution of the protein-DNA constructs had to be searched for experimentally in order to work with them in the optical tweezers setup. Serial dilutions were performed to 1/10, 1/100 and 1/1000 in order to determine which was the best concentration to use. Dig-beads were generated by coupling anti-digoxigenin antibodies (Roche) to 3.2 μm proteinG-coated beads (Spherotech). After reacting with the protein-DNA construct, a dig-bead was placed in the optical trap and then brought in close proximity to a 2.10 μm streptavidin-coated bead (Spherotech) which was held in place at the end of a pipette, until a tether between the two beads was attained (Cecconi et al., 2008).

Atomic force microscope measurements

AFM was performed in air with a Digital Instrument Nanoscope III, in tapping mode, using tips from Nanosensors (point probes, type NCH-100). The products of the reaction with the DNA handles were diluted to a final concentration of 2 nM in deposition buffer (10 mM HEPES, pH 7.5, 10 mM NaCl, 2 mM $MgCl_2$). Twenty μl of this solution were deposited onto freshly cleaved mica (Asheville-Schoonmaker Mica Company #472X8 /10XSF) and allowed to adsorb onto the surface for 1 min. The sample was then gently washed with doubly distilled water and dried with a stream of nitrogen.

Mass spectrometry

All mass spectrometry experiments were done in the HHMI facility with the help of Dr. David King. The MALDI technique was used in order determine the molecular weight of the different mutant enzymes, with and without DTDP.

Optical tweezers

A custom optical tweezers machine was made in the laboratory of Prof. Carlos Bustamante and used for these experiments (Figure 5). Such a device called "miniTweezers" contains several patented innovations (Bustamante and Smith, 2006; Bustamante and Smith, 2007). The name "miniTweezers" derives from its compact size. It is a cost-effective analytical optical trap capable of resolving picoNewton forces, Angstrom distances and millisecond event times.

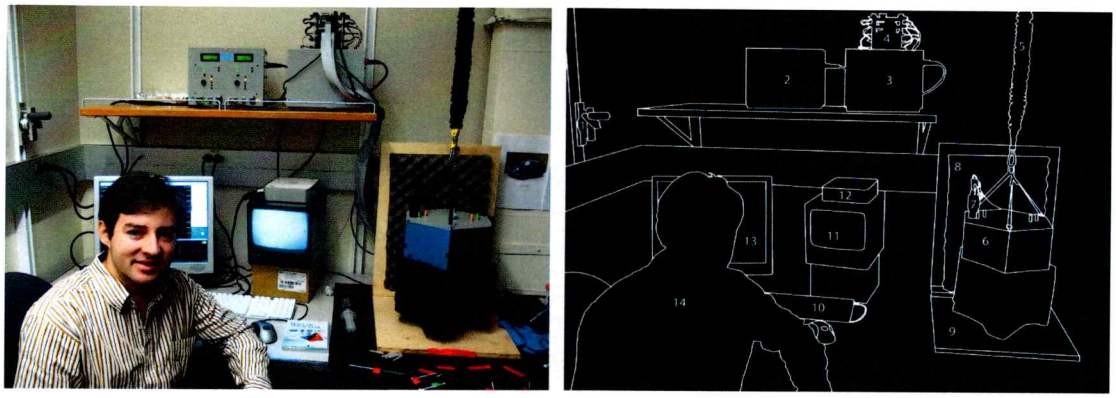


Figure 5. MiniTweezers machine. Left. The complete system to operate the optical tweezers with an operator. The optical tweezers are used inside a box for acoustic isolation, and it is suspended from the ceiling with a spring (bungee cord) to minimize vibrations from the laboratory floor. Right. Key to the items of the picture on the left. 1. Winch for bungee cord. 2. Laser power supply. 3. Controller power supply. 4. Controller circuit board (analog to digital converter board). 5. Bungee cord. 6. Optics head. 7. CCD television camera. 8. Sound isolation. 9. Base of wood isolation box. 10. Keyboard. 11. Television monitor. 12. Macintosh computer (Mac-mini). 13. Computer screen. 14. Operator (picture, thanks to Jesse Dill)

This instrument measures forces on a trapped object in 3 dimensions, namely an X and Y force which are transverse to the laser-beam axis, and also a third Z-axis force parallel to the optic axis. Quantification of these forces is based on the conservation of the light momentum as explained by Smith et al. (2003). Light which passes through the trapped object is deflected to the side if there is a force on the object. PSD detectors (silicon photo cells) measure both the changes in power deflections of the light beams leaving the trap region (for transverses forces), and changes in power concentrations of the light beams leaving the trap region (for longitudinal forces). To do this (see Figure 6), the output light beams from the beads are projected by relay lenses and split by a non-polarizing beam splitter cubes and land on two separate detectors that measure the transverse forces (Fx and Fy) and the axial force (Fz) separately. The optical traps are moved relative to a pipette by moving two optical fibers (fiber wigglers). There is a web site (http://tweezerslab.unipr.it) where the construction of miniTweezers is described.

Single molecule experiments

Two types of experiments were performed with the mutants studied; at constant velocity and at constant force. The constant velocity experiments are based on moving the trap at a constant velocity relative to a fixed reference point (a bead on a pipette, for example). Generally 100 nm/sec was used, unless another velocity is mentioned. In the constant force experiments the light deflection is held constant in time. Here the traps (fiber wigglers) follow the motion of the molecule in time and the force (beam deflection) remains constant throughout the changes in extension and folded state of the

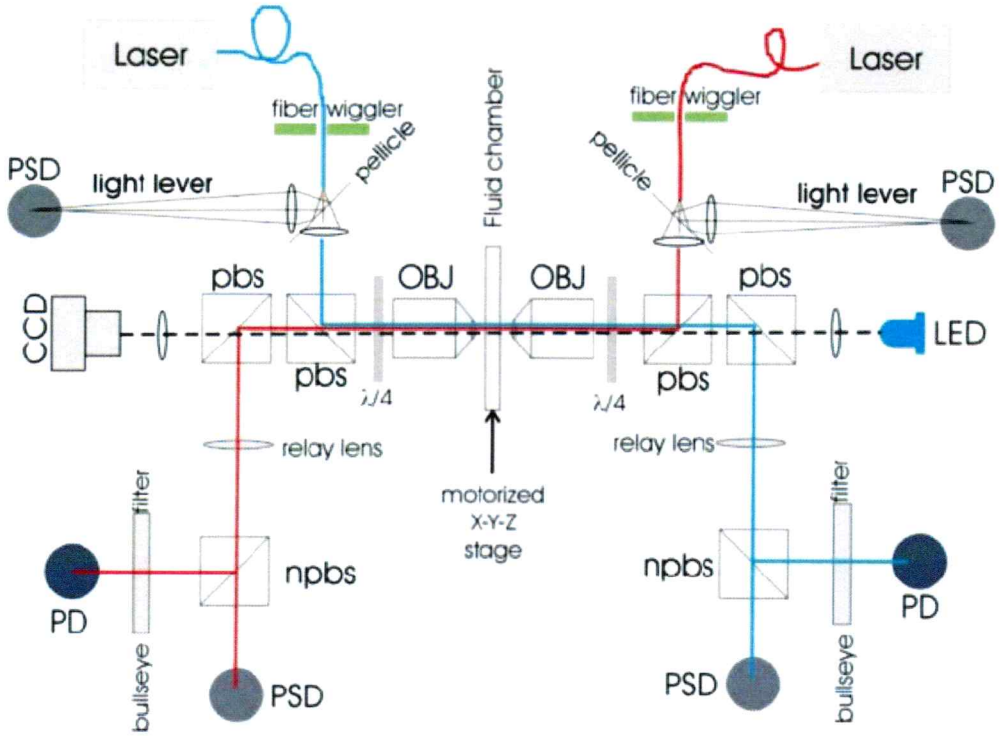


Figure 6. Diagramatic representation of a set of miniTweezers. The optical components that form the miniTweezers are shown. Optics components: *Laser* = Fiber-coupled diode laser, *pellicle* = Thorlabs BP108, *pbs*=polarizing beam-splitter cube, *npbs*=non-polarizing hybrid beam-splitter, *OBJ* = Olympus UPlanApo-60x-Water, *PSD* = position-sensitive photo detector UDT DL-10, *PD* = photodiode UDT 10-DI, *bullseye filter* = custom made on laser printer, *l/4* = quarter-wave plate, *CCD* = TV camera, *LED* = high brightness blue, *motorized stage* = Newport 561D-XYZ stage, Thorlabs Z606 motors. Figure taken from http://tweezerslab.unipr.it/ (Smith and Rivetti, 2010).

protein. The buffer used in the single molecule experiments was 100 mM Tris pH 7.8 and 10 mM $MgCl_2$. In the experiments in presence of ADP, 1 mM ADP was used.

Single molecule data acquisition

The data acquisition was at 1 msec in a program written in C++.

Single molecule data processing

The data processing was performed with the OPTIMAL program (Jesse Dill) that works in MATLAB. The hopping processing was done with a program made by Phillip Elms.

For the analysis of the kinetic data, the Bell model (Bell, 1978) was used, which predicts an experimental dependence of the rate constant on the external force:

$$k = k_0 e^{\frac{\Delta x^{\pm} F}{K_B T}}$$

The linearized equation is:

$$\ln k = \ln k_0 + \frac{\Delta x^{\pm} F}{K_B T}$$

Where k is the rate constant, k_0 is the rate constant at zero force, F the force, $\Delta x^{\pm}$ is the distance to the transition state, T the temperature in Kelvin and K_B the Boltzman constant.

The ln(k) *versus* F was plotted.

Analysis of the pulling data at constant velocity

For the analysis of the pulling data at constant velocity, the force at which protein unfolding occurs was determined and also the distance of this event. In order to determine the latter, curve was adjusted to a line before and after the unfolding and the distance between these lines was determined.

Analysis of the data with hopping at constant force

The velocity constant (k) was obtained from the distance data at constant force in an extension versus time plot (Figure 31). Here the hopping from the folded to the unfolded state is observed, so a line drawn in the middle between the two states is used to define the extension of the folded and unfolded states. It is then possible to determine the lifetimes of these states from the average time the molecule spend at these two positions. Once the lifetimes were quantified and the data is shown to follow a single exponential, the kinetic constants can be obtained from the inverse of the corresponding lifetimes.

Worm like chain fit

The equation of the worm like chain shows the dependence of the extension *versus* force for a flexible polymer in a thermal bath. A long polymer tends to contract as thermal forces try to randomize any alignment of its chain segments. This force has an entropic origin and its magnitude is given by (Bustamante et al., 2004):

$$F = \frac{k_B T}{p}\left[\frac{1}{4}\left(1-\frac{x}{L_C}\right)^{-2} - \frac{1}{4} + \frac{x}{L_C}\right]$$

where p is the persistent length of the chain (a value of 0.65 nm is used , see Cecconi et al., 2005); x is the end-to-end extension, Lc is the contour length; calculated by multiplying the number of amino acids by 0.36 nm/aa. Then the folded distance must be subtracted from the contour length, which corresponds to the distance between the attachment points of the DNA handles in the absence of force.

For each rip or mechanical unfolding in the plots of extension *versus* force (constant velocity experiments), the length of the rip is determined and is placed on a scatter plot of the rip length *versus* force. A best fit to the worm like chain values is made from these data by using the OPTIMAL program (Jesse Dill).

Thermodynamic studies

Through the observation of "hopping" or reversible unfolding between 2 states, we may determine the force necessary for the enzyme to spend half of the time in each state.

At constant temperature and pressure the reversible work is:

$W_{rev}(F) = \Delta G^{o} + k_B T \ln Keq(F) + \Delta G(F)_{stretching}$ (Liphardt et al., 2001)

Where $\Delta G(F)_{stretching}$ accounts for the free energy needed to stretch the unfolded state of the enzyme under force (it is obtained experimentally as the area under the WLC curve).

At $F=F_{1/2}$ the expression $k_B T \ln Keq(F) = 0$ since at $F_{1/2}$, Keq = 1, and

$W_{rev}(F) = \Delta G^{o} + \Delta G(F)_{stretching}$

Where $W_{rev}(F)$ is approximately $F_{1/2} * \Delta x(F_{1/2})$

and the ΔG^{o}, i.e. the standard free energy for the process in the abscence of force is given by;

$\Delta G^{o} = F_{1/2} * \Delta x(F_{1/2}) - \Delta G(F)_{stretching}$

RESULTS

1. Experimental setup

The experimental setup used was the one proposed by Cecconi et al. (2008). This corresponds to attaching the protein to polystyrene beads by means of DNA handles. The DNA handles are modified with digoxigenin and biotin, to attach them in a directional way to the polystyrene beads (Figure 7).

The polystyrene beads are covered with streptavidin and anti-digoxigenin antibodies. The DNA handles are attached to the protein by means of disulfide bonds. One of the polystyrene beads is attached to a pipette by suction and the other bead is the one that is trapped in the tweezers. The bead attached to the pipette usually is the one that has the streptavidin and is of 2.1 μm in diameter. The one that is attached to the tweezers' laser is 3.1 μm diameter. A different size is used to make the bead detection easier.

The maximum forces that may be generated in this experimental setup depend on the strength of the interaction between the digoxigenin and the anti-digoxigenin on the one hand, and biotin and streptavidin on the other.

The experimental results show that the average force of rupture of the interaction of digoxigenin and anti-digoxigenin is around 40 pN, and 120 pN for the rupture between the streptavidin and biotin with a pulling velocity of around 100 nm/sec. The optical tweezers machine may generate forces of up to 140 pN.

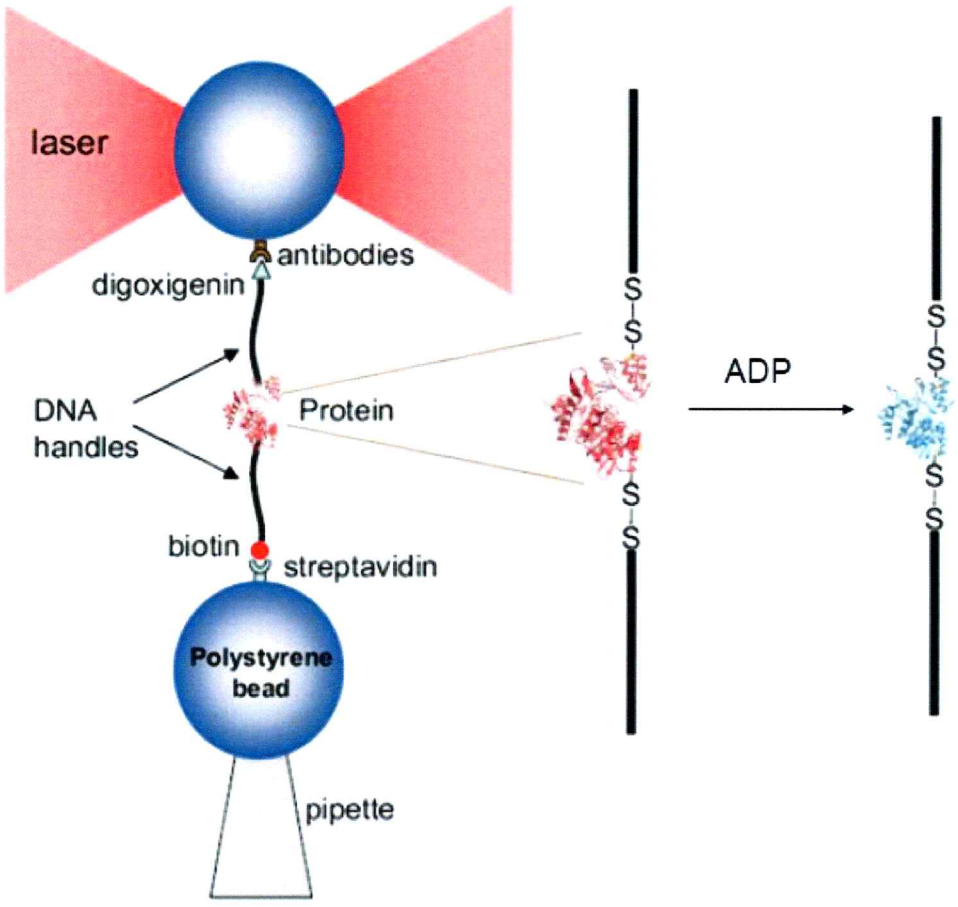


Figure 7. Experimental setup. The tlGK protein is attached to little balls of polystyrene by means of DNA handles. These DNA handles are derivatized in its 5' ends with digoxigenin and biotin, to bind the respective beads (antidigoxigenin coated beads and streptavidin coated beads, respectively). The streptavidin bead is attached to a micropipette by suction and has a diameter of 2.1 μm. This system is inside a laminar flow camera, which allows buffer change (with and without substrate). Figure modified from Cecconi et al. (2007).

2. Protein purification

Since the enzyme is present in a thermophilic organism, the experimental strategy uses a thermal shock as the first step to eliminate the proteins that are not temperature resistant. As seen in Figure 8, with this step a quite highly enriched fraction of the protein is obtained. The protein does not have a purification tag, so the following step used was a hydrophobic column chromatography (Toyopearl), because tlGK is quite hydrophobic. Next, an ionic exchange column was used (Figure 8) after which the enzyme co-purifies with a 10 kDa protein. Then, after a size exclusion column step two glucokinase activity peaks were detected: one that elutes first where the enzyme tlGK elutes together with the 10 KDa protein and the other with the pure enzyme of around 55 kDa, which is used in the experiments below (lane 6 of Figure 8).

This purification method was utilized with the various cysteine mutant enzymes used for DNA handles attachment, the purity of which was quite similar. The molecular weight of all the mutants was verified by mass spectrometry to determine the absence of undesirable mutations and to prove that it was the correct molecule.

3. DNA handles attachment

Several tlGK mutants were used for DNA handle attachment. Some of them were not modified in both cysteines and others were. This may be due to the degree of exposure that the engineered cysteines have. When performing the reaction, five species may be obtained in most cases: DNA alone, DNA-DNA, DNA-protein, DNA-protein-DNA and

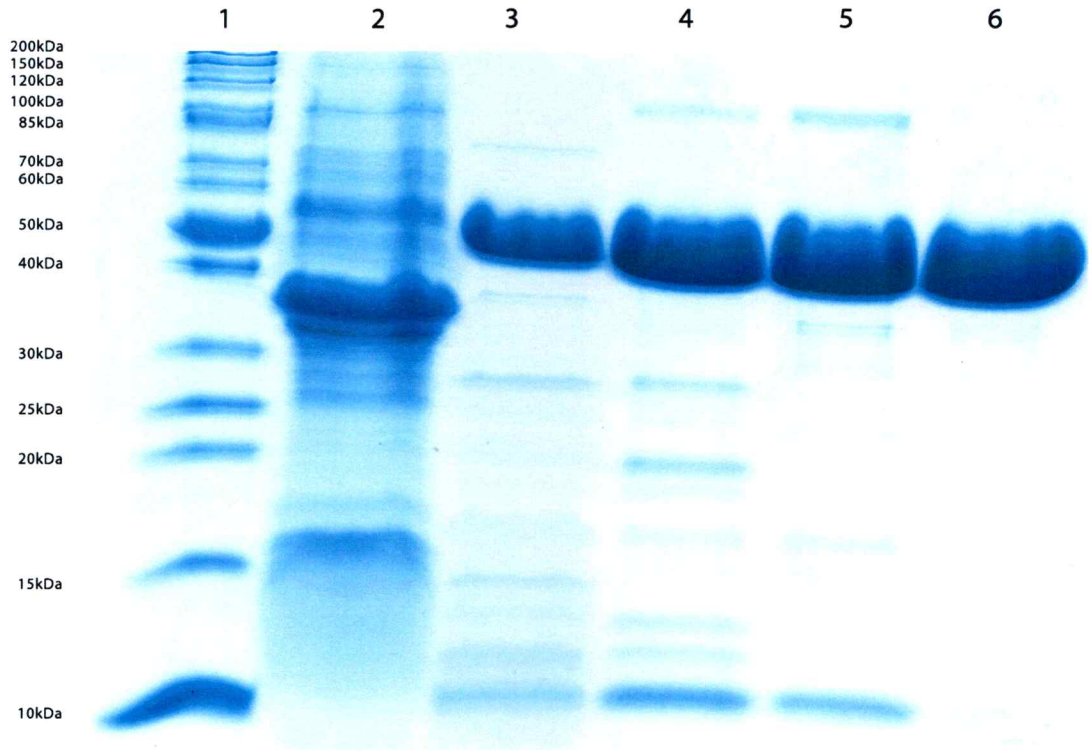


Figure 8. tlGK purification. The tlGK protein was purified in 4 steps. Lane 1: Molecular weights proteins standards. Lane 2: lysed bacteria supernatant. Lane 3. sample after thermal shock (80° C for 1 h). Lane 4: sample after hydrophobic column (Toyopearl). Lane 5: sample after ionic exchange column (hitrapQ) Lane 6: sample after size exclusion chromatography (S75). A 15% polyacrylamide gel was run at 130 V for 1 h.

protein alone. When running a gradient PAGE gel and staining for DNA (sybr green) four species are observed. Also, under the AFM the reactions show large quantities of undesirable species. What is important to mention is that only those that have double marks, that is to say digoxigenin and biotin, will be trapped by the polystyrene beads. Therefore, both DNA-DNA and DNA-protein-DNA will be observed in the optical tweezers. The efficiency of "fishing" in the optical tweezers is not very high; one can trap only about one fiber per day.

The first step of the DNA derivatization is a modification of the protein with DTDP and different conditions were tested. As observed in Figure 9, if the reaction with DTDP is left overnight and later the protein is derivatized with DNA, oligomeric protein species start to form via cysteine attachment. If the reaction time is 6-8 h, oligomeric species are still present and several species are formed after DNA derivatization (Figure 10). The best condition found was reaction of the proteins with DTDP for 2 h, and then derivatize with DNA (Figure 11). Mass spectrometry of the derivatized protein with DTDP gave the correct molecular weight, without the presence of dimers. Also PAGE in the presence of SDS, without reducing agent, show very low amount of dimeric species. Upon DNA derivatization the four expected species are observed (Figure 12). AFM images show the different species obtained and their proportions (Figure 13). These derivatization conditions were used with all the mutant proteins studied.

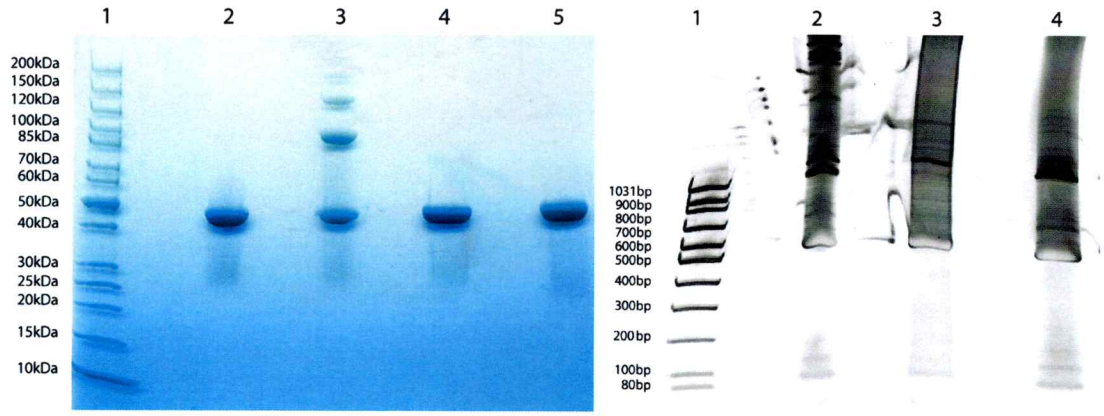


Figure 9. Overnight reaction with DTDP and handles attachment. Left. Polyacrylamide gel that shows the staining for proteins (Comassie). Lane 1: molecular weight standards; lane 2: protein before DTDP reaction; lane 3: protein after DTDP reaction overnight; lanes 4 and 5: protein after DTDP reaction plus reducent agent (DTT). Right. Polyacrylamide gel with staining for DNA (sybr green). Lane 1: DNA standard; lane 2: protein after reaction with DNA; lane 3: DNA handle derivatized with biotin; lane 4: DNA handle derivatized with digoxigenin. A 4-20% polyacrylamide gradient gel was run at 130 V for 1 h.

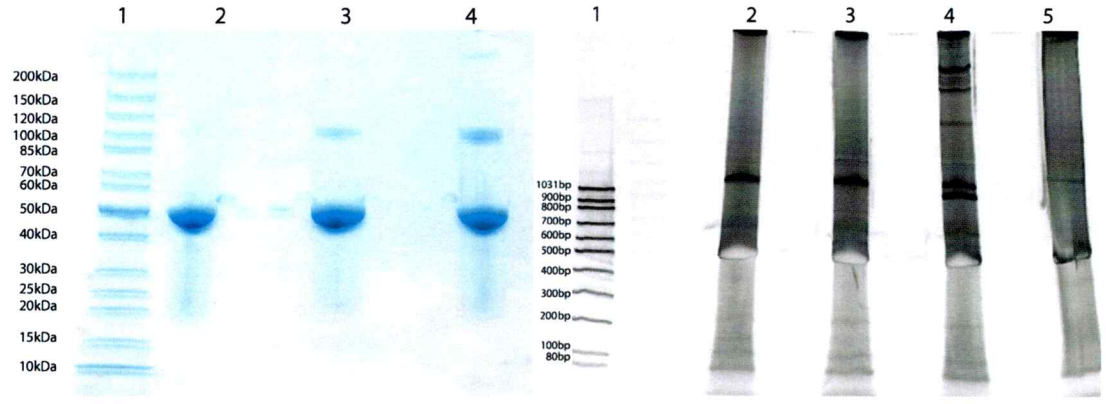


Figure 10. Handles attachment during 6-8 h reaction with DTDP. Left. Polyacrylamide gel showing the staining for proteins (Coomassie). Lane 1: molecular weight standards; lane 2: protein before DTDP reaction; lane 3: protein after DTDP reaction 6-8 h plus reducing agent (DTT); lane 4: protein after DTDP reaction 6-8. Right. Polyacrylamide gel stained for DNA (sybr green). Lane 1: DNA standard; lane 2: DNA handle derivatized with biotin; lane 3: DNA handle derivatized with digoxigenin; lane 4: protein after reaction with DNA; lane 5: protein after reaction with DNA plus reducing agent (DTT). A 4-20% polyacrylamide gradient gel was run at 130 V for 1 h.

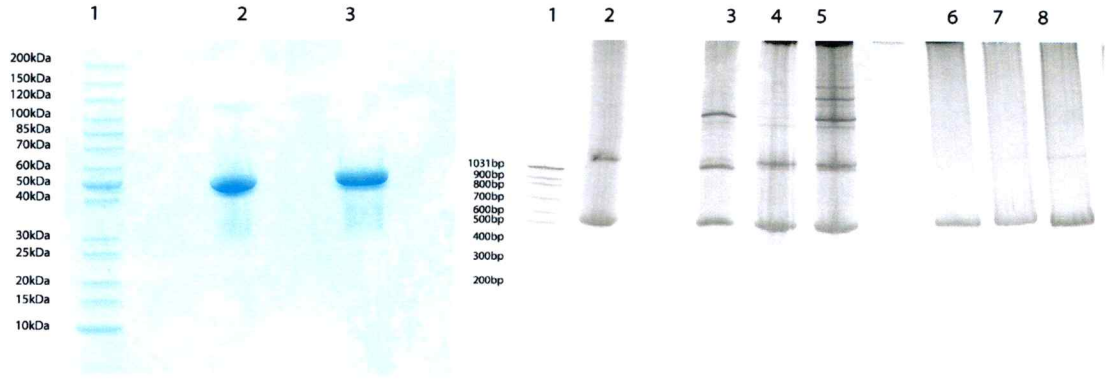


Figure 11. Handles attachment during 2 h reaction with DTDP. Left. Polyacrylamide gel that shows staining for proteins (Coomassie). Lane 1: molecular weight standards; lane 2: protein after DTDP reaction; lane 3: protein before DTDP reaction. Right. Polyacrylamide gel with staining for DNA (sybr green); lane 1: DNA standard; lane 2: DNA handle derivatized with biotin and DNA handle derivatized with digoxigenin; lane 3, 4 and 5: protein after DNA reaction; lane 6, 7 and 8: protein after DNA reaction plus reducent agent (DTT). A 4-20% polyacrylamide gradient gel was run at 130 V for 1 h.

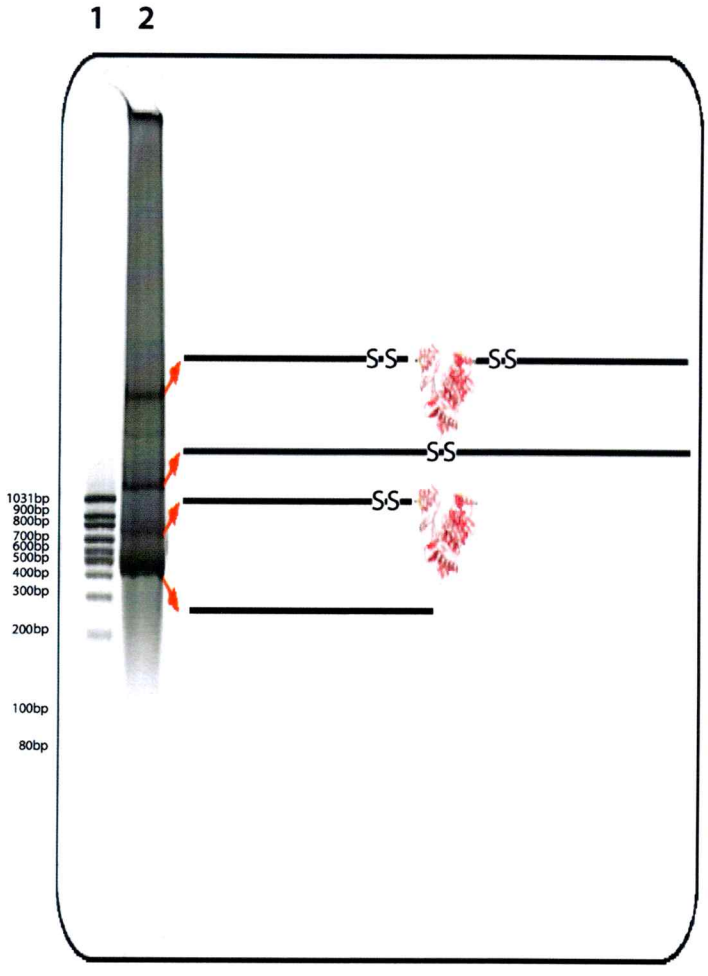
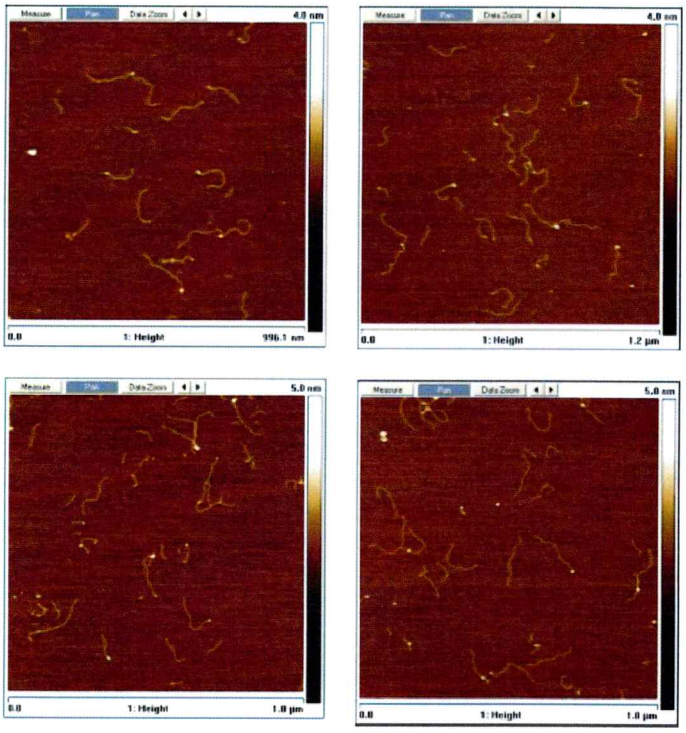


Figure 12. DNA handles attachment. Polyacrylamide gel stained for DNA (sybr green), showing the species obtained after DNA attachment. In this example the derivatization of the T57C/T418C mutant is shown. Four species are observed, as shown in a schematic way at the right of the gel. DNA-protein-DNA, the one with the greater molecular weight, DNA-DNA, DNA-protein and the one of lesser molecular weight is DNA alone. Lane 1: DNA molecular weight standards; lane 2: protein after DNA reaction. A 4-20% polyacrylamide gradient gel was run at 130 V for 1 h.



Species	Percentage
	5.4%
	15.4%
	74.2%
	5.0%

Figure 13. AFM studies of the sample after reaction with DNA. Left. Images of AFM taken in tapping mode, where the different species obtained after the reaction with DNA are shown. Right. Table that shows the percentage of each species.

4. In-gel activity assay

The activity test on gels allowed us to observe in a direct way if the proteins with DNA handles were active or not. There are no previous studies that demonstrate that a protein is active with DNA handles. This is because the quantity of product generated is quite small; DNA handles are generated through PCR (where pmoles of DNA are obtained).

The experimental strategy was to run a native gel (BN PAGE), technique developed by Schägger and von Jagow (1991), which uses Coomassie blue binding instead of SDS binding, to give a negative homogenous charge to the protein. It has been demonstrated that Coomassie blue does not alter the activity of the protein (Schägger and von Jagow, 1991). The activity gel was carried out for all the cysteine mutants that were generated.

As seen in Figure 14 left, the native gel allows observing the protein ladder staining and tlGK alone without the handles (lanes 1 and 4), but it is not possible to observe the DNAs-Protein constructs before the activity assay in-gel, because of the gel resolution. Once the gel is incubated with the enzymatic reaction mixture, only the proteins that have enzymatic activity generate the formazan blue precipitate. As seen in Figures 14 and 15, tlGK generates dimeric species, but in low proportions. DNA does not stain with formazan and neither does the ladder of native proteins used as standards, demonstrating that the method is specific for glucokinase. When DNA staining is observed, it correlates with the formation of formazan (Figure 15).

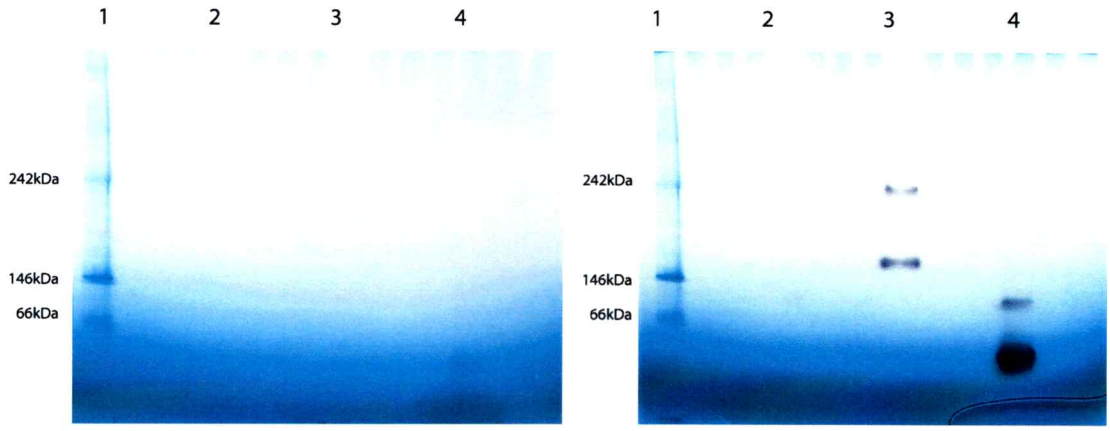


Figure 14. In-gel activity assay. Left. Blue native gel before the activity assay. Right. Gel after the activity assay where the formation of the blue formazan precipitate is observed in species that have glucokinase activity. In both gels, lane 1: molecular weight standards; lane 2: DNA handle derivatized with biotin and digoxigenin; lane 3: protein after reaction with DNA (mutant T57C/T418C); lane 4: tlGK alone. Native 4-16% Bis-tris PAGE (Invitrogen) at 150V for 1.5 h.

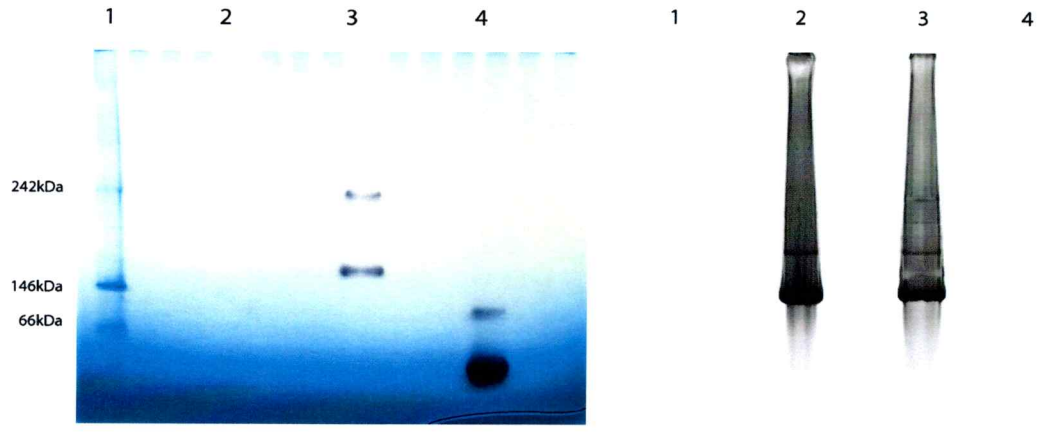


Figure 15. In-gel activity assay and DNA staining. Left. Polyacrylamide gel after the activity assay. The formation of the blue formazan precipitate is observed in species that have glucokinase activity. Right. Polyacrylamide gel stained for DNA (sybr green). In both gels, lane 1: molecular weight standards; lane 2: DNA handles derivatized with biotin and digoxigenin; lane 3: protein after reaction with DNA (mutant T57C/T418C); lane 4: tlGK alone. Native 4-16% Bis-tris PAGE (Invitrogen) at 150V for 1.5 h.

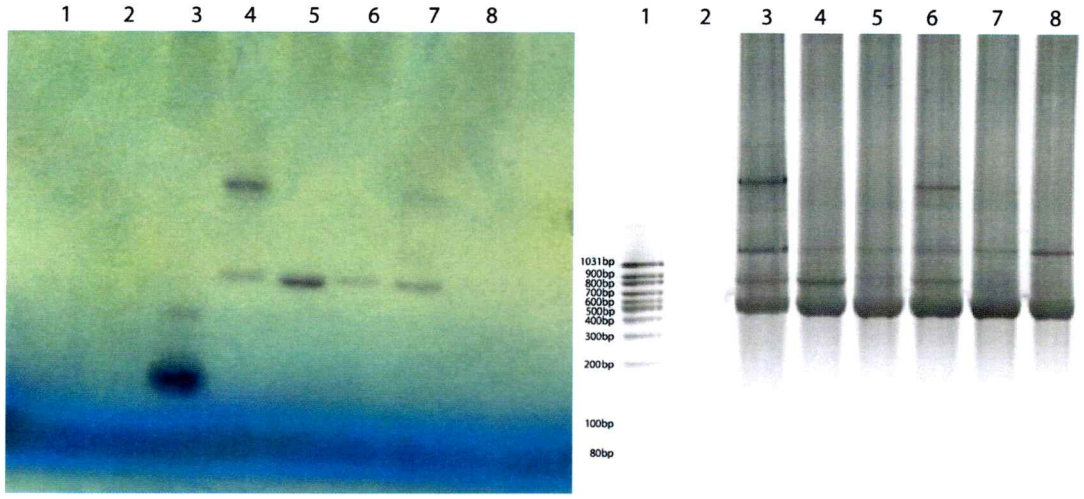


Figure 16. In-gel activity assay of the different cysteine mutants. Left. Polyacrylamide gel that shows some of the mutants with activity staining. Left. Polyacrylamide gel after the activity assay. The formation of the blue formazan precipitate is observed. Right. Polyacrylamide gel with staining for DNA (sybr green). In both gels, lane 1: DNA molecular weight standards; lane 2: protein alone; lane 3: T57C/T418C; lane 4: A26C/T57C; lane 5: S4C/S77C; lane 6: S4C/T57C; lane 7: DNA handle derivatized with biotin; lane 8: DNA handle derivatized with digoxigenin. Native 4-16% Bis-tris PAGE (Invitrogen) at 150V for 1.5 h.

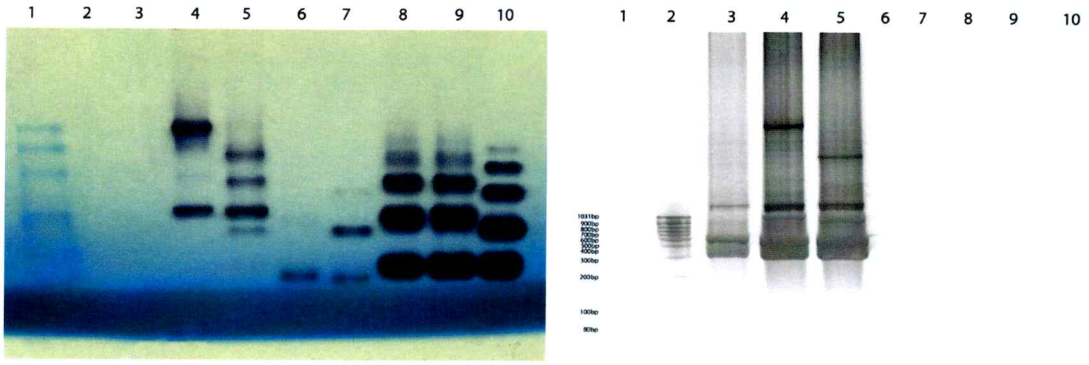


Figure 17. In-gel activity assay of the different mutants before and after DNase treatment. Left. Polyacrylamide gel after the activity assay. Right. Polyacrylamide gel stained for DNA (sybr green). In both gels, lane 1: molecular weight standards of native proteins; lane 2: molecular weight standards of DNA; lane 3: DNA handle derivatized with biotin and digoxigenin; lane 4: Q197C/T418C; lane 5: S4C/R467C; lane 6: Q197C/T418C after derivatization with DNA handles and treated with DNase I; lane 7: S4C/R467C after derivatization with DNA handles and treated with DNase I; lane 8, 9 and 10: tlGK left overnight in the absence of reducing agent. Native 4-16% Bis-tris PAGE (Invitrogen) at 150V for 1.5 h.

This activity gel was run with all the derivatizations that were performed. As seen in Figure 16 and 17 the T57C/T418C, Q197C/T418C, S4C/R467C and the S4C/T57C mutants were derivatized in the 2 cysteines and the DNA-protein-DNA species is observed. Other mutants like A26C/T57C, S4C/S77C, could not be derivatized in both cysteines.

The activity gels were also used to detect the presence of dimers once the protein has been derivatized with DNA (Figure 17). This was done by adding DNase to a sample derivatized with DNA, to establish if the recovered species are dimers or monomers. A control gel was run by adding DNase to the tlGK. Since tlGK solutions were shown to form dimers independently, these species was also observed after the addition of DNase (Figure 18). When DNase was added to the derivatized species, only monomers are observed. The only exception was the mutant where the cysteines are located at the N- and C-terminal ends. This mutant has one of its cysteines so exposed, that even after 2 h of reaction dimers and monomers are observed. Nonetheless, purification was achieved during attachment of the protein species to the tweezers. In this way, monomeric species can be chosen.

5. Purification of the DNA-protein-DNA constructs

It is important to use pure species when using the optical tweezers, since many different possibilities of bonding can occur. As shown in Figure 19, proteins with hydrophobic patches can interact with the beads, which are hydrophobic; this

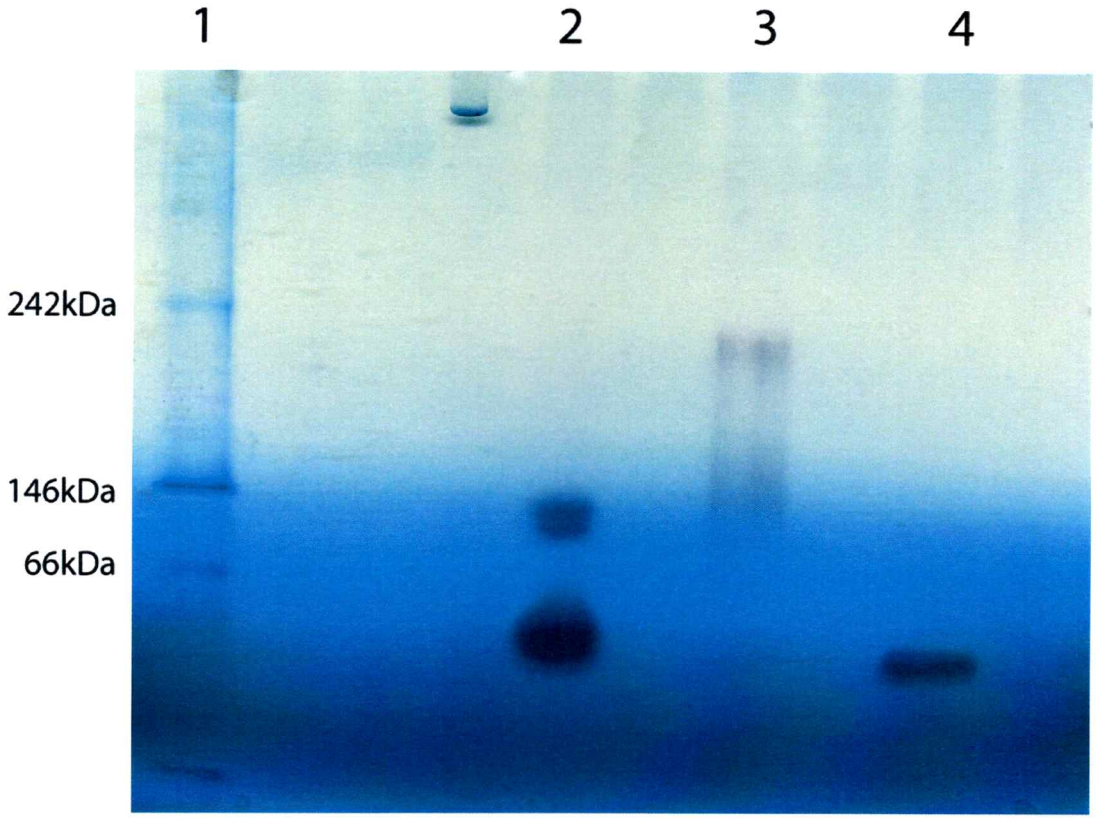


Figure 18. Assay with DNase and state of protein oligomerization with DNA handles. Lane 1: molecular weight standards of native proteins; lane 2: protein alone; lane 3: protein after derivatization with DNA handles; lane 4: protein after derivatization with DNA handles and treated with DNase I. Native 4-16% Bis-tris PAGE (Invitrogen) at 150V for 1.5 h.

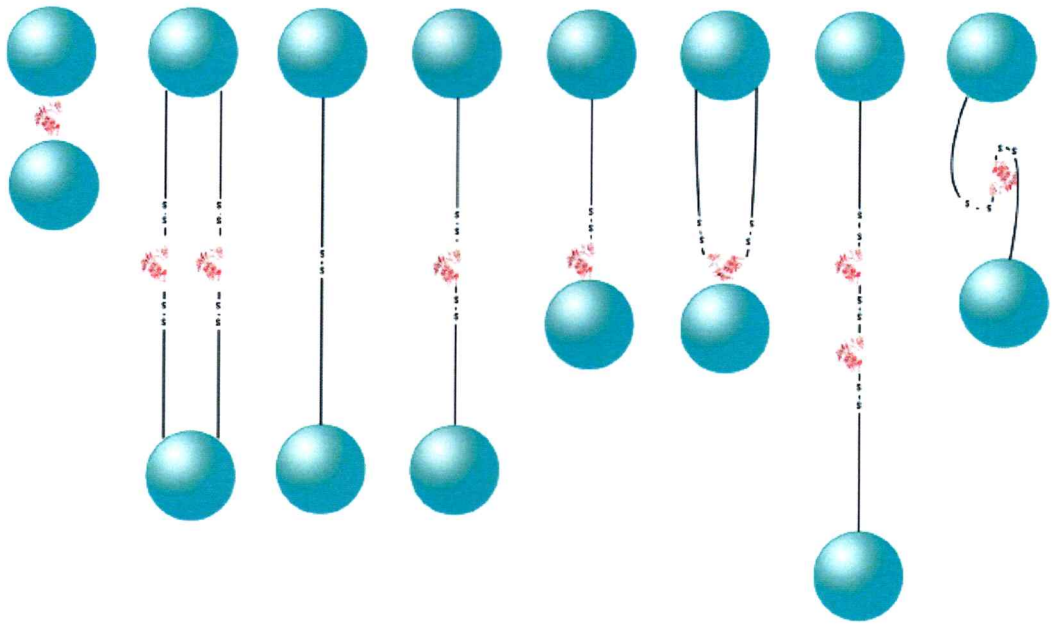


Figure 19. Different binding possibilities of the DNA constructs to the beads in the optical tweezers. Different binding possibilities that can exist for the protein molecules after the derivatization with DNA handles. 1. If the protein has hydrophobic zones, it could be attached directly to the beads. 2. There might be more than one fiber attached to the beads. 3. When there is DNA-DNA with both modifications (biotin and digoxigenin) it may attach to the beads. 4. The correct molecule of DNA-protein-DNA, 5. There may be only one DNA handle attached to the protein and if the protein has non-specific interaction with the beads, there might be binding. 6. The protein that was derivatized with two handles of the same chemistry and besides has non-specific interaction with the bead. 7. A protein dimer derivatized with DNA handles. 8. The protein has interaction with some DNA sequences.

interaction is decreased by the DNA handles. One of the major impurities that are observed when working with the optical tweezers is the presence of DNA-DNA without protein. The behavior of this species in the optical tweezers is known, so it is possible to detect it but, consequently, the fraction of correct fishing events observed decreases in their presence.

In order to increase the efficiency of fishing, the sample was purified by electroelution. As seen in Figure 20, the sample before the electrolelution shows four bands containing DNA. When the band that contains DNA-protein-DNA is cut and purified by electroelution, this species is greatly enriched (Figure 20). This purification procedure was carried out with all the mutants used and similar results were obtained. The example with the T57C/T418C mutant is shown.

Atomic force microscopy (AFM) was also performed with the purified mutant enzymes. A sample containing 5% DNA-protein-DNA (Figure 21) is enriched over 70%; the results being similar with all the mutants used. This resulted in a significant increase in the fraction of correct fishing events in the optical tweezers. Before electroelution only about one fiber was trapped per day, but after electroelution this number increased to 7-10 fibers per day. It is important to point out that the mechanical and enzymatic properties of the proteins do not change after electroelution.

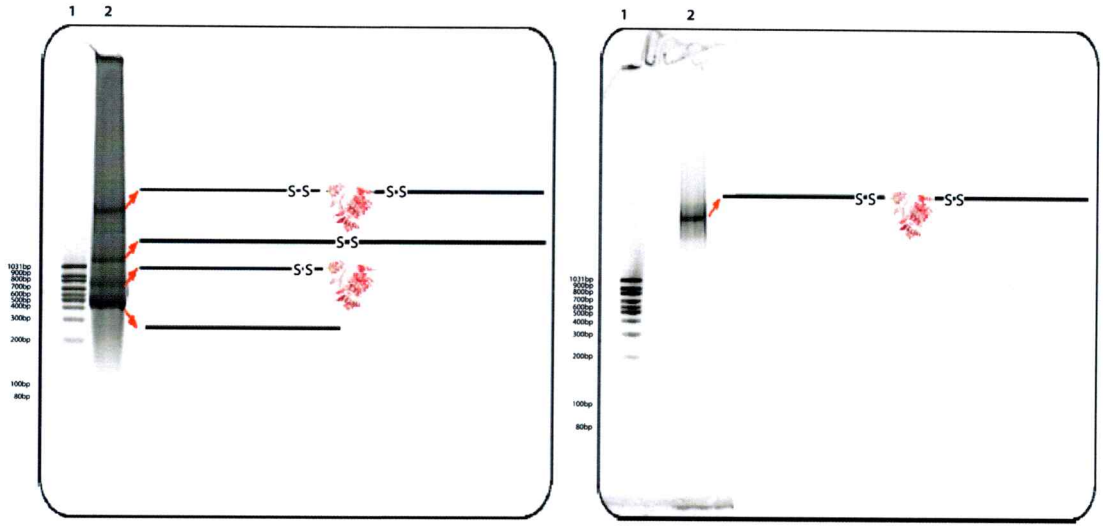
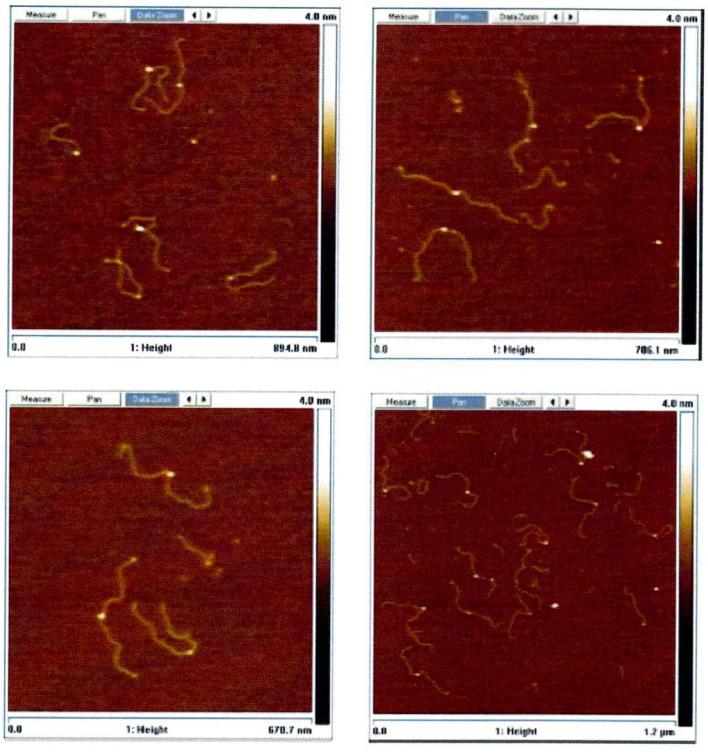


Figure 20. Purification by electroelution. Left. Gel that shows the reaction of DNA handles previous to the purification by electroelution, with a scheme that represents each species. Right. Gel after electroelution. The major DNA-protein-DNA species is observed. Gels were stained with sybr green.



Species	Percentage
	72.9%
	14.0%
	11.7%
	1.4%

Figure 21. AFM studies of the sample after DNA reaction and after electroelution. Left. Images of AFM taken in tapping mode. The different species obtained after the reaction with DNA and after electroelution. Right. Table that shows the percentage of each species; the major species is DNA-protein-DNA.

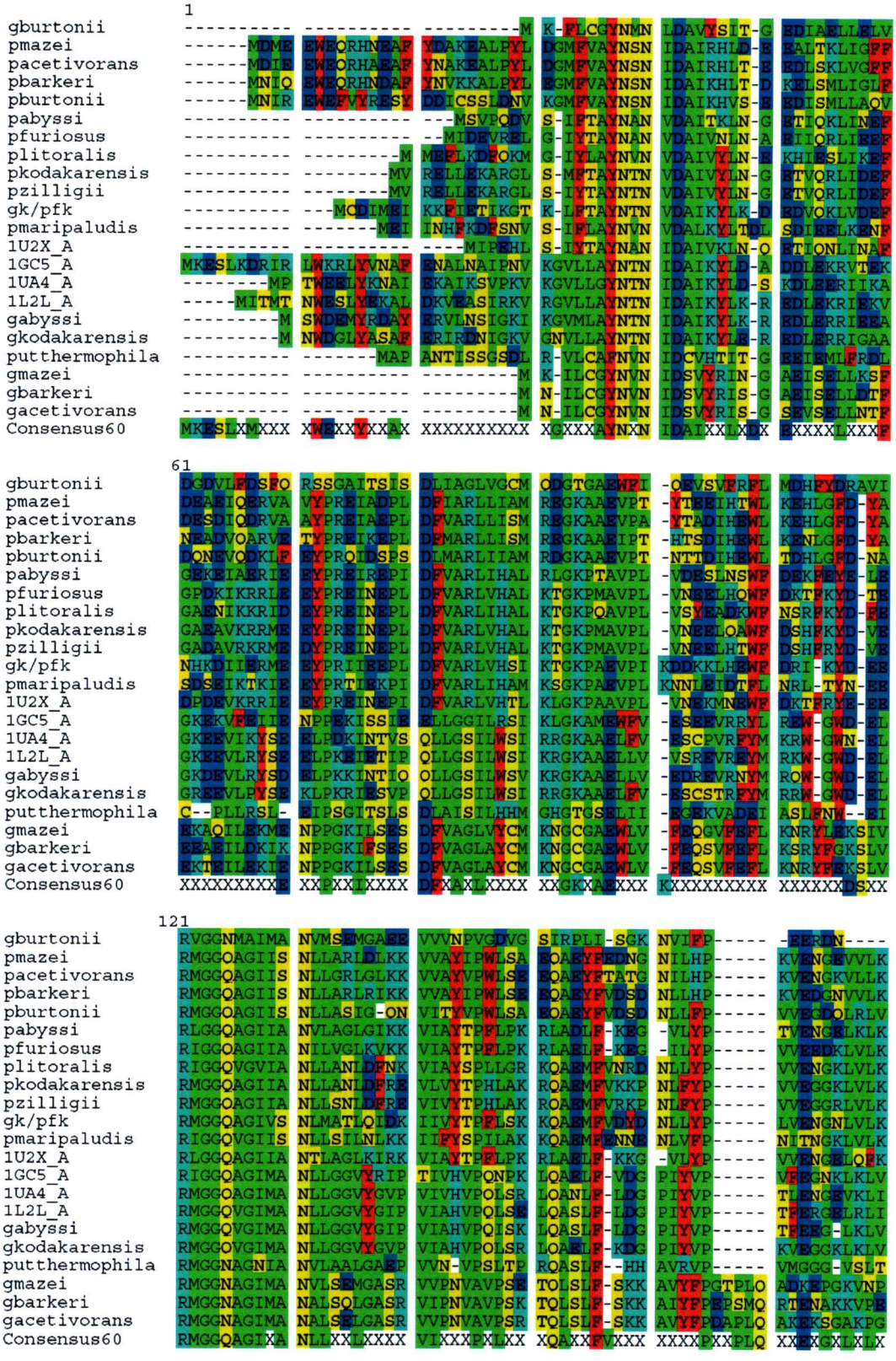
6. Mechanical unfolding of glucokinase and the effect of substrate

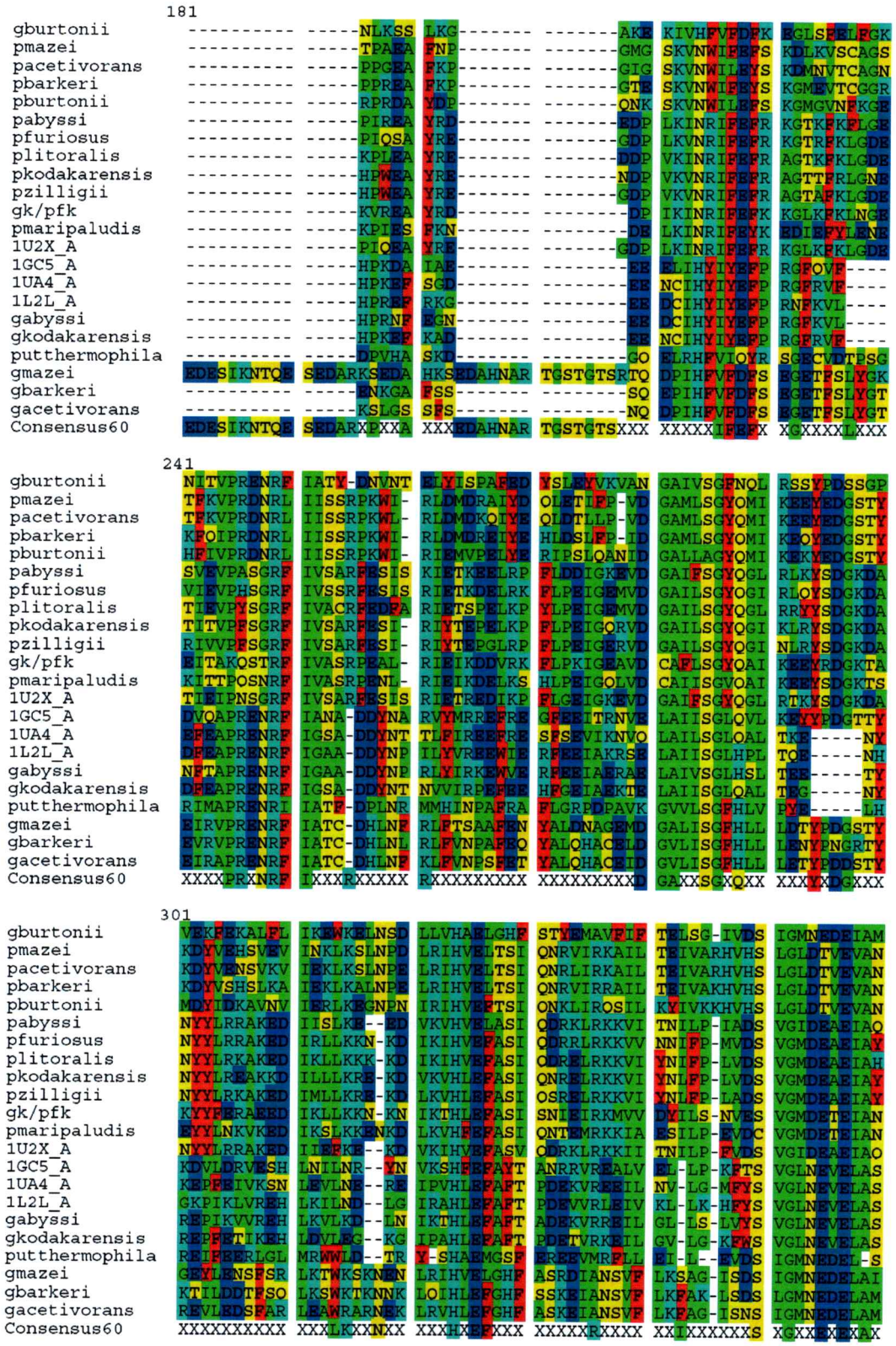
Choice of cysteine mutants

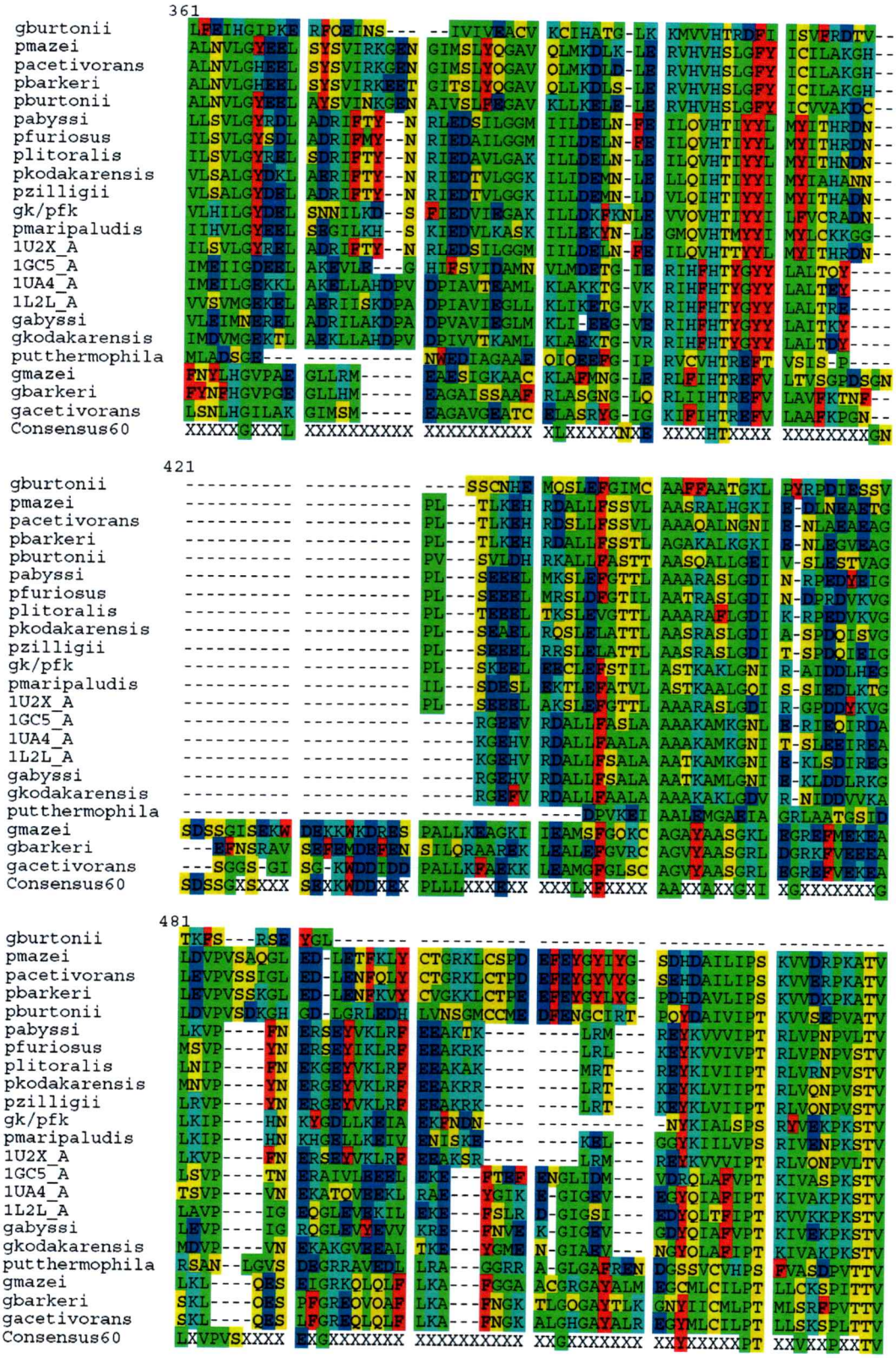
To choose the location of the cysteines in the protein several points were taken into account. The sites should be in zones that do not have a high degree of amino acid conservation, should be relatively exposed to the solvent so that the DNA handles can react, and should be in zones perpendicular to the hinge of the protein. A study done by Bertz and Rief (2009) with the maltose binding protein, shows that the substrate has no effect in the force required to unfold the enzyme when it is pulled from the N and C termini, but a clear effect is seen when it is pulled from a zone perpendicular to the hinge of the protein. The zones that were used to locate the cysteines were in each domain of the protein in zones perpendicular to the hinge. Moreover, we sought to find a pulling axis that would allow following the unfolding process at equilibrium; the effect of substrate at the mechanical level in conditions of equilibrium has not been carried out up to now.

Sequence alignments of proteins related with tlGK and cysteine mutations

After making an alignment of sequences and observing the positions of the residues that could be mutated to cysteines, according to the criteria mentioned above, the following residues were selected (Figure 22).


1
gburtonii
pmazei
pacetivorans
pbarkeri
pburtonii
pabyssi
pfuriosus
plitoralis
pkodakarensis
pzilligii
gk/pfk
pmaripaludis
1U2X_A
1GC5_A
1UA4_A
1L2L_A
gabyssi
gkodakarensis
putthermophila
gmazei
gbarkeri
gacetivorans
Consensus60
61
121




361
gburtonii
pmazei
pacetivorans
pbarkeri
pburtonii
pabyssi
pfuriosus
plitoralis
pkodakarensis
pzilligii
gk/pfk
pmaripaludis
1U2X_A
1GC5_A
1UA4_A
1L2L_A
gabyssi
gkodakarensis
putthermophila
gmazei
gbarkeri
gacetivorans
Consensus60
421
gburtonii
pmazei
pacetivorans
pbarkeri
pburtonii
pabyssi
pfuriosus
plitoralis
pkodakarensis
pzilligii
gk/pfk
pmaripaludis
1U2X_A
1GC5_A
1UA4_A
1L2L_A
gabyssi
gkodakarensis
putthermophila
gmazei
gbarkeri
gacetivorans
Consensus60
481
gburtonii
pmazei
pacetivorans
pbarkeri
pburtonii
pabyssi
pfuriosus
plitoralis
pkodakarensis
pzilligii
gk/pfk
pmaripaludis
1U2X_A
1GC5_A
1UA4_A
1L2L_A
gabyssi
gkodakarensis
putthermophila
gmazei
gbarkeri
gacetivorans
Consensus60

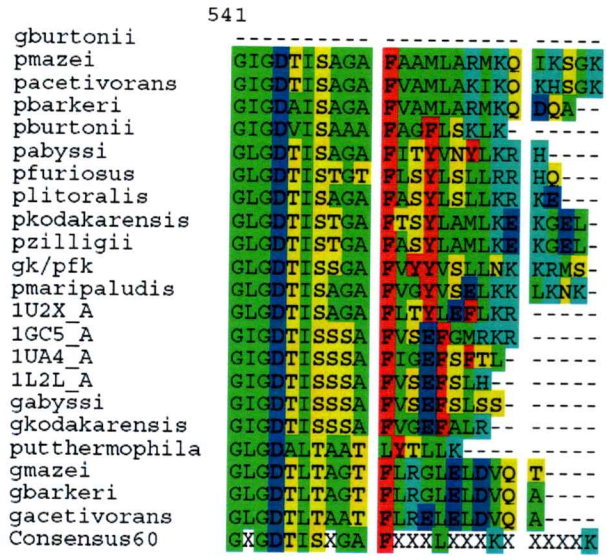


Figure 22. Sequence alignment of tlGK with other ADP-dependent glucokinases and phosphofructokinases. Alignment of ADP-dependent phosphofructokinases and glucokinases of archaeas was carried out. Names: gburtonii, glucokinase from *Methanococcoides burtonii*; pmazei: phosphofructokinase from *Methanosarcina mazei*; pacetivora: phosphofructokinase from *Methanosarcina acetivorans*; pbarkeri: phosphofructokinase from *Methanosarcina Barkeri*, pabyssi: phosphofructokinase from *Pyrococcus abysii*, pfuriosus: phosphofructokinase from *Pyrococcus furiosus;* plitoralis: phosphofructokinase from *Thermococcus litoralis*; pkodakaren: phosphofructokinase from *Thermococcus kodakarensis*; pzilligii: phosphofructokinase from *Thermococcus zilligii*; gk/pfk: glucokinase and phosphofructokinase from *Methanocaldococcus jannaschii*; pmaripalud: phosphofructokinase from *Methanococcus maripaludis*; 1u2x: phosphofructokinase from *Pyrococcus horikoshii*; 1ua4: glucokinase from *Pyrococcus furiosus*; 1l2l: glucokinase *Pyrococcus horikoshii*; 1gc5 glucokinase from *Thermococcus litoralis*; gabyssi: glucokinase from *Pyrococcus abysii;* gkodakaren:

glucokinase from *Thermococcus kodakarensis,* putthermop: phosphofructokinase from *Methanosaeta thermophila*; gmazei: glucokinase from *Methanosarcina mazei;* gbarkeri: glucokinase from *Methanosarcina Barkeri;* gacetivora: glucokinase from *Methanosarcina acetivorans.* Alignment carried out by Felipe Merino (University of Chile) in ClustalW (figure placed with his authorization). The program SeView (Gouy et al., 2010) was used to show the alignment. The color code is the following: nonpolar, aliphatic R groups (G, A, P, V, L, I and M) in green; aromatic R groups (F, Y and W) in red; polar, uncharged R groups (S, T, C, N and Q) in yellow; positively charged R groups (K, H and R) in light blue; and negatively charged R groups (D and E) in blue. Consensus60 (last row) is a consensus sequence, which shows residues that are identical in more than 60% on the listed sequences. The numbers are the enumeration of the amino acids corresponding to the tlGK (1GC5_A) sequence.

N-C terminal mutant: S4C and R467C

Mutations perpendicular to the hinge of the protein: S4C/T57C, T57C/T418C and Q197C/T418C

None of these mutants have a high degree of sequence conservation in Archae organisms at the sites where the cysteine mutations from the Archae at the sites selected.

Degree of exposure of the residues chosen to place the cysteines

The residues chosen to place the cysteines were all above 40% exposed to the solvent: S4, 47%; T57, 53%; Q197, 95%; T418, 87%; R467, 99% and their modification was feasible. Other candidate residues were exposed below 40% and could not bind the DNA handles: A26, 7% and S77, 29% (as observed in Figures 16 and 17)

7. Mechanical unfolding of tlGK

How can we know when we have a single molecule?

When performing a mechanical unfolding in the optical tweezers, we often worry if we really have only one trapped molecule between the 2 polystyrene beads, or if perhaps we have multiple tethers contributing to the force-extension behavior. To find out the answer, we use the properties of the DNA handles bound to the protein. Double-stranded DNA has very specific characteristics when it is stretched (Smith et al., 1996). One of them is that it suddenly elongates by "overstretching" at a force around 67 pN (Figure 23). This force can vary slightly (+/- 3 pN) with the content of CG/AT and with

the DNA length (Steve Smith, personal communication) .The overstretching force can also change slightly with the ionic strength of the buffer or with crosslinking of the strands. Overstretching is a highly cooperative transition from B-DNA to s-DNA (stretched) form (Smith et al., 1996; Wenner et al., 2002). The DNA goes on stretching at a constant force, usually at 67 pN, into a stable form with 5.8 angstroms rise per base pair, that is, 70% longer than B-form dsDNA. Another important point to consider is whether or not the correct molecule was trapped (DNA-protein-DNA). When this is the case, the DNA overstretching should correspond to 70% of the B-form length for 2 handles in series. This is of importance because if the protein is sticky and sticks to the polystyrene beads, only half of the DNA will stretch. In order to know if the construct has the correct length and if it contains one protein molecule, the properties of the DNA handles are also advantageous.

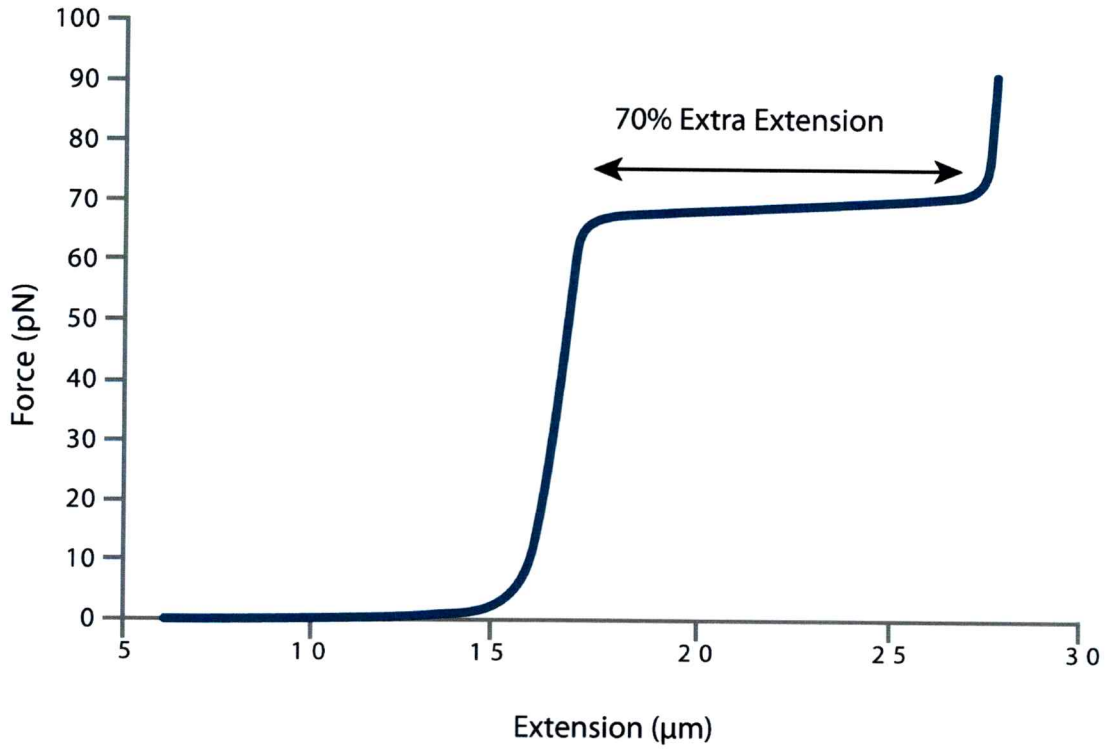


Figure 23. DNA force extension curve. Typical extension curve *versus* force of a DNA molecule. Around 67 pN, the overstretching of DNA is observed, it corresponds to 70% of the total extension of the B-form contour length DNA. The overstretching extension is used to determine that there is a single molecule, and allows knowing that we have the correct DNA molecule. Lambda DNA in 500 mM NaCl, 20 mM Tris, pH 8, at room temperature. Figure modified from Tinoco and Bustamante (2002)

Mechanical unfolding of the N-C termini mutant S4C/R467C

The protein termini mutant has a theoretical contour length of 166 nm because it has 463 amino acid residues. When unfolded under force it shows several intermediates. At least 5 species of different extension are observed, which may indicate 5 different intermediates. The unfolding force is about 10 pN, a value lower than that of many other proteins studied. The constant velocity experiments show a great dispersion, which may be due to various reasons. When the protein is allowed to refold to be unfolded again, many times it is observed that it does not pass through intermediates of the same extension as in the previous unfolding event.

At least in some cases, this effect may be due to the incorrect refolding of the protein. Studies were conducted where the waiting time given to the molecule to refold at low force is increased. Generally the waiting time was 5 sec, but increasing it from 45 sec to 5 min did not change the variety of intermediates significantly.

A jump study at constant force was done: a sudden change from low forces to high forces was exerted in order to observe the species that appeared. In all the experiments performed the total expected length was never obtained, indicating that we could be starting from some intermediate species that the enzyme populates under force.

We could not determine the effect of ADP binding on the stability of the protein using force jumps in this particular mutant, because of the dispersion of the traces.

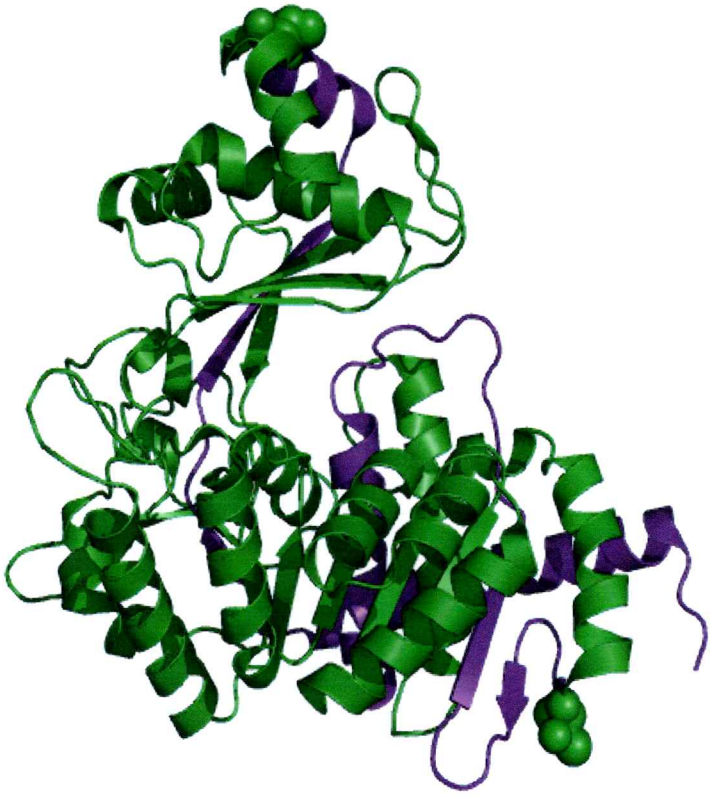


Figure 24. Stretching sites in the T57C/T418C mutant. In the ribbon tlGK representation the amino acids between residues 57C and 418C are in green, the rest of the structure is in violet. The side chains of amino acids T57 and T418 are highlighted with spheres.

Because of the complexity of the traces and the dispersion of the results obtained in constant velocity and forces jumps experiments, the effect of the substrate on the stability of the protein could not be analyzed.

Mechanical unfolding of the T57C/T418C mutant

The T57C/T418C mutant has a theoretical contour length of 129.96 nm corresponding to the 361 amino acid residues from where the protein is pulled (Figure 24).

The experiments at constant velocity with this mutant show that there are two extension intermediates upon mechanical unfolding (Figure 25). One of these corresponds to 80% of the total length of the structure and the other to 20% of the total length of the tlGK under this particular attachment points for pulling. In this mutant we are not taking into account great part of N and C termini of the protein (they are not between the DNA handles), and the total extension of unfolding in this mutant is reached, therefore it may indicate that parts of the N and/or C termini of the protein are unfolded at very low forces and when pulling the complete protein (S4C/R467C) we do not see the total expected unfolding. This mutant shows the total expected unfolding through the adjustment to the WLC model.

When adding the ADP substrate the unfolding intermediary that corresponds to 20% of the total extension of the pulling axis given by the WLC model is not observed (Figure 26 right). Maybe this portion of the protein is protected by the substrate and it

cannot unfold because the appropriate forces are not generated by the optical tweezers (one can get up only to 67 pN, where the DNA overstretching occurs).

To determine the length unfolded, the rip is bound between two WLC curves. As seen in Figure 27 left, if the parameters of the complete mechanical unfolding of the T57C/T418C mutant are used, there is no good fit (persistent length of 0.65 nm, contour length of 129.96 nm, folded distance 5 nm). The fit was done taking into account 80%, 40% and the 20% of the total contour length and it adjusts much better to the 3 cluster groups of rips (Figure 27 right).

The same analysis was done in presence of ADP, showing that the 40% species is proportionally increased, relative to the other species (Figure 28).

To more clearly visualize the species that are generated when stretching the T57C/T418C mutant, a histogram was constructed that shows the number of rip events in relation to the relative distance, normalized by the WLC with the parameters of complete unfolding of the protein (Figure 29). This figure shows that the majority of the species have an extension of 80% and 40%; the species corresponding to an extension of 20% is less populated.

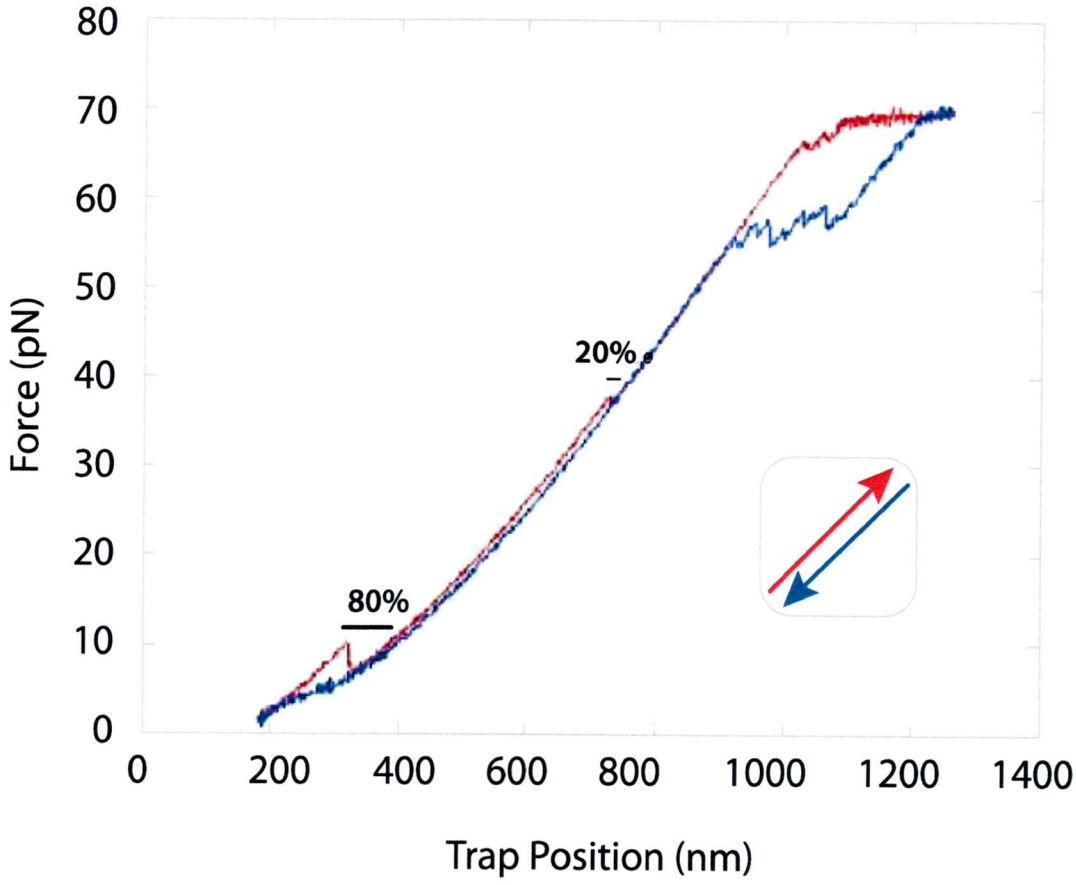


Figure 25. Extension *versus* force curve for the T57C/T418C mutant. An extension curve *versus* force of the T57C/T418C mutant is illustrated. The 80% and the 20% intermediates in the absence of ADP are observed. The stretching is shown in red, and the relaxing in blue.

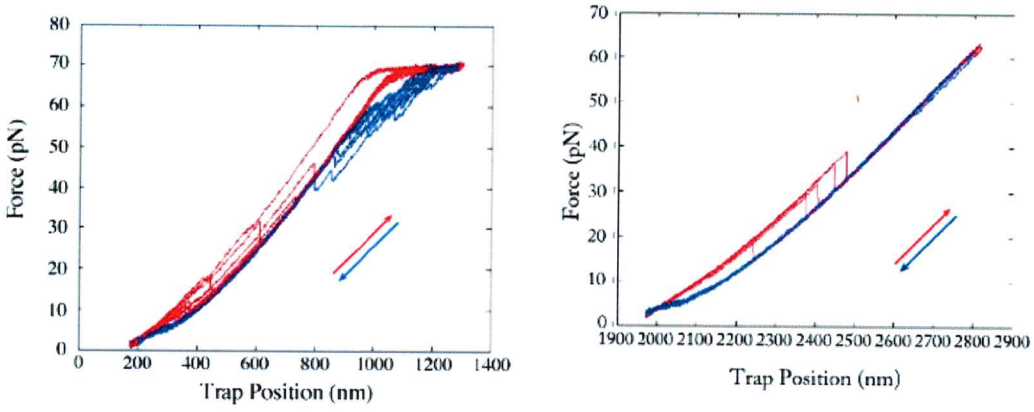


Figure 26. Extension *versus* force curve of the T57C/T418C mutant with and without ADP. Left. Extension curves *versus* force of the T57C/T418C mutant in absence of ADP with overstretching of the DNA. Right. Extension curves versus force of the T57C/T418C mutant in the presence of ADP. The stretching is shown in red and the relaxing of the curves in blue. Several unfolding-refolding events are shown for the same fiber.

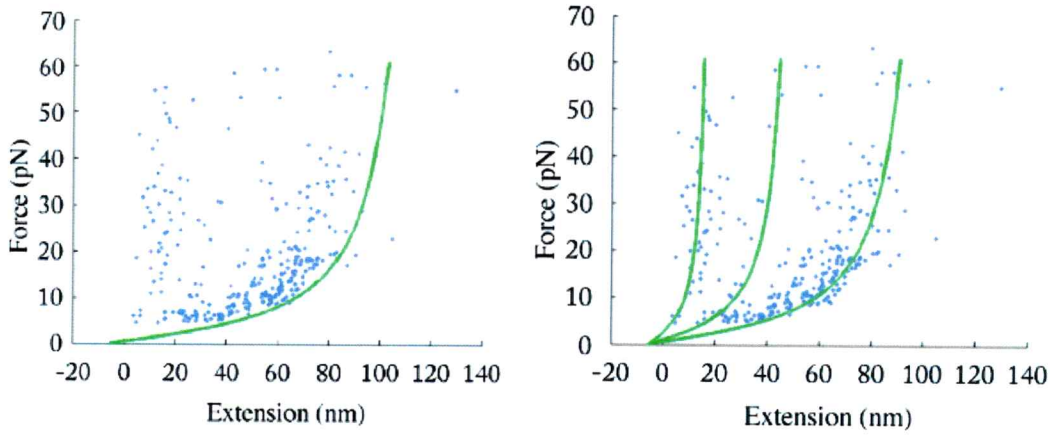


Figure 27. Analysis of the unfolding rip of the T57C/T418C mutant. From the extension curves versus force, the lengths of the unfolding events were graphed in order to determine the extension of the event and the force at which it occurred. Left. The curve that corresponds to the theoretical WLC is shown, starting from the total length of the protein. A persistent length of 0.65 nm was used, a folded distance of 5.4 nm and a contour length corresponding to 361 amino acid residues of 130 nm (0.36 nm/amino acid residue). Right. The WLC was adjusted to different contour lengths, keeping the other parameters constant. The curve to the right corresponds to a contour length of 103 nm, the one in the middle to 52 nm and the one on the left to 26 nm.

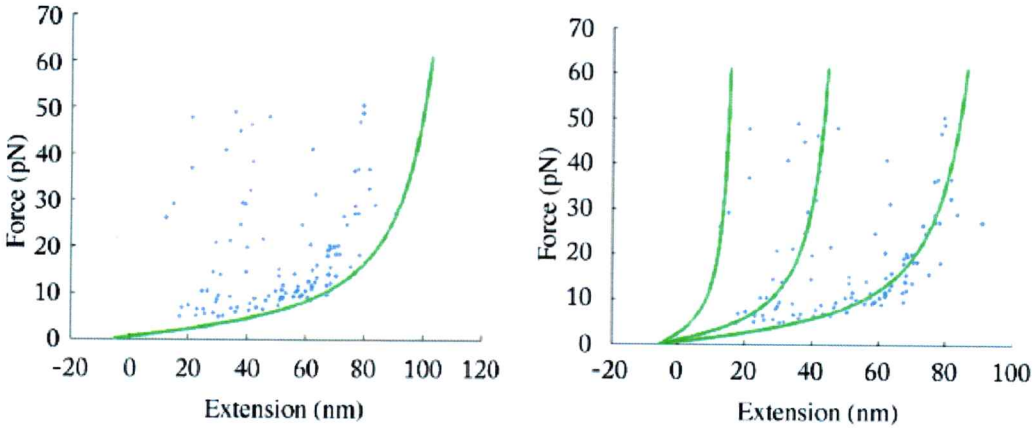


Figure 28. Analysis of the unfolding rip of the T57C/T418C mutant in the presence of ADP. Legend as in Figure 27. The concentration of ADP used was 1 mM.

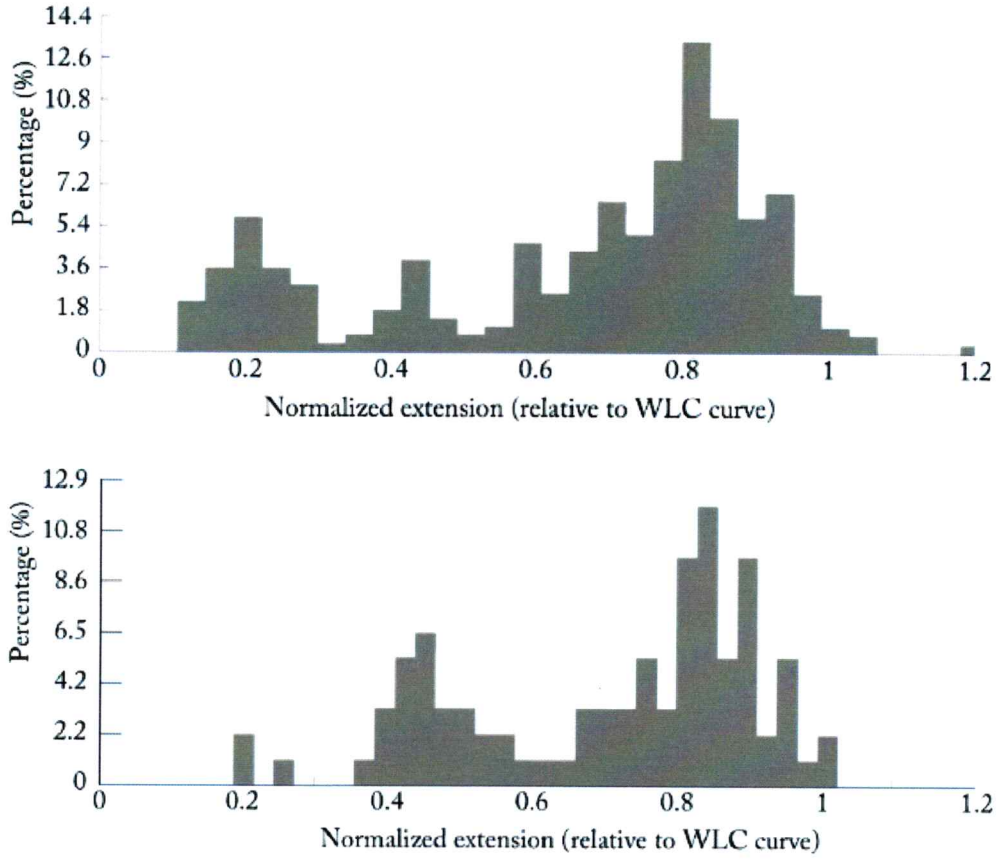


Figure 29. Histograms of the normalized extension of the unfolding events to the WLC curve. Every extension event was normalized, taking the total value of the extension, to the WLC curve with the parameters of the protein from the bonding points of the DNA handles. A persistent length of 0.65 nm was used, and a folded distance of 5.4 nm and a contour length corresponding to 361 amino acid residues of 130 nm (0.36 nm/amino acid residue). Above. The histogram of the normalized extension in the absence of ADP. The majority of the species have an extension of 80%, 40% and 20% of the corresponding total extension of the protein. Below. The histogram of the normalized extension in the presence of ADP.

Mechanical unfolding of the S4C/T57C mutant

The S4C/T57C mutant has a contour length of 19.08 nm corresponding to the 53 amino acid residues emcompased by the pulling locations (Figure 30).

At constant velocity only a reversible unfolding at 28 pN is observed (Figure 31) which corresponds to about 5.5 nm. The total expected length for this mutant is 9 nm at 28 pN. If the initial alpha helix of the protein were unstructured it gives an extension of 5.5 nm, so it is possible that this helix was unstructured already by the time the DNA handles were placed (only one unfolding event is observed). Also, this initial alpha helix is not conserved in proteins with high degree of sequence identity nor in proteins related phylogenetically (Figure 22). In addition, the activity gel shows that this mutant enzyme is active with the DNA handles (Figure 16). Moreover, this mutant was also generated without the initial alpha helix; although it could be only partially purified, it showed ADP-dependent glucokinase activity.

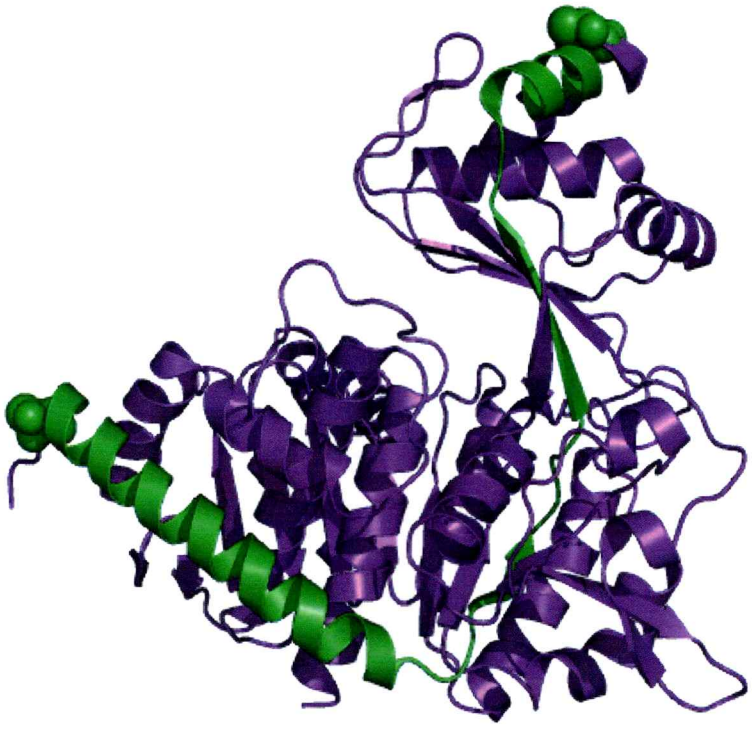


Figure 30. Stretching sites of the S4C/T57C mutant. In the ribbon representation of tlGK, the amino acid residues between the residues 4C and 57C are in green, the rest of the structure is in violet. The amino acids S4 and T57 are highlighted with spheres the side chains.

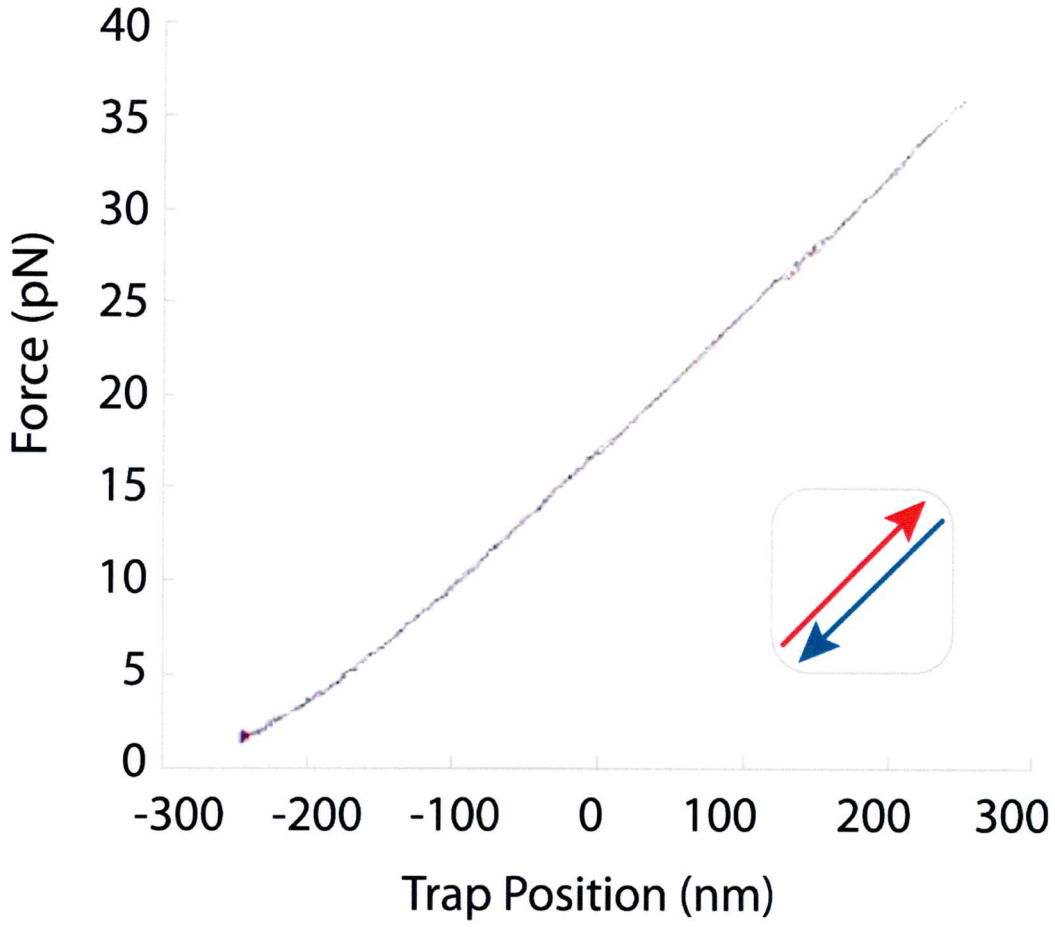


Figure 31. Extension *versus* force curve for the S4C/T57C mutant. An extension curve *versus* force of the S4C/T57C mutant is shown, where around 28 pN there are reversible unfolding events. The stretching is shown in red, and the relaxing in blue.

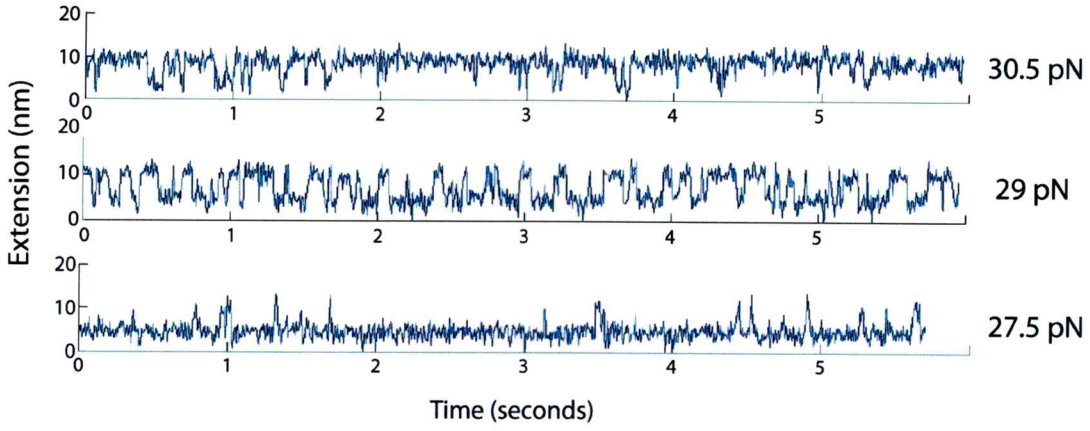


Figure 32. Extension *versus* time of the S4C/T57C mutant at different forces. The hopping of the S4C/T57C mutant is shown at three different forces. At high forces (30.5 pN) the protein spends more time in the state of greater extension. Around 29 pN the species spends 50% of the time in each state and at 27.5 pN most of the time in the state of lower extension.

Through the studies at constant force of the S4C/T57C mutant we may quantify the dependence of the protein stability on the applied external force. We may also obtain kinetic data and the equilibrium constant. When the unfolding transitions of proteins occur at equilibrium, a phenomenon called "hopping" is observed, which means that at a same force the protein may explore two or more structures. In the case of the tlGK two states are explored and we call them folded and unfolded states (Figure 32).

The time the protein spends in each state was quantified and a graph of time *versus* the number of counts that the enzyme spends in a given state for each time interval was constructed (Figure 33). The best fit to a single exponential curve was obtained at each force. The inverse of time allows us to get the rate constants (k) (Figure 33). Also, a population histogram that shows the number of events in which the enzyme was found in each state was constructed (Figure 34).

For each particular fiber (or DNA-protein-DNA construct) this analysis was done to determine the kinetic constants at each force. Also, for each fiber the kinetics was determined at different forces. In order to establish the robustness of the values of kinetic constants obtained by the constant force experiments, the same fiber was allowed to refold under low force (5 pN) and it was then subjected to high forces (around 28pN), to observe the hopping kinetics. The variability among fibers and in one same fiber is low.

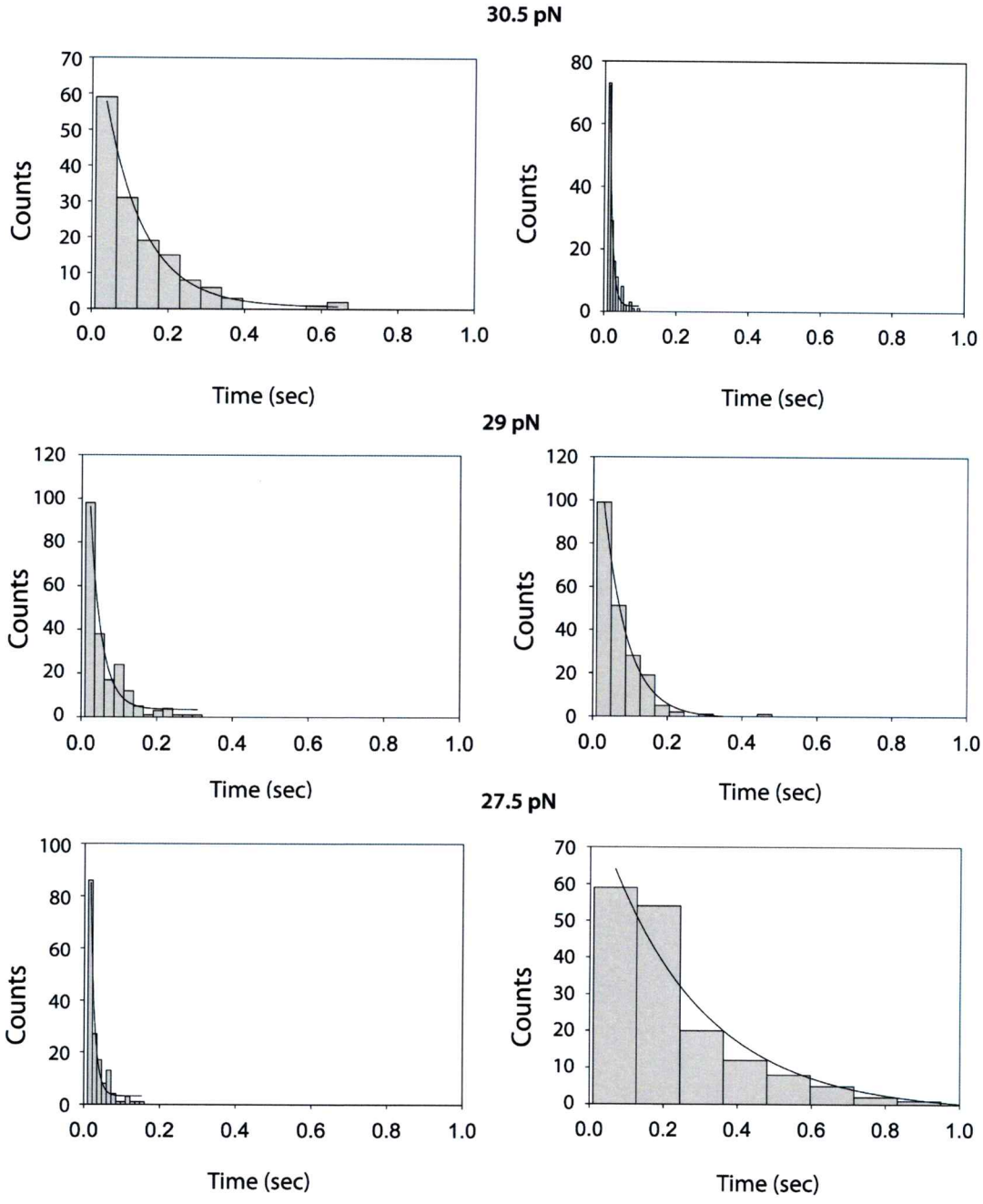


Figure 33. Determination of the unfolding and refolding rate constants at different forces. The time spent in each transition was determined. The graph of counts *versus* time for each transition is shown. Left. Refolding transitions. Right. Unfolding transitions. The data was fit to a single exponential curve.

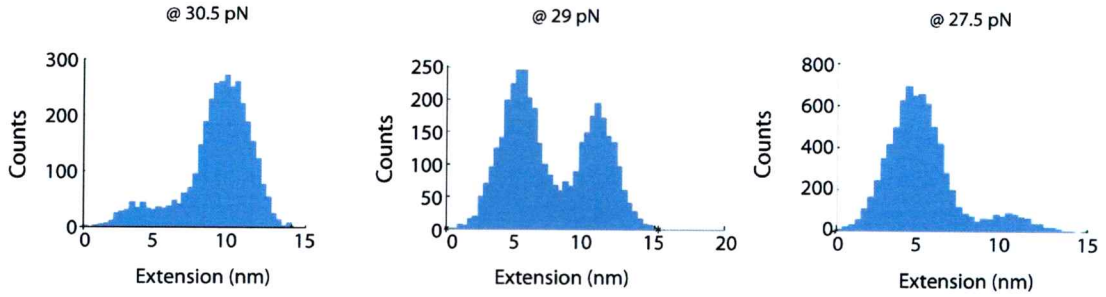


Figure 34. Histograms of population of the species of greater and lesser extension for the hopping of the S4C/T57C mutant at different forces. The histograms of species population of lesser extension and greater extension were carried out at different forces. At low forces (27.5 pN), the more populated species is the one of lesser extension; at greater forces (30.5 pN) that of greater extension. At around 29 pN the populations are similar.

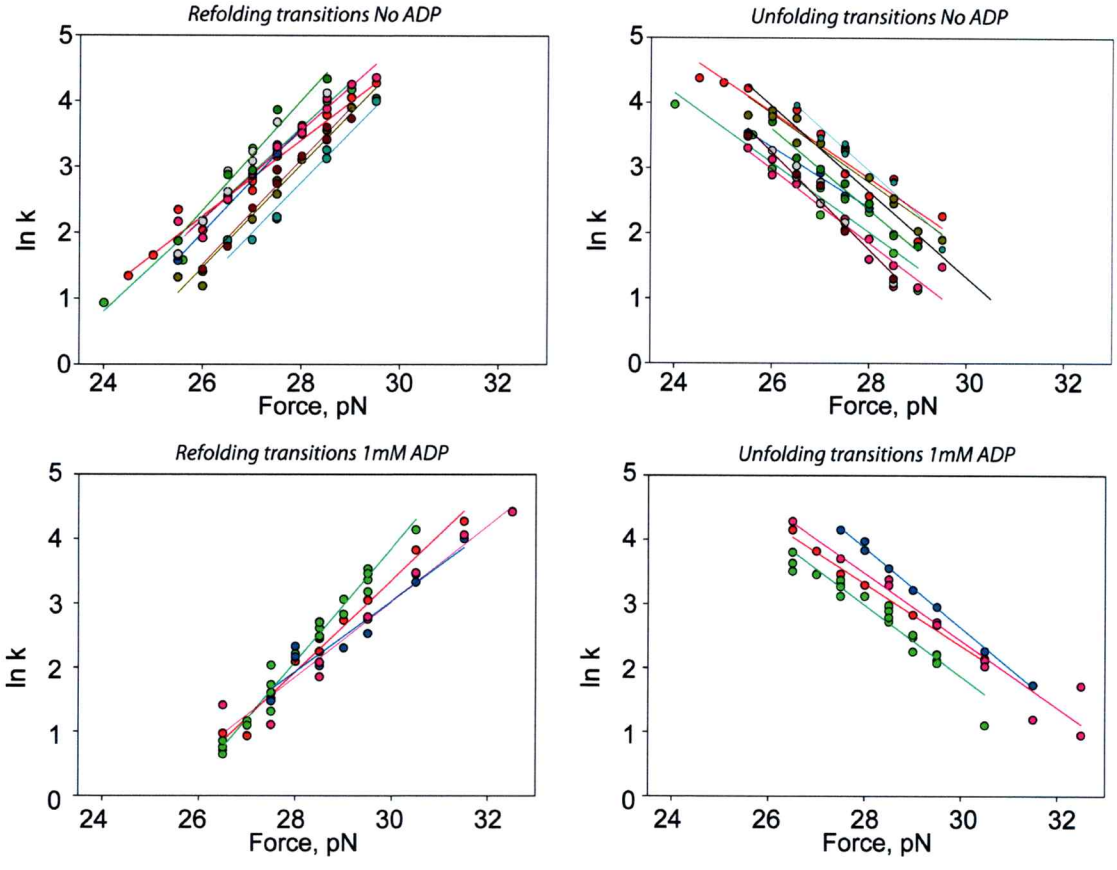


Figure 35. Effect of the force on the velocity constant for the S4C/T57C mutant in the presence and in the absence of ADP. Above Left. Unfolding transitions of different fibers of the S4C/T57C mutant. Right. Refolding transitions of different fibers of the S4C/T57C mutant Below. The transitions in the presence of ADP. Left, unfolding. Right, refolding.

Utilizing Bell's (1978) equation we may determine the distance to the transition state, if we know the dependence of the rate constants with the force:

$k=k_0e(\Delta x^{\neq}F/k_BT)$

If the equation is linearized we obtain: $\ln k = \ln k_0 + (\Delta x^{\neq}F/k_BT)$

Each fiber has small variations, since we are working *in singulo*, therefore each different enzyme molecule will be called a fiber, corresponding to a distinct DNA-protein-DNA set. We will see the typical variations of each fiber. In the case of the tlGK we observe that the variations are slight, when comparing fiber to fiber (Figure 35, Table 1).

Table 1. Variability from fiber to fiber in the transition state distance values in the absence and presence of ADP

A. Unfolding transition state distances in the absence of ADP

Fiber name	Slope ± SD	Transition state distance ± SD (nm)	Color of fiber in plot on Figure 35
Fiber 1	0.6997 ± 0.0469	2.87 ± 0.19	Green
Fiber 2	0.5817 ± 0.0306	2.38 ± 0.13	Red
Fiber 3	0.7984 ± 0.0309	3.27 ± 0.13	Blue
Fiber 4	0.6790 ± 0.0329	2.78 ± 0.13	Pink
Fiber 5	0.7950 ± 0.0494	3.26 ± 0.20	Gray
Fiber 6	0.8356 ± 0.0754	3.42 ± 0.31	Dark red
Fiber 7	0.8395 ± 0.1006	3.44 ± 0.41	Dark green
Fiber 8	0.7815 ± 0.0329	3.20 ± 0.13	Dark yellow
Fiber 9	0.7750 ± 0.0702	3.18 ± 0.29	Dark cyan

Mean distance to the unfolding transition state = 3.09 ± 0.35 nm

B. Refolding transition state distances in absence of ADP

Fiber name	Slope ± SD	Transition state distance ± SD (nm)	Color of fiber in plot on Figure 35
Fiber 1	0.5354 ± 0.0582	2.20 ± 0.24	Green
Fiber 2	0.5062 ± 0.0456	2.08 ± 0.19	Red
Fiber 3	0.4665 ± 0.0388	1.91 ± 0.16	Blue
Fiber 4	0.5669 ± 0.0439	2.32 ± 0.18	Pink
Fiber 5	0.7571 ± 0.0639	3.10 ± 0.26	Gray
Fiber 6	0.7463 ± 0.07495	3.06 ± 0.31	Dark red
Fiber 7	0.6240 ± 0.03891	2.56 ± 0.16	Dark green
Fiber 8	0.5263 ± 0.04071	2.16 ± 0.17	Dark yellow
Fiber 9	0.6664 ± 0.0648	2.73 ± 0.27	Dark cyan

Mean distance to the refolding transition state = 2.46 ± 0.43 nm

C. Unfolding transition state distances in presence of ADP

Fiber name	Slope ± SD	Transition state distance ± SD (nm)	Color of fiber in plot on Figure 35
Fiber 1	0.4843 ± 0.0243	1.99 ± 0.10	Red
Fiber 2	0.5570 ± 0.0276	2.28 ± 0.11	Green
Fiber 3	0.6224 ± 0.0157	2.55 ± 0.06	Blue
Fiber 4	0.5251 ± 0.0431	2.14 ± 0.18	Pink

Mean distance to the unfolding transition state 2.24 ± 0.24 nm

D. Refolding transition state distances in presence of ADP

Fiber name	Slope ± SD	Transition state distance ± SD (nm)	Color of fiber in plot on figure 35
Fiber 1	0.7156 ± 0.0393	2.93 ± 0.16	Red
Fiber 2	0.8854 ± 0.0269	3.63 ± 0.11	Green
Fiber 3	0.5516 ± 0.0696	2.26 ± 0.29	Blue
Fiber 4	0.5899 ± 0.0407	2.42 ± 0.17	Pink

Mean distance to the refolding transition state 2.81 ± 0.62 nm

The comparison of the force dependence of the kinetics of the S4C/T57C mutant was done in the presence and in the absence of ADP. For the unfolding reaction, the distances to the transition state were $x^{\ddagger} = 3.09 \pm 0.35$ nm in the absence and 2.81 ± 0.62 nm in the presence of ADP, respectively. The corresponding values during refolding were 2.24 ± 0.24 nm and 2.46 ± 0.43 respectively. The sum of the unfolding transition state and the refolding one gives us 5.33 nm in the absence of ADP and 5.27 nm in its presence, the same value of the total extension observed at forces around 28 pN.

The kinetics of unfolding with and without ADP show a greater difference (Figure 36) than the kinetics of refolding (Figure 37), as revealed by the velocity constant at zero force. This means that the substrate is disturbing the folded state of the enzyme and, in a much less significant way, the unfolded state. Moreover, the slopes of plots of the logarithm of rate constant *versus* force remain constant with and without ADP, showing that the distance to the transition state is similar in both cases.

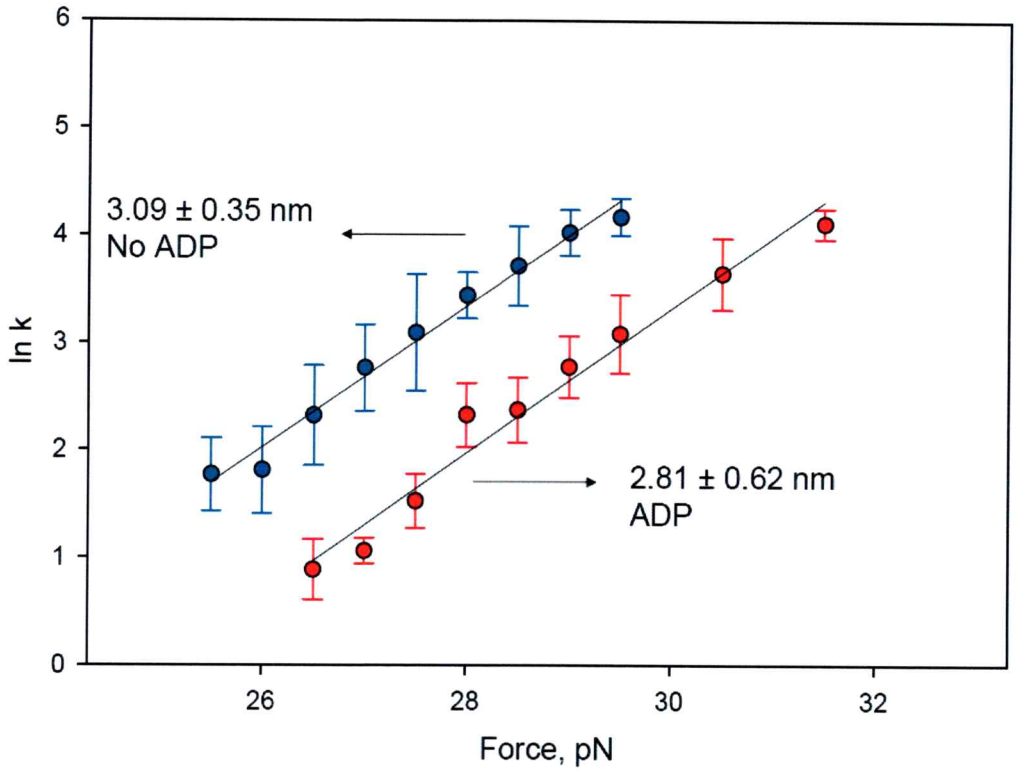


Figure 36. Determination of the distances to the transition state in the unfolding transitions. The Bell equation was used ($\ln k = \ln k_0 + \Delta X^{\neq} F/k_B T$) to determine the distance to the transition state. The distances to the transition state do not change in the presence of (red curve) or absence (blue curve) of ADP.

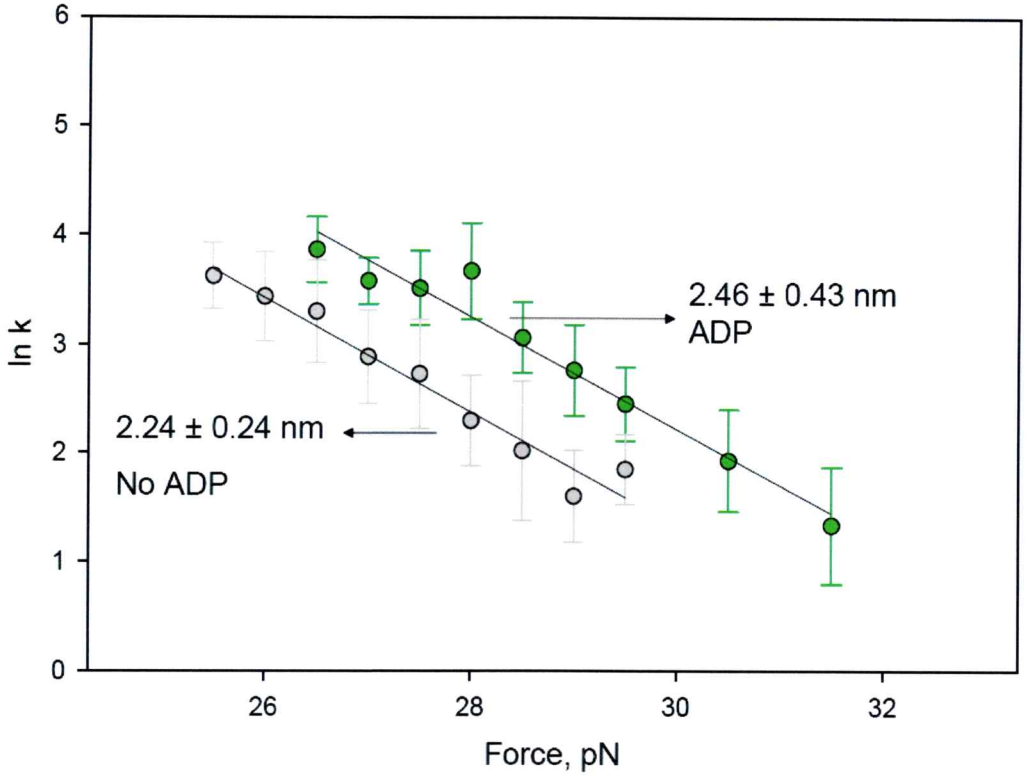


Figure 37. Determination of the distances to the transition state in the refolding transitions. The Bell equation was used ($\ln k = \ln k_0 + \Delta X^{\neq} F/k_B T$) to determine the distances to the transition state. The distances do not change in the presence (green curve) or the absence (gray curve) of ADP.

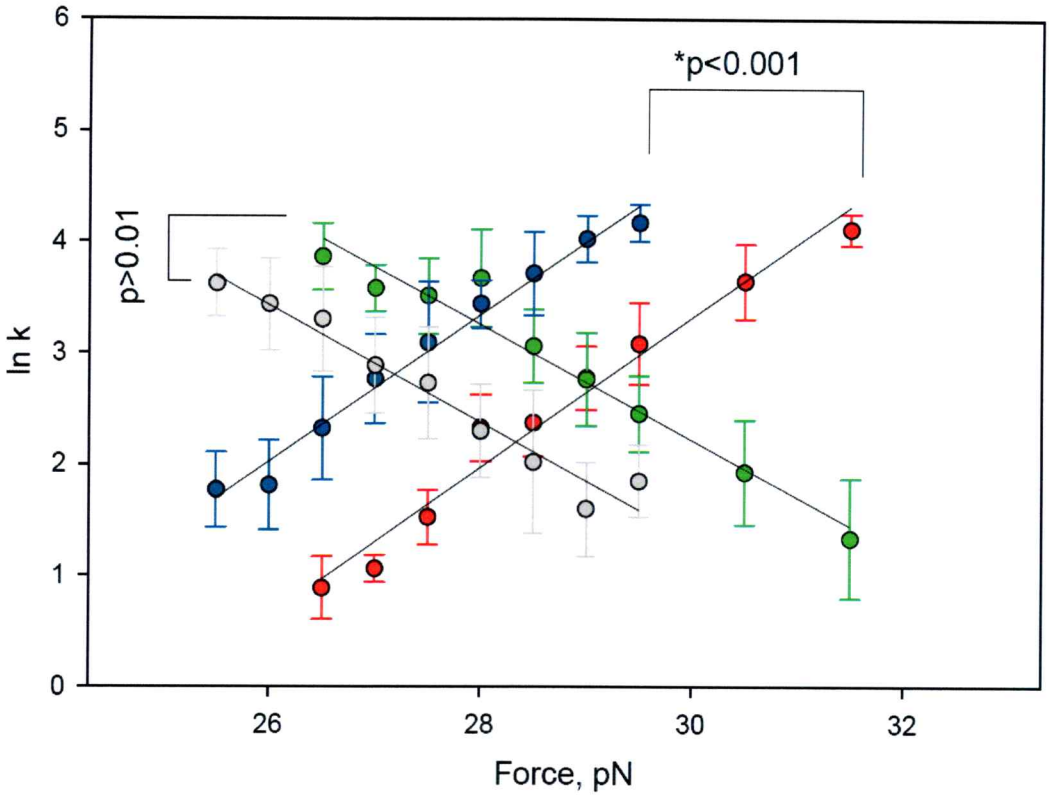


Figure 38. Relation between the velocity constant and the force for the hopping of the S4C/T57C mutant. The Bell equation was used ($\ln k = \ln k_0 + \Delta X^{\neq} F/k_B T$) to determine the distance to the transition state for the unfolding (blue without ADP and red with ADP) and refolding reactions (green without ADP and gray with ADP) and the rates of unfolding and refolding were extrapolated to zero force. A t-Student analysis was performed to verify if the data are significantly different in relation to the transitions.

Table 2. Statistical significance of the differences in the unfolding and refolding distances to the transition states.

t-Student of the data with and without ADP

Force (pN)	P (between $x_u^{\mp}$ ADP/ $x_u^{\mp}$ NO ADP)	P (between $x_f^{\mp}$ ADP/ $x_f^{\mp}$ NO ADP)
26.5	0.0000134788	0.0220820929
27	0.0000053677	0.0186261145
27.5	0.0000000092	0.0003030295
28	0.0000187696	0.0002698618
28.5	0.0000000021	0.0000305546
29	0.0000197905	0.0013781247
29.5	0.0002439856	0.0143052147
	All $p<0.001$	$p>0.01$

The statistical study between the values of k_u and k_f with and without substrate shows that the differences in k_u are more significant than those in k_f (Figure 38, Table 2). The extrapolation to zero force of k_u and k_f in the presence of ADP were 5×10^{-8} and 3×10^{-7} s^{-1}, respectively

Specificity of the ADP binding

To discard any effect of unspecific binding, we used another nucleotide, UTP, to observe the mechanical effect. The kinetic parameters obtained in the presence of UTP do not show significant differences with the ones determined in its absence.

The substrate glucose was also used. No difference is observed with glucose either in the fibers either with or without ADP.

Thermodynamic analysis of the mechanical unfolding of S4C/T57C in the presence and in the absence of ADP.

To determine the ΔG of protein stability in the absence and in the presence of ADP, the procedure described in Materials and Methods was used, based on Liphardt et al. (2001).

For this purpose, the value of $F_{1/2}$ is used, which was determined as the intersection in the graphics of lnk *versus* force. In the absence of ADP the value of $F_{1/2}$ is 27.1 pN, while in its presence is 29.1 pN. The Δx of extension to that force was around 5.5 nm.

By means of the equation

$\Delta G^{o} = F_{1/2} * \Delta x(F_{1/2}) - \Delta G(F)\text{stretching}$

The value of ΔG^{o} can be obtained.

By means of the equation $\Delta G^{o} = F_{1/2} * \Delta x(F_{1/2}) - \Delta G(F)\text{stretching}$ the value of ΔG^{o} can be obtained.

The value of ΔG(F) stretching (unfolded state stretching) can be obtained determining the area under the curve in a graphic of extension *versus* force, through fitting of the data to the WLC model of polymer elasticity.

To determine the ΔG(F) stretching, it is important to know the proper values to adjust the WLC model and determine the area under the curve. The folded distance of

the S4C/T57C mutant should be of 9.18 nm, assuming that the initial alpha helix is extended due to the effect of the DNA handles. This distance was determined from the crystalline structure of the tlGK (Ito et al., 2001, PDB 1GC5). The length of the extended alpha helix is 3.54 nm; this helix extends as far as the residue A26. From the residue A26 up to the T57C there is a fully extended length distance of 5.64 nm, giving a total of 9.184 nm. A persistent length of 0.65 nm, and a contour length of 19.08 (equivalent to 53 amino acid residues) were used. The area under the curve of the WLC was determined from zero force up to the $F_{1/2}$ force, where 50% of the unfolding and refolding occurs.

In the case of the mutant in the absence of ADP, the $F_{1/2}$ is 27.1 pN, giving a $\Delta G(27.1)$ stretching of 71.26 pNnm

In the case of the mutant in the presence de ADP, the $F_{1/2}$ is 29.1 pN, giving a $\Delta G(29.1)$ stretching of 76.93 pNnm

ΔG^{o} in absence of ADP= 27.1x5.5 – 71.25 = 77.79 pNnm = 11.2 kcal/mol

ΔG^{o} in presence of ADP= 29.1x5.5 – 76.93 = 83.11 pNnm = 12.0 kcal/mol

DISCUSSION

Experimental setup

The experimental setup used was similar to that of Cecconi et al (2008). This setup is quite adequate to study the effect of the substrate in the mechanical stability of tlGK enzyme. Particularly informative was the study of the S4C/T57C mutant, which shows an equilibrium behavior at 28 pN with an extension of 5.5 nm. Other techniques, like magnetic tweezers, do not have the spatial resolution to observe an extension of 5.5 nm at that force, the same as for the Atomic Force Microscope where the resolution at low forces drops drastically (Neuman and Nagy, 2008).

Moreover, the equilibrium condition occurs at forces smaller than the rupture force of the interaction of anti-digoxigenin and digoxigenin (around 40 pN), making this phenomenon perfectly accessible to the constant force technique. Although the rupture force for digoxigenine-anti-digoxigenine is lower than the DNA overstretching force, we were able to observe the DNA overstretching because there is a distribution of rupture forces (Smith et al., 1996).

It is essential to show that we are working with a single protein molecule. Also, in some cases, some fibers were more resistant than others, which may mean that the antibodies of digoxigenin are polyclonal.

Proteins with handles are active

It has not been previously demonstrated directly that proteins kept with DNA handles are active. The DNA handles are quite large compared to the protein, and could

induced protein unfolding. In this work, by using a novel in-gel activity assay, we demonstrate that proteins are active with the handles and show that the in gel assay is more informative than the simple observation of DNA staining. This method may be extensible to several other enzymes, and can be a general approach to determine the enzymatic activity of modified proteins in-situ.

This method is also useful to observe the presence of dimers or monomers after the DNA derivatization. This is of importance because binding is performed by means of cysteines; many proteins may generate disulfide bridges between them and generate oligomers. This is why it was important to find the specific conditions to generate monomeric species when derivatizing with DNA. The protein is reduced at first with DTT, and later this agent is removed and the protein is derivatized with DTDP to protect the cysteines and make them more reactive to DNA. Excess of DTDP is used, but if the protein does not derivatize quickly with DTDP once the DTT is removed, the oligomeric species may start to form. Therefore, DTDP has to be in excess and in the same column which will be useful for buffer exchange so as totally derivatize the protein with DTDP. Also, if the protein is left in DTDP for a long time, oligomeric species start to form, so two-hours is the optimum time to derivatize with DNA.

Purification of the DNA-protein-DNA fibers

Generally, to trap a protein fiber, one could spend a whole day fishing in the optical tweezers, getting none or very few fibers. Usually there is DNA-DNA fishing, and the possibility of observing false positive is, therefore, high. It is important to

mention that the digoxigenin interaction with anti-digoxigenin has an unfolding force of about 40 pN, and because the antibodies against digoxigenin are polyclonal, certain artifacts can be observed. About 8-10% hopping around 8 pN of an extension of about 20 nm is observed in fibers that are DNA-protein-DNA as well as in fibers that are only DNA-DNA. This artifact can be attributed to the interaction of antidigoxigenin and digoxigenin and it does not occur in 100% of the cases because the antibodies are polyclonal. This is the only artifact that is observed in a consistent manner. Also, the fact that the antibodies are polyclonal makes some fibers to be more resistant than others, and one may observe the overstretching of the DNA in a better way, assuring that one is dealing with a single protein. It is possible go to higher forces and observe the overstretching at 67 pN, so one can be certain to be dealing with just one molecule. If there are two molecules, the overstretching force will increase.

Electroelution makes it possible to purify one of the bands of one gel in a selective way. Generally, extracting the protein out of the gel is quite difficult by means of mashing the acrylamide gel. On the other hand, electroelution consists in extracting the protein of the gel using an electric field. The protein, after the electroelution, remains active with the DNA handles and the purification of the protein increases considerably. This means that once it is taken to the optical tweezers, the efficiency for fishing the DNA-protein-DNA complex is significantly better than what one can obtain with an unpurified sample. One additional advantage is that this approach completely eliminates the DNA-DNA product.

Summing up, the purification allows a considerable enrichment of the DNA flanked protein and a corresponding increase in data collection.

Derivatization with cysteines

Mutations were chosen based on several parameters. A sequence alignment was made in order to determine the residues that are not conserved. Only those mutations whose degree of cysteines exposure was over 40%, based on the crystalline structure, were able to be derivatized with DNA. It is interesting to mention that the mutants could be derivatized with DTDP (first step for the DNA handles attachments), since it is a smaller molecule, but the derivatization with DNA did not work, showing a steric problem for DNA binding, and confirming that over 40% of exposure was needed for the derivatization. Two types of mutants were chosen, those that go from the N to the C terminal, and those whose cysteines are placed in opposite positions to the hinge of the protein. It is interesting to mention that the proteins that are derivatized in the hinge run all in the same position in the native gel, but the protein that is derivatized in the N- and C-terminals runs slightly more than the ones that are derivatized in the hinge. This is because in native gels, besides the charge, the macromolecule shape is quite important. In all the hinge mutants the DNA is in opposite ends, but in the mutants of the N- and C-terminals, the DNAs are in the same domain, making the positions in the gel slightly different.

Single molecule unfolding and refolding studies

The mutant at the terminals (S4C/R467C) shows a great amount of intermediates in the constant velocity experiments. At least 5 different intermediates with different extension are observed. The complete unfolding of the protein was never observed based on the theoretical length obtained with the WLC model. One possible explanation is that the enzyme is not completely folded even at low forces. With the optical tweezers it is possible to observe events at a few pNewtons. It is also important to mention that at a low force the instrumental noise is high. This is because of the presence of the DNA linkers. At a low force they are not as stiff as at high forces and some information is lost. Moreover, studies in bulk (Merino, 2008) have demonstrated that tlGK folding is not completely reversible in the presence of GdHCl, an observation that could explain in part the heterogeneity of the sample at the individual molecule level.

The events of unfolding in this thermophilic protein occur at rather low forces. At unfolding velocities of 100 nm/sec proteins generally unfold at 20-30 pN in the optical tweezers, and with the S4C/R467C mutant, unfolding occurs at around 10 pN at that velocity, but with a great dispersion of forces. This may be because the species that is mechanically unfolded is an intermediate species, or possibly a molten globule.

The T57C/T418C mutant, with the cysteines in opposite positions at the hinge of the protein, has an interesting behavior. Three intermediates of different extensions are observed and are unfolded in a wide distribution of forces. These intermediates correspond to 20%, 40% and 80% of the total extension of the protein. The intermediate

that corresponds to 40% of the complete structure of the protein seems to be same as the 80% one but broken into two species. It is very difficult to assign what region of the protein corresponds to each one of the intermediates. tlGK is a reentrant protein, and has a complex topology. It is important to mention that the addition of the intermediates corresponds to the complete unfolding of the protein, based on the WLC model. Another interesting observation of this mutant is that in the presence of the ADP the unfolding of the portion of the protein corresponding to 20% of the total extension is not observed. The maximum force at which unfolding can be observed in these experiments is about 67 pN, at higher forces the overstretching of the DNA takes place. Unfolding events of the 20% portion of the protein are not observed all the way up to 67 pN, implying that such part of the protein is being protected by the ADP preventing its unfolding below 67 pN. Another option is that this zone is unfolded at very low forces; this interpretation is, however, less probable, since, if anything, one would expect that the substrate would stabilized the protein. It is quite complex to determine the energy values and unfolding kinetics in this enzyme, to determine which part of the protein is being stabilized by the substrate, since the protein displays several intermediates. This is why other sites of binding were sought that would lead to the observation of two states, and preferably hopping between the two states at equilibrium. The protein with those characteristics is the S4C/T57C mutant, which reversibly unfolds in the absence of ADP at 27.1 pN, with an extension of 5.5 nm. At that force, the expected extension is 9 nm, so about 3.5 nm would be missing. There are several possible explanations for this. One is that maybe the part of the protein corresponding to the 3.5 nm is unfolded at forces

greater than 67 pN (overestretching of DNA). This option may be discarded, since the ΔG of unfolding is about 11.2 kcal/mol, which is quite high for a segment of 53 amino acid residues, or about 10% of the total protein. A linear extrapolation would predict that the ΔG for the whole protein would be around 112 kcal/mol, a value too high compared to those commonly reported for proteins (5-15 Kcal/mol; Creighton, 1990). This high ΔG may be due to the reentrant topology of the tlGK protein, since Shank et al. (2010) have demonstrated that when there is reentrant topology there is folding/unfolding coupling between different segments of the protein.

Another possible explanation for not having a complete extension of the protein is because at first, upon adding the DNA handles, one portion of the protein unfolds. The tlGK protein has an initial alpha helix of about 26 amino acid residues. This alpha helix is makes very few contacts with the rest of the structure, and if the calculation of unfolding extension at 28 pN was made discounting the alpha helix, the resulting value is near 5.5 nm, i.e. the precise extension observed. It is interesting to mention that, if this alpha helix is folded from the beginning, the protein is still active. When running native gels, the enzyme with the DNA handles attached is active. Besides, a preliminary experiment (with a mutant), where the alpha helix was removed from the protein and was expressed in BL21, reveals that the glucokinase found in the supernatant is active in an ADP-dependent manner. Finally, when observing the alignment of the sequences of proteins related with the tlGK, it is seen that this alpha helix is not conserved among the many proteins of the family. With this mutant, hopping occurs at greater forces in the presence of the ADP substrate, with a shift of about 2 pN in the force occurring in

50% of the events.

The thermodynamic analysis of the S4C/T57C mutant in the presence and absence of ADP shows an elevated ΔG of unfolding (11.2 and 12 kcal/mol, respectively). The ΔG value for the unfolding of tlGK *in multiplo* is not known, probably because of the difficulty of finding thermodynamic equilibrium conditions. No conditions have been found where unfolding reversibility of the protein is observed (Merino, 2008 and Rivas Pardo, in preparation). The value here determined *in singulo* could be close to the ΔG value for complete unfolding of the protein since most proteins show values of 5-15 kcal/mol (Creighton, 1990).

When studying the dependence of the rate constant on the force (within the force interval where transitions occur) in the presence and in the absence of ADP, one finds that the slopes of the curves are the same, which indicate that the distance to the transition state does not change because of the substrate. This same analysis was done adding glucose (the other substrate of the tlGK) and UTP (as specificity control, because tlGK does not bind UTP) and the same distance to the transition state was found. An interesting point is that when adding glucose, no shift in the force was observed in the presence of ADP, suggesting that ADP binds first and then glucose, in an orderly sequential way. Rivas Pardo (in preparation) has demonstrated that this is also observed in AFM and SAXS experiments, where structural changes are observed in the presence of ADP, but in the presence of glucose no changes are observed. The use of UTP as specificity control, where no changes are observed in the force necessary to obtain 50% of the transition, shows that the effect of ADP on tlGK is specific.

The addition of the distances to the transition state from the unfolded and the folded states gives the exact value of the distance of the mechanical unfolding of the protein at forces of about 28pN. It has been observed that at low forces or low spring constants, there are losses of transitions, and the addition of the distances to the transition state does not add to the total distance of the transitions, probably because the system is less stiff (Elms, 2010). On the other hand, the phenomenon of hopping of the S4C/T57C mutant occurs at high forces (~28 pN) with the DNA handles being more stretched and the system being stiffer; therefore, we observe less noise because of the higher forces and thus we don't need more sophisticated analysis of the data (such as hidden markovian methods) to obtain the agreement between the sum of distances to the transition state and the observed extension change upon unfolding.

Moreover, the rate of the unfolding extrapolated to zero force was an order of magnitude smaller in the presence than in absence of ADP, with no changes in the extrapolated refolding rates. In the presence of glucose or UTP there were no changes in the unfolding rates. These results suggest that the folded state of the protein is stabilized by the substrate with no effect on the distance to the transition state for the unfolding reaction.

CONCLUSIONS

The fact that DNA-protein-DNA was purified using gel eletroelution show a significantly increase of fishing in the tweezers and the yield from tweezers experiments was improved.

The fact that DNA-protein-DNA construct is active with the handles and its activity can be tested directly inside the gel, demonstrates that handle attachment does not interfere with tlGK activity.

The studies of mechanical unfolding of S4C/R467C mutant show high hysteresis.

Unfolding of the T57C/T418C mutant can be fit to the WLC model; changes are observed in the presence of substrate with respect to the number of intermediates.

The length of mechanical unfolding of S4C/T57C mutant does not conform to the WLC model, suggesting that the initial alpha helix is probably unfolded from the moment the DNA handles are attached, but that this partial unfolding does not interfere with the enzyme activity.

Equilibrium studies with the S4C/T57C mutant suggest that the substrate stabilizes the folded state of the enzyme, but it does not change the distance to the transition state.

The ΔG of unfolding of the S4C/T57C mutant (11.2 kcal/mol) is very high if we consider that a small portion of the enzyme is unfolded, indicating the existence of folding/unfolding coupling in the structure, possibly due to the re-entrant nature of the enzyme topology.

REFERENCES

Ashkin, A. (1970) Aceleration and trapping of particles by radiation pressure. Phys. Rev. Lett. 24, 156-159.

Ashkin, A., Dziedzic, J.M., Bjorkholm, J.E., Chu, S. (1986) Observation of a single-beam gradient force optical trap for dielectric particles. Opt. Lett. 11, 288-290.

Bell, G.I. (1978) Models of the specific adhesion of cells to cells. Science, 200, 618-627.

Bertz, M., and Rief, M. (2009) Ligand binding mechanics of maltose binding protein. J. Mol. Biol. 393, 1097-1105.

Bradford, M.M. (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein dye-binding. Anal. Biochem. 72, 248-252.

Bustamante, C. (2008) *In singulo* biochemistry: When less is more. Annu. Rev. Biochem. 77, 45-50.

Bustamante, C., Chemla, Y.R., Forde, N.R., and Izhaky, D. (2004) Mechanical processes in biochemistry. Annu. Rev. Biochem. 73, 705-748.

Bustamante, C., Cheng, W., and Mejia, Y.X. (2011) Revisiting the central dogma one molecule at a time. Cell 144, 480-497.

Bustamante C., Macosko J.C., and Wuite G.J. (2000) Grabbing the cat by the tail: manipulating molecules one by one. Nat. Rev. Mol. Cell Biol. 1, 130-136.

Bustamante, C., and Smith S.B. (2006) Light-force sensor and method for measuring axial optical-trap forces from changes in light momentum along an optical axis. U.S. Patent 7133132 B2.

Bustamante, C., and Smith S.B. (2007) Optical beam translation device and method utilizing a pivoting optical fiber. U.S. Patent 7274451 B2.

Cecconi, C., Shank, E.A., Bustamante, C., and Marqusee, S. (2005) Direct observation of the three-state folding of a single molecule. Science, 309, 2057-2060.

Cecconi, C., Shank, E.A., Marqusee, S., and Bustamante, C. (2007) Studying Protein Folding with Laser Tweezers. Proceedings of the International School Enrico Fermi, Società Italiana di Fisica, 145-160

Cecconi, C., Shank, E.A., Dahlquist, F.W., Marqusee, S., and Bustamante, C. (2008) Protein-DNA chimeras for single molecule mechanical folding studies with the optical tweezers. Eur. Biophys. J. 37, 729-733

Collin, D., Ritort, F, Jarzynski, C., Smith, S.B., Tinoco, I., and Bustamante, C. (2005) Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. Nature, 437, 231-234.

Creighton, T.E. (1990) Protein Folding. Biochem. J. 270, 1-16.

Crooks, G.E. (1999) Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. Phys. Rev. E., 60, 2721-2726.

Deniz A.A., Mukhopadhyay S., and Lemke E.A. (2008) Single-molecule biophysics: at the interface of biology, physics and chemistry. J R Soc Interface., 5, 15-45.

Dudko, O.K., Hummer, G. and Szabo, A. (2008) Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. Proc. Natl. Acad. Sci. U S A. 105, 15755-15760.

Elms, P. (2010) An investigation of the mechanical properties of the molten globule state of apomyoglobin. Doctoral Thesis, University of California, Berkeley.

Gouy, M., Guindon, S., and Gascuel, O. (2010) SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol., 27, 221-224.

Guerrero, R. (2008) Disponibilidad de enzimas para la síntesis de glicógeno por la vía indirecta en oocitos de rana chilena (*Caudiverbera caudiverbera*): Localización subcelular y caracterización parcial de lactato deshidrogenasa. Professional degree Thesis, University of Chile

Hann E., Kirkpatrick N., Kleanthous C., Smith D.A., Radford S.E., and Brockwell D.J. (2007) The effect of protein complexation on the mechanical stability of Im9. Biophys J., 92, L79-L81

Inoue, H., Nojima, H., and Okayama, H. (1990) High efficiency transformation of *Escherichia coli* with plasmids. Gene 96, 23-28.

Ito, S., Fushinobu, S., Jeong, J-J., Yoshioka, I., Koga, S., Shoum, H., and Wakagi, T. (2003) Crystal structure of an ADP-dependent glucokinase from *Pyrococcus furiosus*: Implications for sugar-induced conformational change in ADP-dependent kinase. J. Mol. Biol. 331, 871-883.

Ito, S., Fushinobu, S., Yoshioka, I., Koga, S., Matsuzawa, H., and Wakagi, T. (2001) Structural basis for the ADP-specificity of a novel glucokinase from a hyperthermophilic archaeon. Structure 9,205-214.

Jarzynski, C. (1997) Nonequilibrium equality for free energy differences. Phys. Rev. Lett. 78, 2690-2693.

Junker, J.P., Hell, K., Schlierf, M., Neupert, W., and Rief, M. (2005) Influence of substrate binding on the mechanical stability of mouse dihydrofolate reductase. Biophys. J. 89, L46-L48.

Junker J.P., Ziegler F., and Rief M. (2009) Ligand-dependent equilibrium fluctuations of single calmodulin molecules. Science, 323, 633-637.

Koga, S., Yoshioka, I., Sakuraba, H., Takahashi, M., Sakasegawa, S., Shimizu, S., and Ohshima, T. (2000) Biochemical characterization, cloning, and sequencing of ADP-dependent (AMP-forming) glucokinase from two hyperthermophilic archae, *Pyrococcus furiosus* and *Thermococcus litoralis*. J. Biochem. 128, 1079-1085.

Kramers, H.A. (1940) Brownian motion in a field of force and the diffusion model of chemical reactions. Physica, 7, 284-304.

Liphardt, J., Dumont, S., Smith, S.B., Tinoco, I., and Bustamante, C. (2002) Equilibrium information from Nonequilibrium Measurements in an Experimental Test of Jarzynski's Equality. Science, 296, 1832-1835.

Liphardt, J., Onoa, B., Smith, S.B., Tinoco, I., and Bustamante, C. (2001) Reversible unfolding of single RNA molecules by mechanical force. Science, 292, 733-737.

Merino, F. (2008) Adaptaciones de arqueas hipertermófilas a ambientes de altas temperaturas: estudios de estabilidad térmica y mecanismo cinético en quinasas dependientes de ADP. Professional degree Thesis, University of Chile.

Moffitt, J.R., Chemla, Y.R., Aathavan, K., Grimes, S., Jardine, P.J., Anderson, D.L., and Bustamante, C. (2009) Intersubunit coordination in a homomeric ring ATPase. Nature, 457, 446-450.

Neuman, K.C., and Nagy, A. (2008) Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. Nature Methods, 5, 491-505.

Rivas Pardo, J.A. (in preparation) Propiedades cinéticas y estructurales, determinantes moleculares de la especificidad de ligando y estabilidad de la glucoquinasa dependiente de ADP de *Thermococcus litoralis*. Doctorate Thesis, University of Chile

Schagger, H., and von Jagow, G. (1991) Blue native electrophoresis for isolation of membrane protein complexes in ezymatically active form. Anal. Biohem. 199, 223-231.

Shank, E.A. (2004) An investigation of the single molecule behavior of RNase H. Doctorate Thesis, University of California, Berkeley.

Shank, E.A., Cecconi, C., Dill, J.W., Marqusee, S., and Bustamante, C. (2010) The folding cooperativity of a protein is controlled by its chain topology. Nature, 465, 637-640.

Smith, S.B., Cui, Y., and Bustamante, C. (1996) Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. Science, 271, 795-799.

Smith, S.B., Cui, Y., and Bustamante, C. (2003) Optical-trap force transducer that operates by direct measurement of light momentum. Mehods Enzymol. 361, 134-162.

Smith S.B., and Rivetti C. (2010) TweezersLAB. http://tweezerslab.unipr.it.

Tinoco, I. (2004) Force as a useful variable in reactions: unfolding RNA. Annu. Rev. Biophys. Biomol. Struct. 33, 363-385.

Tinoco, I., and Bustamante, C. (2002) The effect of force on thermodynamics and kinetics of single molecule reactions. Biophys. Chem. 101-102, 513-533.

Tsuge, H., Sakuraba, H., Kobe, T., Kujime, A., Katunuma, N., and Ohshima, T. (2002) Crystal structure of the ADP-dependent glucokinase from *Pyrococcus horikoshii* at 2.0-Å resolution: A large conformational change in ADP-dependent glucokinase. Protein Sci. 11, 2456-2463.

Wenner, J.R., Williams, M.C., Rouzina, I., and Bloomfield, V.A. (2002) Salt dependence of the elasticity and overstretching transition of single DNA molecules. Biophys. J. 82, 3160-3169.

Willard, L., Ranjan, A., Zhang, H., Monzavi, H., Boyko, R.F., Sykes, B.D., and Wishart, D.S. (2003) VADAR: a web server for quantitative evaluation of protein structure quality Nucleic Acids Res., 31, 3316-3319.

Zhang, X., Halvorsen, K., Zhang, C.Z., Wong, W.P., and Springer, T.A. (2009) Mechanoenzymatic cleavage of the ultralarge vascular protein von Willebrand factor. Science, 324, 1330-1334.