DEPARTAMENTO DE GENÓMICA Y PROTEÓMICA

N-ACETIL-GLUTAMATO QUINASA DE ESCHERICHIA COLI: ESTRUCTURA TRIDIMENSIONAL E IMPLICACIONES FUNCIONALES

FERNANDO GIL ORTIZ

UNIVERSITAT DE VALENCIA
Servei de Publicacions
2003

Aquesta Tesi Doctoral va ser presentada a Valencia el día 4 de Juliol de 2003 davant un tribunal format per:

- Dra. Dª. Concepción Abad Mazario
- Dr. D. Ignacio Fita Rodríguez
- Dr. D. Carlos Gómez-Moreno Calera
- Dr. D. José M. González Ros
- Dr. D. Armando Albert De la Cruz

Va ser dirigida per:
Dr. D. Vicente Rubio Zamora

Depòsit legal:
I.S.B.N.:84-370-5771-10
Edita: Universitat de València
Servei de Publicacions
C/ Artes Gráficas, 13 bajo
46010 València
Spain
Telèfon: 963864115

UNIVERSITAT DE VALENCIA
Departamento de Bioquímica y Biología Molecular

INSTITUTO DE BIOMEDICINA DE VALENCIA
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

N-acetil-L-glutamato quinasa de *Escherichia coli*: Estructura tridimensional e implicaciones funcionales

Memoria presentada por
FERNANDO GIL ORTIZ
para optar al grado de
Doctor en Biología

Director
Prof. VICENTE RUBIO ZAMORA

Valencia, Abril de 2003

El autor ha disfrutado de una beca para la realización de tesis doctorales de la Fundación Ramón Areces.

El trabajo contenido en el capítulo 1 se ha financiado con la ayuda PM97-0134-C02-01 de la Dirección General de Enseñanza Superior del Ministerio de Educación y Ciencia, y el de los demás capítulos con la ayuda Rayos X de la Fundación Ramón Areces, y con la GV01-259 de la Generalitat Valenciana (capítulos 2, 3 y 5) o BMC2001-2182, del Ministerio de Ciencia y Tecnología (capítulo 4).

El autor agradece a la Unión Europea y a los sincrotrones de Hamburgo (DESY-EMBL) y Grenoble (ESRF-EMBL) por la asistencia financiera y apoyo para la recogida de datos.

AGRADECIMIENTOS

Todo comenzó el verano del 97, un joven y sonriente licenciado en biología comenzaba sus andares por el mundo de la ciencia. Tras seis años, con unas cuantas neuronas menos, he acabado la tesis aunque aun no me lo crea del todo. Durante todo este tiempo han pasado muchísimas cosas, unas buenas y otras malas, pero sin duda alguna lo mejor que me ha pasado es conocer a un montón de buenos amigos/as, a la mayoría de los cuales no podré olvidar nunca. En este corto espacio voy a intentar dar las gracias a toda esta gente maravillosa, espero no dejarme a nadie y si lo hago espero que me perdonéis.

En primer lugar, como no podría ser de otra forma, doy las gracias a mi director de tesis, el Prof. Vicente Rubio Zamora, por la oportunidad que me brindó al hacerme un hueco en su laboratorio, por aguantarme durante todos estos años y por su ayuda sin la cual esta tesis no se hubiera hecho realidad.

Un agradecimiento muy especial tengo para mis compañeros de laboratorio, con los que he estado tantos años por la incalculable ayuda que me han prestado día a día, así como por su amistad y cariño. A Belén, por ser tan buena persona y amiga, y estar siempre dispuesta a ayudar a los demás. Deseo de corazón que seas muy feliz con tu Javierín por que te lo mereces. A Leonor, por ser tan buena compañera y amiga, por su simpatía argentina y por esos momentos de terapia compartida. A Santi y Clara, con quien tantos años he pasado, y como no a nuestra buena amiga Santa Orosia. También a nuestro super técnico Jorge, alias Pocholo, y a los más jovencitos del labo, Sandra y Juanma.

No quiero olvidarme de las personas que estuvieron en el laboratorio y de los que guardo grandes recuerdos, Alberto Marina, por su inestimable ayuda, Chelo Climent, Matxalen Uriarte, Michelle y Mari Luz. Un recuerdo muy especial tengo para mi amiga Amparillo, con la que he pasado tan buenos momentos, por su amistad y aquellas charlas sin sentido en Viveros.

Con especial cariño quiero dar las gracias al Prof. Ignacio Fita por su amistad y su constante ayuda y apoyo durante todo el desarrollo de esta tesis. También a mis amigos del Instituto de Biología Molecular de Barcelona, que tan bien me ha tratado siempre, por su ayuda en mis incursiones sincrotroneras, sobretodo a Nuria, Cristina, Eva, las Raqueles, Alex y en especial a Wendy Garcia y Xavi Carpena.

A mi gran amiga Gracia Lopez, una persona inolvidable, por darme su amistad, por escucharme y hacerme sonreír cuando tenía problemas y por los innumerables buenos momentos que hemos pasado juntos.

A todos mis amigos de la peña de Viveros, unas personas increíbles, por su amistad y apoyo en todo momento, y por hacer de la hora de la comida uno de los mejores momentos del día: a Susana, por ser tan buena amiga y por aguantarme en mis desvaríos, a Anita, por su gran simpatía, a Rosa, por esas discusiones al límite, a Toni y Paz, por ser una pareja genial, a Salome, por su ingenuidad y sentido del humor, a JM, por su inmejorable sangría, a Ana Gorostidi y a Lidia. Y como no a los habituales de la hora de comer, Paloma, Hugo y Dani.

A mi gran amigo Nachin, por ser una persona estupenda y por esas charlas tan profundas, indispensables en cualquier gran acontecimiento.

A la segurata Yolanda, sin duda la más eficiente del lugar, por ser tan buena amiga y por encontrar siempre algún aparato para apagar antes de irme.

A toda la gente del Instituto de Biomedicina de Valencia, en especial Dr. Cubells, por su simpatía, a Manolo, por su inestimable ayuda informática, a la gente de administración, en especial a Marita y Mari Luz, y a las chicas del escuadrón de limpieza, Amparo, Susana, Pili y Luisa.

También quiero dar las gracias a todos mis amigos del Instituto de Investigaciones Citológicas, que tanto me ayudaron cuando estuve allí, en especial a Filo Bote e Inmaculada Micó, dos personas inolvidables, y a mi gran amigo Juanqui, por dejarse ganar al futbito.

No quiero olvidarme de los más grandes amigos que una persona pueda tener, Casque, Rafa, Pepe, Manolo, Fer, Mon y Luis, por estar siempre a mi lado y por ser casi como mis propios hermanos. Y como no, al resto de la peña de las nueve y media, Noelia, Eva y Alicia.

A Chayanne, por acompañarme en esas las largas noches en el IBV, y a Don Facundo, por impregnarnos de su alegría y sabor cubano.

A mi novia Noemí, una persona maravillosa, sin duda quien más me ha apoyado y alentado, por estar a mi lado durante todo este tiempo, por creer en mi y por mimarme tanto.

Por ultimo, doy las gracias a mis Padres y a mis hermanos, Richar, Alex, Yacky y Rocko, para los cuales cualquier cosa que diga es poco.

Esta tesis está dedicada a mis padres, a mis hermanos y a Noemí.

ÍNDICE

Introducción general **1**

Capítulo 1 ..... **15**

N-acetyl-L-glutamate kinase from *Escherichia coli*: cloning of the gene, purification and crystallization of the recombinant enzyme and preliminary X-ray analysis of the free and ligand-bound forms.

Capítulo 2 ..... **23**

Structure of acetylglutamate kinase, a key enzyme for arginine biosynthesis and a prototype for the amino acid kinase family, during catalysis.

Capítulo 3 ..... **55**

A crystallographic glimpse of a nucleotide triphosphate (AMPPNP) bound to a protein surface. External and internal AMPPNP molecules in crystalline N-acetyl-L-glutamate kinase.

Capítulo 4 ..... **67**

The course of phosphorus in the reaction of N-acetyl-L-glutamate kinase, determined from the structures of crystalline complexes, including a complex with an AlF_4^- transition state mimic

Capítulo 5 ..... **95**

Towards structural understanding of feed-back control of arginine biosynthesis: cloning and expression of the gene for the arginine-inhibited N-acetyl-L-glutamate kinase from *Pseudomonas aeruginosa*, purification and crystallization of the recombinant enzyme and preliminary X-ray studies.

Resumen de los resultados **105**

Discusión general **121**

Conclusiones **131**

Bibliografía **135**

correspondiente a la introducción general, resumen de los resultados y discusión general

Introducción General

INTRODUCCIÓN GENERAL

Esta tesis comprende cinco manuscritos, cuatro de ellos ya publicados y otro en fase de evaluación editorial, sobre el enzima N-acetil-L-glutamato quinasa (abreviado de aquí en adelante como acetilglutamato quinasa, o, más brevemente, como NAGK), del que se describe aquí su estructura tridimensional, sitios de unión de sustratos, mecanismo catalítico, y curso del fósforo en la reacción. Además se describe la unión superficial, fuera del centro activo, de una segunda molécula de nucleótido, y se avanza hacia el conocimiento de la regulación por retroalimentación ("feed-back") de la fosforilación de acetilglutamato. La aproximación metodológica es cristalográfica, basada en el uso de la difracción de rayos X. Los enzimas estudiados, de *Escherichia coli* y de *Pseudomonas aeruginosa*, son producidos en todos los casos en forma recombinante, a partir de un plásmido, en *E. coli*.

¿Por qué me he ocupado de la acetilglutamato quinasa? El grupo en el que he desarrollado mi trabajo de tesis había realizado numerosos estudios mecanísticos y de estructura física, pero no cristalográficos, sobre el enzima carbamil fosfato sintetasa (Rubio & Grisolia, 1977; Rubio et al., 1981; Rubio et al., 1991; Alonso et al., 1992; Alonso & Rubio, 1995; Rubio et al., 1998), una proteína compleja que cataliza una complicada reacción en tres pasos, dos de los cuales son fosforilaciones de bicarbonato y carbamato, y había concluido que era posible que un enzima mucho más sencillo, la carbamato quinasa, representara un buen modelo de los dominios de fosforilación de la carbamil fosfato sintetasa (Rubio et al., 1990; Rubio, 1994; Rubio & Cervera, 1995). Por ese motivo, a mitad de los años 90, el grupo inició un proyecto

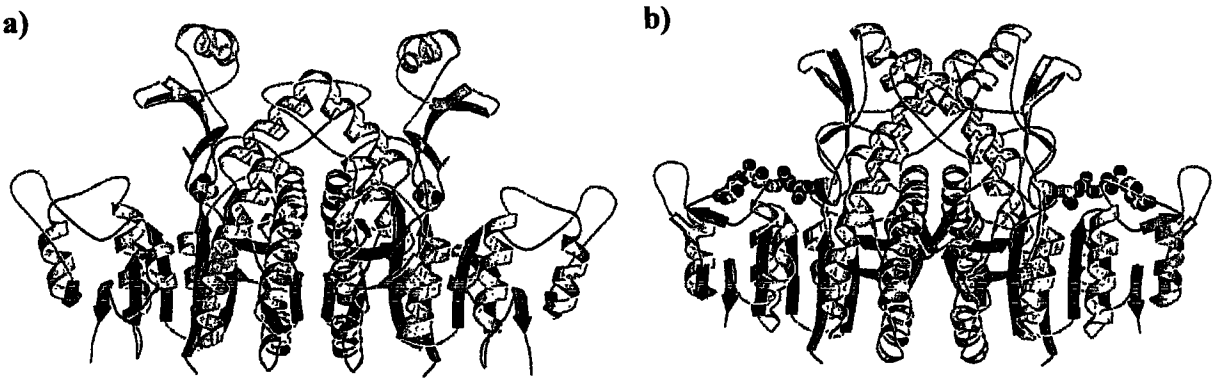


Figura 1. Representación de cinta de la estructura de las carbamato quinasas de *Enterococcus faecalis* con un ión sulfato unido (a) y de *Pyrococcus furiosus* con MgADP unido (b). Se muestra en rojo la hoja β molecular y en azul y verde los demás elementos de una y otra subunidad de cada dímero. Figuras modificadas de Marina et al., 1999 y Ramón-Maiques et al., 2000.

cristalográfico basado en intentar determinar la estructura tridimensional de la carbamato quinasa (Marina et al., 1994). Este estudio permitió determinar la estructura de dos carbamato quinasas (Fig. 1), una mesofílica clásica, de *Enterococcus faecalis* (Marina et al., 1999), y otra, identificada inicialmente por otro grupo como una carbamil fosfato sintetasa atípica (Durbecq et al., 1997), y que resultó ser realmente una carbamato quinasa hipertermofílica (Uriarte et al., 1999; Ramón-Maiques et al., 2000). Estos estudios, junto con los de otro grupo norteamericano (Waldrop et al., 1994; Thoden et al., 1997), demostraron a las claras que carbamato quinasas y carbamil fosfato sintetasa tienen estructuras diferentes. Los estudios de nuestro grupo revelaron con la carbamato quinasa un tipo nuevo de plegamiento α/β abierto, no representado en las bases de datos estructurales (Marina et al., 1999).

El paso siguiente del grupo ha sido intentar establecer si ese plegamiento se circunscribe a las carbamato quinasas o se encuentra en otras proteínas. La búsqueda de identidades de secuencia en las bases de datos estableció como posible que la acetilglutamato quinasa fuera un enzima homólogo a la carbamato quinasa (Marina et al., 1998), presentando con ella una identidad de secuencia de 18-22 %. Al fin y al cabo ambos enzimas forman parte del grupo 2.7.2. de la Enzyme Commission (fosfotransferasas que tienen como aceptor un grupo carboxilato) (Fig. 2), y catalizan una reacción similar, la fosforilación reversible, a partir de ATP, de un grupo carboxilato (acetilglutamato quinasa) o carbamato (carbamato quinasa), es decir, la formación de un anhídrido mixto fosfórico-carbónico o fosfórico-carbámico.

E.C.2.7.2. Fosfotransferasas con un grupo carboxilo como aceptor

2.7.2.1	Acetato quinasa
2.7.2.2	Carbamato quinasa
2.7.2.3	Fosfoglicerato quinasa
2.7.2.4	Aspartato quinasa
2.7.2.6	Formato quinasa
2.7.2.7	Butirato quinasa
2.7.2.8	Acetilglutamato quinasa
2.7.2.10	Fosfoglicerato quinasa (GTP)
2.7.2.11	Glutamato 5-quinasa
2.7.2.12	Acetato quinasa (difosfato)
2.7.2.14	Isobutirato quinasa

Figura 2. Enzimas del grupo 2.7.2. de la Enzyme Commission (http://www.chem.qmul.ac.uk/iubmb/enzyme/EC2/7/2/). El listado ha sido modificado para excluir las entradas transferidas a otros grupos y la glutamato 1-quinasa, de la que los autores de la descripción se han retractado por carta al Dr. Rubio.

Mi trabajo ha sido someter a prueba esta posibilidad, y ha resultado en la confirmación plena de la hipótesis de que carbamato quinasas y acetilglutamato quinasas comparten el mismo patrón estructural. De hecho, como se verá más adelante, demostramos aquí que lo comparten en

un grado muy elevado. Es más, en el ínterin las actuales iniciativas de formación de familias estructurales basadas en la existencia de similitudes de secuencia ha generado una familia de proteínas a la que se denomina familia "aminoácido quinasa" (PF00696 de la base de datos Pfam; http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00696)(Bateman et al., 2002) (Fig. 3) y que incluye los enzimas carbamato quinasa, acetilglutamato quinasa, aspartoquinasa, glutamato 5-quinasa y UMP quinasa bacteriana. Como se verá a lo largo de esta memoria de tesis, nuestros alineamientos hacen prever que con muy alta probabilidad todos estos enzimas van a compartir el modo de plegamiento establecido inicialmente para la carbamato quinasa, y sus sitios de unión de sustratos van a ser muy similares y van a ocupar lugares equivalentes en la estructura.

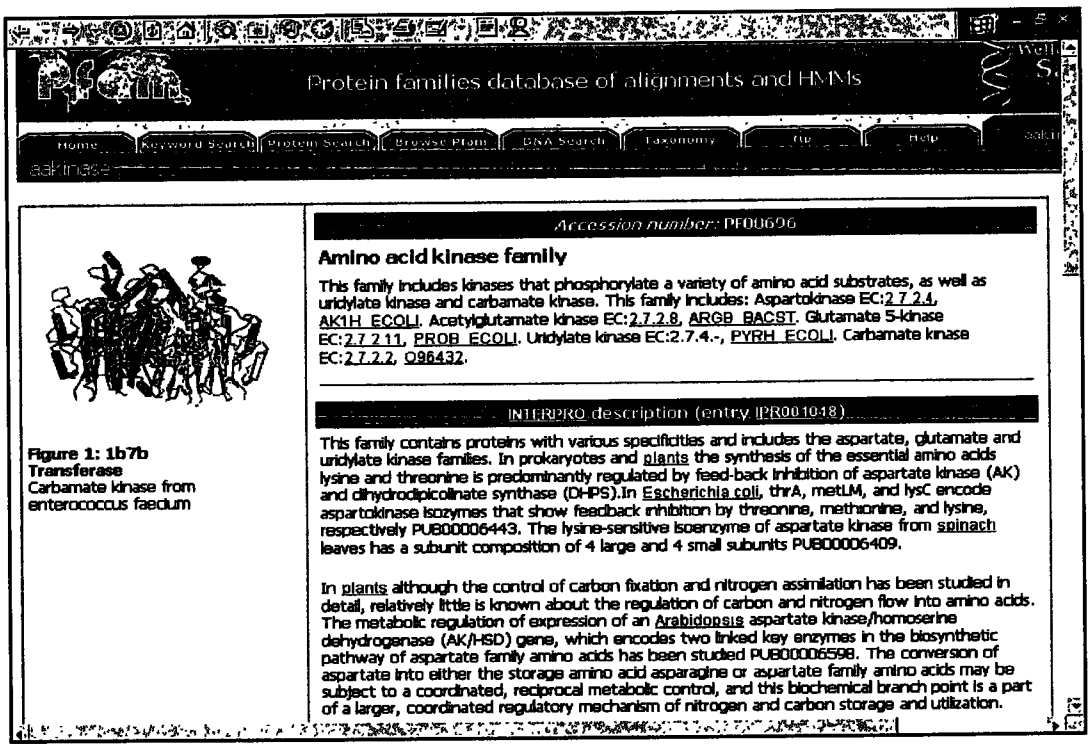


Figura 3. Página web de la base de datos Pfam donde se muestra la familia aminoácido quinasa (PF 00696).

Si estos enzimas, que comparten el mismo tipo de reacción catalizada, van a compartir también estructura, centros activos, modo de unión de sustratos y mecanismo catalítico, es importante caracterizar en detalle estos aspectos al menos en un enzima de la familia para tratar de entender la fisiología molecular de todos ellos. Esta es la labor que hemos realizado aquí con la acetilglutamato quinasa. Aunque la carbamato quinasa de *E. faecalis* permitió ya vislumbrar (Marina et al., 1999) y la de *P. furiosus* delinear (Ramón-Maiques et al., 2000) el sitio para ADP, sólo la estructura de la NAGK nos ha permitido definir con claridad, de forma completa, los lugares de unión de ambos sustratos, los residuos involucrados en la catálisis, y las distintas

fases del proceso catalítico y de la transferencia de fosforilo. Anteriormente, en la estructura de la carbamato quinasa de *E. faecalis* (Marina et al., 1999) se identificó un sulfato unido al enzima que podría representar uno de los fosfatos del ATP (Fig. 4), comprobándose que alrededor de

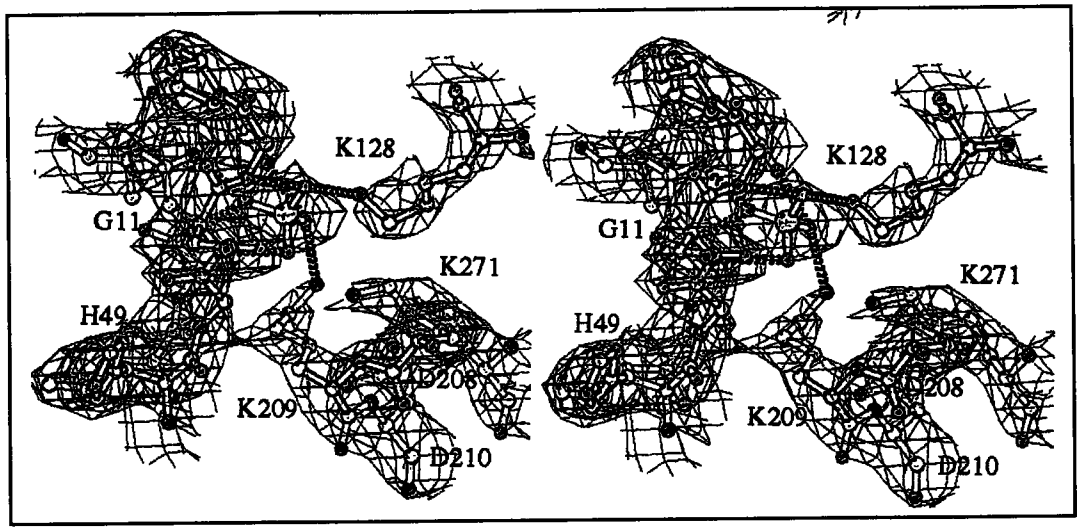


Figura 4. Vista estereoscópica del mapa de densidad electrónica $2F_o$-F_c y modelo molecular de bolas y bastones del anión sulfato presente en la carbamato quinasa de *Enterococcus faecalis* y de los residuos que lo rodean. El azufre es verde, oxígeno en rojo, nitrógeno en azul y carbono en amarillo. Tomado de Marina et al., 1999.

dicho sulfato se concentraba un gran número de residuos altamente conservados, la mutación de algunos de los cuales condujo a la pérdida de la actividad. Por tanto, se sugirió que el centro activo se correspondía con el bolsillo de unión del sulfato. Al determinar la estructura de la carbamato quinasa de *P. furiosus* se observó (Ramón-Maiques et al., 2000) la presencia de MgADP unido en una cavidad más amplia dentro de la que se incluye el sitio para el sulfato (Fig. 5). La superficie de la zona próxima al β-fosfato del ADP hacía suponer hacia dónde se uniría el carbamato, sustrato que es fosforilado por este enzima. Sin embargo, la inestabilidad tanto del carbamato (Caplow, 1968) como de su producto fosforilado, el carbamil fosfato (Allen & Jones, 1964), han hecho imposible por ahora, obtener complejos ternarios cristalinos de este enzima. De este modo, como se verá más adelante, la acetilglutamato quinasa ha venido a llenar este vacío en nuestro conocimiento constituido por la ubicación del sitio para el sustrato a fosforilar, permitiendo definir con toda precisión dónde y cómo se une el acetilglutamato y, por extensión, cómo se unen carbamato y carbamil fosfato en las carbamato quinasas. Es más, no resulta difícil imaginar, dadas las similitudes con el acetilglutamato, cómo se unirán glutamato y aspartato en otros dos enzimas de la familia, glutamato 5-quinasa y aspartoquinasa.

La transferencia enzimática de fosforilo es un proceso que ha sido objeto de considerable investigación previa (resumida en Knowles, 1980). Un gran número de enzimas transfiere un

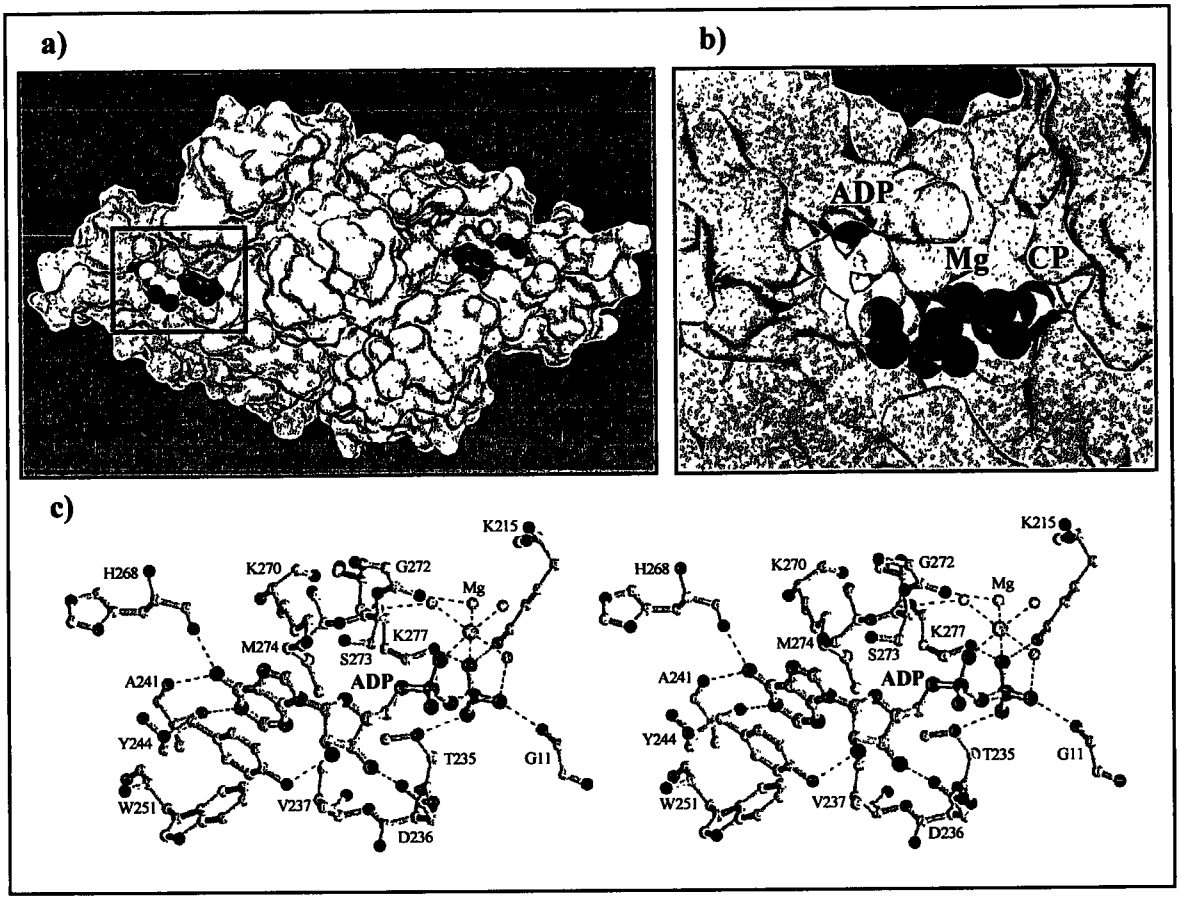


Figura 5. Representación de la superficie molecular del dímero de la carbamato quinasa de *Pyrococcus furiosus* (a). El rectángulo rojo indica el lugar donde se encuentra unida la molécula de ADP en una subunidad que se muestra en más detalle en (b) junto con una molécula de carbamilfosfato modelado. En (c) se muestra en modelo de bolas y bastones las interacciones que establece el MgADP con los residuos de la proteína. Tomadas de Ramón-Maiques et al., 2000.

grupo fosforilo de uno a otro sustrato, y en gran parte de los casos el donador del fosforilo es el ATP u otro nucleótido trifosfato. La transferencia del fosforilo puede suceder en dos pasos con formación de un fosfoenzima intermedio, con retención de configuración en el fosforilo transferido y mecanismo cinético generalmente de tipo ping-pong (Knowles, 1980 y 1982; Fry, 1992). Sin embargo, es más usual que la transferencia suceda en un solo paso, con inversión de configuración en el fosforilo (Fig. 6), y con evidencias cinéticas de formación de un complejo ternario del enzima con sus dos sustratos (mecanismo cinético secuencial) (Knowles, 1980 y 1982; Fry, 1992). En estas transferencias se plantea la duda de si el proceso es disociativo o asociativo (Wimmer & Rose, 1978; Knowles, 1980; Mildvan & Fry, 1987; Fry, 1992; Cleland & Hengge, 1995; Mildvan, 1997; Matte et al., 1998; Fersht, 1999). En una transferencia disociativa (Fig. 6, panel superior) se produciría a partir de ATP el anión metafosfato intercalado entre el ADP y el grupo a fosforilar, con el que eventualmente reaccionaría el metafosfato, generando el éster o anhídrido fosfórico en el producto. En la transferencia puramente asociativa (Fig. 6, panel

inferior) se produce un intermediario o estado de transición pentavalente, de modo que el enlace entre el átomo de O puente β-γ y el P_γ del ATP se alarga, a la vez que se va acortando la distancia con el grupo atacante (generalmente un átomo de O del sustrato a fosforilar), formándose gradualmente el nuevo enlace. Un nivel adicional de discusión lo constituyó la cuestión de si la transferencia sucede en línea o no (Knowles, 1982; Fersht, 1999), prediciéndose en el primer caso que el $O_{\beta\text{-}\gamma}$, el P_γ y el O atacante del sustrato a fosforilar deben ubicarse en una misma línea recta. La inversión de configuración indica claramente que la transferencia enzimática sucede en línea.

Aunque existen procedimientos químicos que permiten poner a prueba en mayor o menor medida los distintos mecanismos posibles de transferencia de fosforilo, una forma muy directa de estudio se puede derivar del análisis estructural, si se tiene la fortuna de poder observar el

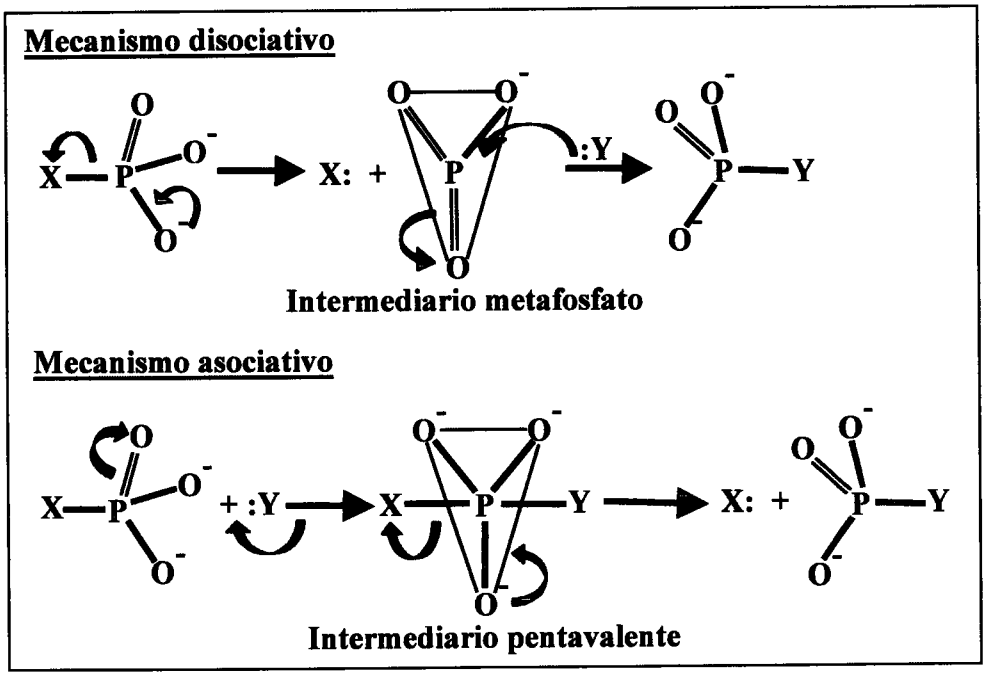


Figura 6. Representación esquemática de los mecanismos generales de transferencia del grupo fosforilo en un solo paso con inversión de configuración. Modificado de Mildvan, 1997.

complejo ternario con ambos sustratos o productos de la reacción. Ello sin embargo no es generalmente posible dada la reactividad intrínseca de dichos complejos, en los que tendría lugar la reacción. Por este motivo puede recurrirse a una de tres estrategias: 1) utilizar análogos que remeden el complejo ternario genuino, pero que sean inertes (Jin et al., 1999); 2) utilizar mutantes del enzima que hayan perdido su capacidad catalítica (Li et al., 1998); 3) interrumpir súbitamente la reacción mediante congelación rápida a bajísima temperatura (Kack et al, 1998). De las tres estrategias anteriores hemos adoptado la primera, ya que la tercera requiere

instalaciones complejas de que no disponíamos y se ha coronado con el éxito en muy pocas instancias (Chu et al., 2000). En cuanto a la estrategia 2), tiene el defecto de que requiere generar y cristalizar mutantes, así como la incertidumbre de no tener la seguridad absoluta de que la mutación introducida tenga un único efecto, la inactivación catalítica del enzima, sin afectar al modo de unión de los sustratos o al plegamiento de la proteína. Como se verá a lo largo de esta memoria, hemos adoptado el primer abordaje, la preparación de complejos cristalinos de la NAGK con análogos inertes de los sustratos.

El primer análogo utilizado es el adenilil imidodifosfato (AMPPNP) (Fig. 7a), un análogo del ATP considerado inerte en las reacciones de transferencia de fosforilo (Yount et al., 1971), en el que el O puente β-γ ha sido reemplazado por un átomo de N. La molécula resultante tiene un gran parecido con la de ATP, difiriendo, sin embargo, en algunos aspectos como el valor del pK del último grupo ionizable (lo que afecta a la afinidad aparente por el Mg o Mn, ver capítulo 3), el ángulo P_β-N-P_γ (127.2°, mientras que en el ATP el ángulo correspondiente P_β-O-P_γ es de 128.7°), la longitud de los enlaces P-N (1.68 Å, a comparar con 1.63 Å para el enlace P-$O_{\beta\gamma}$ del ATP) (Fig. 7b), y la presencia de un átomo de H unido al N del AMPPNP (Fig. 7a), pero no al O del ATP. Estas diferencias se consideran generalmente de poca importancia, pero, como se verá en el capítulo 3 de esta memoria de tesis, explican la presencia en la superficie de la NAGK de AMPPNP unido, libre de metal, en un lugar distinto al relevante desde el punto de vista catalítico. Otro aspecto destacable del uso del AMPPNP ha tenido que ver con su carácter de análogo inerte. Como se prueba más adelante (Capítulo 2) no lo es tanto, ya que en nuestro complejo ternario describimos formación parcial de enlace con el átomo de O atacante, aunque sin transferencia completa del fosforilo terminal.

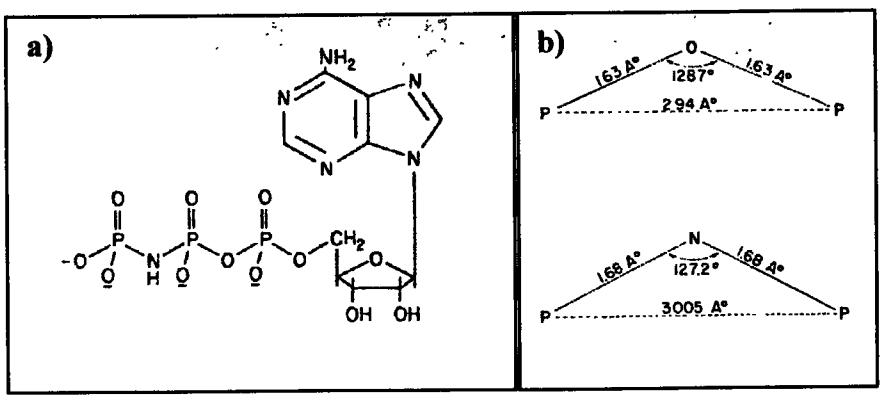


Figura 7. a) Representación esquemática de la molécula de AMPPNP. b) Comparación de las distancias y ángulos de enlace entre los átomos P-O-P del ATP y los átomos P-N-P del AMPPNP. Tomado de Yount et al., 1971.

Puesto que AMPPNP y ATP no son idénticos (Yount et al., 1971), es importante valorar el grado de fidelidad de los complejos que contienen este análogo respecto al complejo genuino. Nosotros hemos abordado esta cuestión utilizando otros análogos distintos, y estudiando también las estructuras cristalinas de sus complejos. Así, el fluoruro de aluminio planar, puede intercalarse entre el ADP y el sustrato a fosforilar, mimetizando así el estado de transición bipiramidal, reemplazando el metal al átomo de P_γ y los átomos de F a los oxígenos ecuatoriales de la bipirámide (Chabre, 1990; Wittinghofer, 1997; Xu et al., 1997; Schlichting & Reinstein, 1999) (Fig. 8). Hemos tenido la fortuna de poder cristalizar un complejo de la NAGK con MgADP, NAG y con una molécula de AlF_4^- interpuesta, abordando así desde otro ángulo la transferencia de fosforilo en este enzima. De hecho, hemos utilizado también otro complejo similar sin AlF_4^-, y un complejo adicional con ADP y sulfato, para conseguir una aproximación a la estructura del complejo del enzima con sus productos.

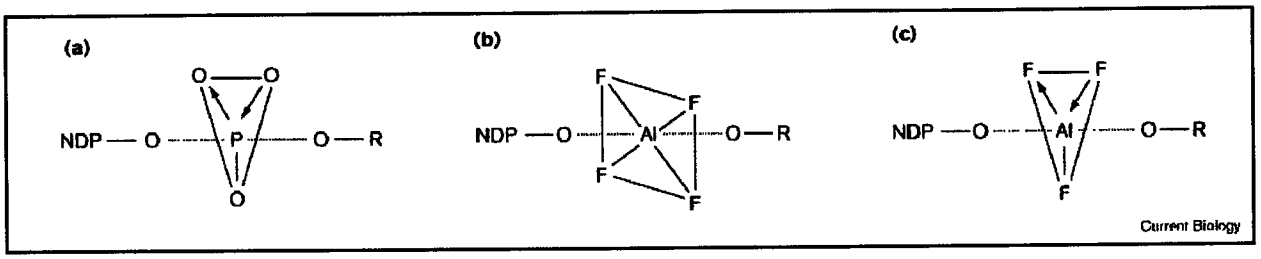


Figura 8. Representación del fosforilo en el estado de transición pentavalente (a) y de la disposición de los análogos planares tetrafluoruro de aluminio (b) y trifluoruro de aluminio. Tomado de Wittinghofer, 1997.

En conjunto, estos cuatro complejos (MgAMPPNP-NAG; MgADP-AlF_4^--NAG; MgADP-NAG; y ADP-SO_4^{2-}) nos han permitido trazar con alto grado de detalle el curso del fosforilo a lo largo de la reacción, y las modificaciones conformacionales en el enzima y de posición en los residuos catalíticos que acompañan la transferencia del fosforilo. En consecuencia, este enzima nos ha ofrecido la rara oportunidad de constatar que se trata de una transferencia en línea, en un sólo paso, de carácter marcadamente asociativo, catalizada por una elevada densidad de grupos que apantallan, neutralizan o acompañan cargas negativas, con una importante ventaja entrópica provista por el centro activo del enzima, que ordena, estructura y orienta productivamente en alto grado los grupos reaccionantes y catalíticos, y, a juzgar por la compresión de los grupos reactivos, con un importante efecto Circe (Jencks, 1975) en el que la energía de unión de los sustratos debe utilizarse en inducir un cambio de conformación que hace complementario el centro activo al estado de transición de la reacción (Pauling, 1946; Fersht, 1999). En ninguno de los enzimas del grupo EC 2.7.2 se ha caracterizado la transferencia de

fosforilo con este grado de profundidad. Así, aunque se ha publicado la estructura de la fosfoglicerato quinasa formando complejos con fosfoglicerato y AMPPNP (Auerbach et al., 1997) este enzima bilobular sufre notables cambios de conformación que aproximan los dos lóbulos entre sí para producir la conformación catalíticamente competente, pero ni en el mejor complejo publicado (Auerbach et al., 1997) tal aproximación ha sido completa, ya que el complejo contiene una molécula de agua interpuesta entre ambos grupos reactivos. Por tanto, nuestros resultados con la NAGK la convierten en un paradigma estructural sobre el mecanismo de transferencia de fosforilo en reacciones que forman acilfosfatos.

Una observación derivada de estos estudios ha tenido que ver con la molécula de AMPPNP. Como se describe en el capítulo 3, durante el análisis del complejo del enzima con MgAMPPNP-NAG encontramos una masa de densidad electrónica no explicada en la periferia del enzima. Aunque esta densidad se encontraba en el eje diádico de la molécula del dímero enzimático, coincidente con un eje cristalográfico, y por tanto en una zona donde la interpretación de las densidades puede ser arriesgada, el hecho de no encontrar tal densidad en los otros tres complejos estudiados a pesar de que exhiben idénticas formas cristalinas y disposición de las moléculas de enzima en el cristal, hace impensable la posibilidad de que esta densidad no tuviera una base física. Así, el trazado de un modelo de AMPPNP dentro de esta densidad ha permitido visualizar una molécula de este nucleótido, exenta de metal, extendida y con interacciones con la proteína muy limitadas, restringidas a la región del γ-fosfato. Puede decirse que se trata de la visión cristalográfica de una molécula del nucleótido trifosfato casi libre en solución, lo que, como se discute más adelante, es inusual y permite ofrecer una instantánea cristalográfica única de una molécula de nucleótido colisionando con el enzima y formando débiles interacciones con él, pero suficientes para garantizar el orden cristalográfico.

La acetilglutamato quinasa desempeña un papel clave en la biosíntesis de arginina (revisado en Cunin et al., 1986). Como muestra la Fig. 9, existen dos versiones alternativas de la ruta biosintética de arginina en microorganismos y plantas (no en mamíferos, donde la síntesis es menos activa y la ruta utilizada es distinta, Alonso & Rubio, 1989. En la versión menos usual, que exhibe *E. coli*, la ruta es completamente lineal y el paso limitante es la acetilación del glutamato a partir de acetilCoA, paso catalizado por la acetilglutamato sintetasa (Cunin et al., 1986). En esta ruta la acetilglutamato sintetasa (enzima que por cierto forma también parte de la familia aminoácido quinasa y aparentemente está relacionada con la acetilglutamato quinasa mediante un proceso de duplicación y deriva génica; Caldovic & Tuchman, 2003) es el punto de control mediante retroinhibición por la arginina. Sin embargo, en la versión más usual de la ruta, representada por ejemplo en *Pseudomonas aeruginosa* o en levadura, la acetilglutamato sintetasa

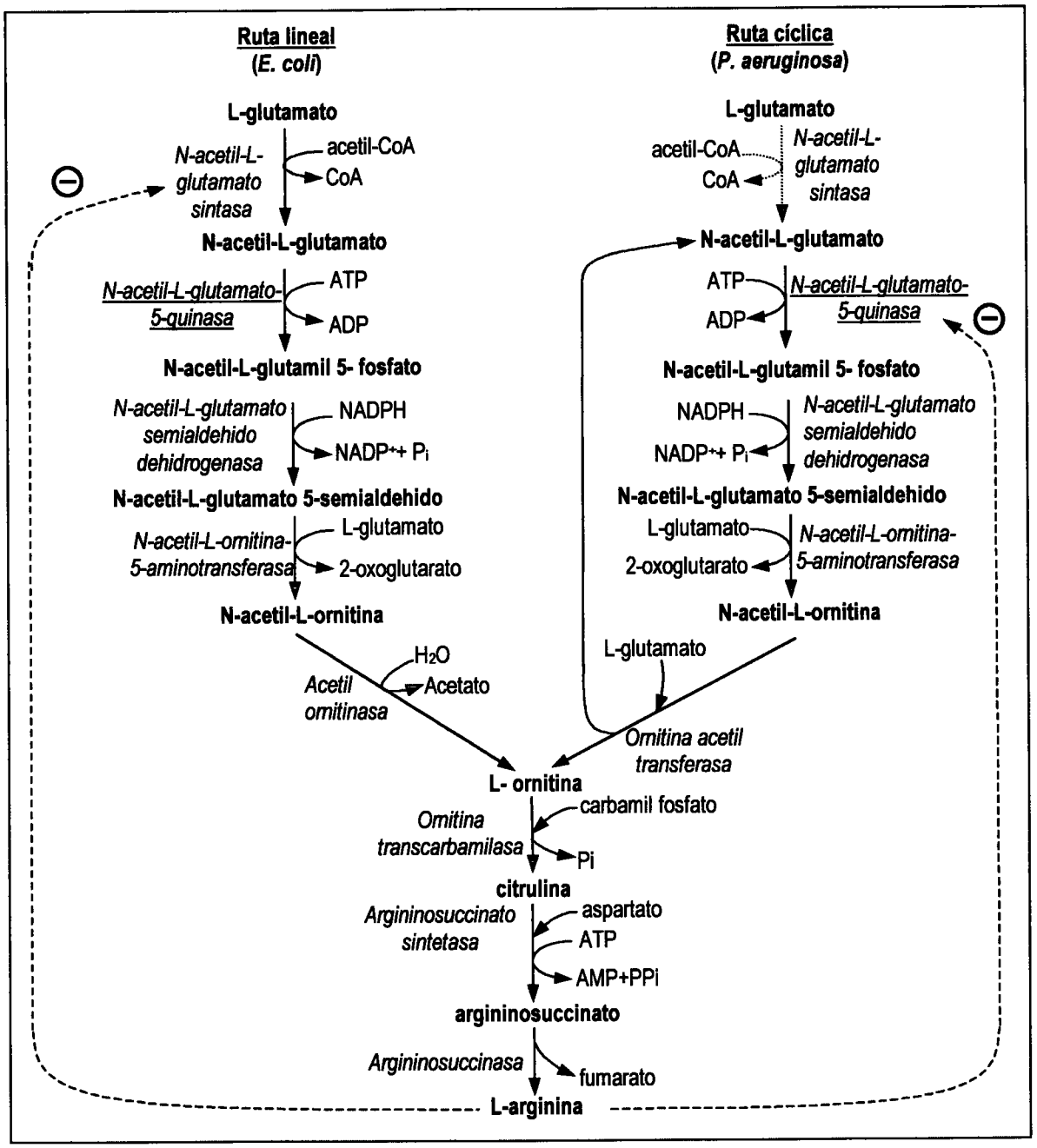


Figura 9. Diagrama representativo de las dos versiones de la ruta de biosíntesis de arginina que aparecen en microorganismos y plantas. Se ha subrayado el enzima objeto de estudio y se ha indicando el enzima inhibido por arginina en cada ruta mediante flechas de trazo discontinuo. La flecha punteada en la reacción de la N-acetilglutamato sintasa de la ruta cíclica denota el carácter anaplerótico de dicha reacción..

ceba el sistema inicialmente y desempeña un papel puramente anaplerótico, puesto que el grupo acetilo se recicla mediante transacetilación a partir de la acetilornitina al glutamato (Cunin et al., 1986), por una transacetilasa abundante y reversible. En consecuencia, en la práctica, la acetilglutamato quinasa es en realidad el paso limitante de esta versión de la vía, siendo este enzima el punto de inhibición por arginina (Haas & Leisinger, 1975a; Haas & Leisinger, 1975b;

Abadjieva et al., 2001). Por tanto, existen dos formas de acetilglutamato quinasa, la seguramente más simple, no inhibible por arginina, representada por la forma de *E. coli* (Vogel & McLellan, 1970; Dénes, 1973) estudiada en la presente memoria, y la posiblemente más compleja, inhibida alostéricamente por arginina, representada por el enzima de *P. aeruginosa* (Haas & Leisinger, 1975a; Haas & Leisinger, 1975b). En esta memoria también hemos dado los primeros pasos para averiguar por qué la arginina inhibe una y no la otra forma de NAGK, y cómo ejerce la inhibición. La estructura del enzima de *E. coli*, que nos ha revelado tanto sobre la forma en que el enzima une los sustratos y cataliza la reacción, no nos ha revelado nada sobre el posible mecanismo de inhibición. Por tanto, el laboratorio ha decidido proceder al estudio estructural de una NAGK inhibible, y en el capítulo 5 de esta memoria se presenta la clonación, caracterización y cristalización del enzima de *P. aeruginosa*. Posiblemente, si somos capaces de determinar las bases estructurales de la inhibición, obtengamos información de valor para otros enzimas de la misma familia aminoácido quinasa, ya que, entre ellos, además de la NAGK, la glutamato 5-quinasa y la aspartoquinasa catalizan reacciones similares y están sometidas a inhibición "feed-back" por el producto final de la ruta en que participan (Fig. 10).

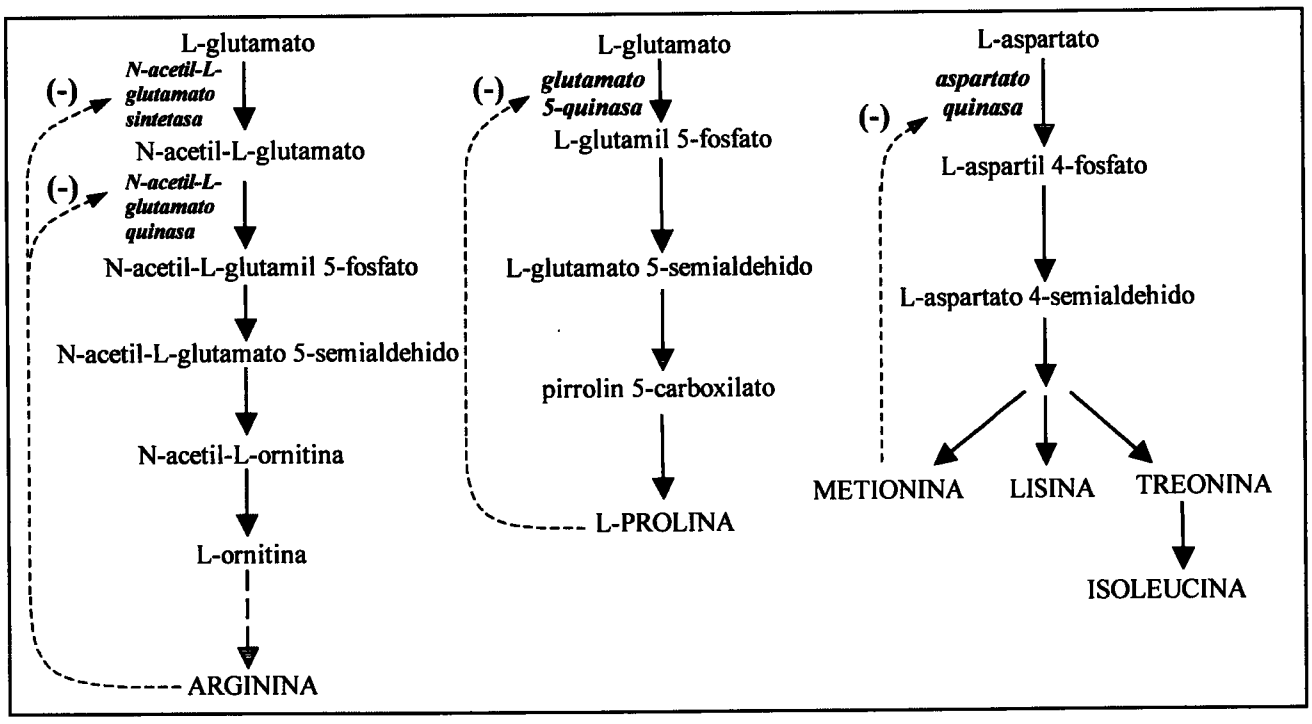


Figura 10. Rutas de biosíntesis de tres de los enzimas de la familia aminoácido quinasa mostrando la inhibición "feed-back" por el producto final de los enzimas homólogos, miembros de la familia aminoácido quinasa, N-acetil-L-glutamato sintetasa, N-acetil-L-glutamato quinasa, glutamato 5-quinasa y asparto quinasa. No se indican los nombres de los demás enzimas de la ruta.

Los manuscritos de que se compone esta memoria de tesis son los siguientes:

Capítulo 1)

Título: N-Acetyl-L-glutamate kinase from *Escherichia coli*: cloning of the gene, purification and crystallization of the recombinant enzyme and preliminary X-ray analysis of the free and ligand-bound forms.

Autores: Gil, F., Ramón-Maiques, S., Marina, A., Fita, I. y Rubio, V.

Referencia: *Acta Cryst.* (1999) **D55**, 1350-1352

Capítulo 2)

Título: Structure of acetylglutamate kinase, a key enzyme for arginine biosynthesis and a prototype for the amino acid kinase family, during catalysis.

Autores: Ramón-Maiques, S., Marina, A.*, Gil-Ortiz, F.*, Fita, I. y Rubio, V.

Referencia: *Structure* (2002) **10**, 329-342. * Equal contribution.

Capítulo 3)

Título: A crystallographic glimpse of a nucleotide triphosphate (AMPPNP) bound to a protein surface. External and internal AMPPNP molecules in crystalline N-acetyl-L-glutamate kinase.

Autores: Gil-Ortiz, F., Fita, I., Ramón-Maiques, S., Marina, A. y Rubio, V.

Referencia: *Acta Cryst.* (2002) **D58**, 1892-1895.

Capítulo 4)

Título: The course of phosphorus in the reaction of N-acetyl-L-glutamate kinase, determined from the structures of crystalline complexes, including a complex with an AlF_4^- transition state mimic.

Autores: Gil-Ortiz, F.*, Ramón-Maiques, S.*, Fita, I. y Rubio, V.

Referencia: Aceptado en *J. Mol. Biol.* * Equal contribution.

Capítulo 5)

Título: Towards structural understanding of feed-back control of arginine biosynthesis: cloning and expression of the gene for the arginine-inhibited N-acetyl-L-glutamate kinase from *Pseudomonas aeruginosa*, purification and crystallization of the recombinant enzyme and preliminary X-ray studies.

Autores: Fernández-Murga, M.L., Ramón-Maiques, A., Gil-Ortiz, F., Fita, I. y Rubio, V.

Referencia: *Acta Cryst.* (2002) **D58**, 1045-1047.

Capítulo 1

N-acetyl-L-glutamate kinase from *Escherichia coli*: cloning of the gene, purification and crystallization of the recombinant enzyme and preliminary X-ray analysis of the free and ligand-bound forms

Trabajo publicado en

***Acta Crystallographica Section D* (1999) D55, Pags. 1350-1352**

N-Acetyl-L-glutamate kinase from *Escherichia coli*: cloning of the gene, purification and crystallization of the recombinant enzyme and preliminary X-ray analysis of the free and ligand-bound forms

Fernando Gil[1], Santiago Ramón-Maiques[1], Alberto Marina[1,2], Ignacio Fita[2] and Vicente Rubio [1]

[1]Instituto de Biomedicina de Valencia (CSIC), C/ Jaime Roig 11, 46010-Valencia, Spain.

[2]IBMB-CSIC, Barcelona, Spain.

ABSTRACT

The gene for *Escherichia coli* N-acetyl-L-glutamate kinase (NAGK) was cloned in a plasmid and expressed in *E. coli*, allowing enzyme purification in three steps. NAGK exhibits high specific activity (1.1 μmol s^{-1} mg^{-1}), lacks Met1, and forms dimers (shown by cross-linking). Crystals of unligated NAGK diffract to 2 Å and belong to space group $P6_122$ or its enantiomorph $P6_522$ (unit-cell parameters a= b=78.6, c= 278.0 Å) with two monomers in the asymmetric unit. Crystals of NAGK with acetylglutamate and the ATP analog AMPPNP diffract to 1.8 Å and belong to space group $C222_1$ (unit-cell parameters a=60.0, b= 71.9 and c= 107.4 Å), with one monomer in the asymmetric unit. NAGK crystallization will allow the determination of proposed structural similarities to carbamate kinase.

INTRODUCTION

N-acetyl-L-glutamate kinase (NAGK; E.C.2.7.2.8) phosphorylates the γ-COOH group of N-acetyl-L-glutamate (NAG) in the second step of microbial arginine biosynthesis (Cunin *et al.*, 1986). In many microorganisms NAGK is a crucial control point, being feed-back inhibited by the final product arginine (Cunin et al., 1986). Little is known about the structure and catalytic mechanism of NAGK. The *P. aeruginosa* enzyme is a homodimer of a 29-kDa subunit, and appears to also form higher oligomers (Haas & Leisinger, 1975). The gene-deduced amino acid sequences indicate that bacterial and chloroplast NAGK polypeptides (Marina *et al.*, 1998) are composed of 258-304 amino acid residues and exhibit considerable sequence identity. Substantial sequence similarity (approximately 20 % identities and 40 % conservative replacements) was also found with carbamate kinase (Marina *et al.*, 1998), the final enzyme of the arginine deiminase

pathway of arginine catabolism, an enzyme whose three-dimensional structure we have recently determined (Marina *et al.*, 1999). Carbamate kinase is a homodimer of a polypeptide of approximately 30 kDa and, similarly to NAGK, catalyzes the transfer of the γ-phosphoryl group of ATP to a COO^- group. These similarities suggest that carbamate kinase and NAGK are structurally and functionally similar (Marina *et al.*, 1999). Since carbamate kinase presents a new α/β fold, NAGK and carbamate kinase might be members of a new structural family which may also include, given the sequence similarities (data not shown), γ-glutamyl kinase and long chain fatty acyl CoA synthetase. The present report is a necessary preliminary step in testing this possibility.

EXPERIMENTAL

The NAGK gene was PCR-amplified from *E. coli* DNA, using Deep Vent DNA polymerase (from New England Biolabs), and the primers 5'CTTATTTACTAGTGTCATGATGAATCCATTAATTATCAAACTGGGC3' and 5'GCTGCGCCGCTCAGCAACAAAACTTAAGCTAAAATCCGC3', which introduce *BspH*I and *Blp*I sites at the initiator ATG and downstream of the stop codon, respectively. The *BspH*I and *Blp*I digested amplified fragment was ligated using T4 ligase into the *Nco*I and *Blp*I sites of plasmid pET-15b (Novagen), and *E. coli* DH5α cells (from Clontech) were transformed. Plasmid pNAGK24 was isolated and was shown by restricition analysis and automated DNA sequencing to harbor the full NAGK gene. *E. coli* BL21(DE3) cells (from Novagen) transformed with pNAGK24 and grown to A_{600} = 0.6 at 310 K in 1.5 l LB broth with 75 μg ml^{-1} ampicillin were induced for 3 h with 1 mM isopropyl-β-D-thiogalactoside and were harvested by centrifugation. Subsequent steps were carried out at 277 K. The cells, suspended as 10 ml g^{-1} cells in 0.1 M sodium phosphate pH 7.0, 0.2 mM dithioerythritol, were disrupted by sonication. After centrifugation (30 min, 35 000 g), the protein in the supernatant which precipitated between 40 and 60 % saturation (at 273 K) of ammonium sulphate was dissolved in 20 ml of buffer A (10 mM sodium phosphate pH 7.0, 0.2 mM dithioerythritol) and dialyzed against the same buffer before application to a Q-Sepharose Fast Flow (Pharmacia Biotech) column (1x18 cm) equilibrated with buffer A. After washing the column with 75 ml of buffer A, a 350-ml linear gradient of 0-0.5 M NaCl in the same buffer was applied. NAGK was eluted as a single peak at approximately 0.12 M NaCl. Fractions rich in enzyme were mixed, dialyzed against 20 mM Tris-HCl pH 7.5, 20 mM $MgCl_2$, 0.2 mM dithioerythritol, and applied to an Affigel Blue (Bio-Rad)

column (2x20 cm), which was equilibrated and washed with 3 column volumes of the dialysis buffer. Pure NAGK was eluted with two column volumes of this buffer supplemented with 2.5 mM ATP and 20 mM NAG, and was precipitated with ammonium sulphate (70 % saturation) and stored as a slurry at 277 K. For crystallization, the enzyme was placed in the appropriate solution by repeated centrifugal ultrafiltration (Microsep 10K, Pall Filtron).

Initial crystallization conditions using the hanging-drop vapor-diffusion method were tested with the sparse-matrix sampling procedure (Jancarik & Kim, 1991), mixing 1.5 µl of reservoir fluid and 1.5 µl of 10 mg ml^{-1} NAGK in 10 mM sodium/potassium phosphate pH 7.0, 1 mM dithioerythritol, either alone or supplemented with 24 mM NAG, 30 mM $MgCl_2$ and 6 mM of the inert ATP analog AMPPNP ("substrates"). The best crystals, reaching about 0.6 mm in the largest dimension were produced in about a week at 295 K. In the absence of substrates, crystals were obtained with 0.1 M sodium citrate pH 5.6, 26-32 % PEG 4K (Sigma) and 0.1-0.3 M ammonium acetate; in the presence of substrates, crystals were obtained with 0.1 M sodium acetate pH 4.6, 27-32 % PEG monomethylether 2K (Hampton) and 0.1-0.3 M ammonium

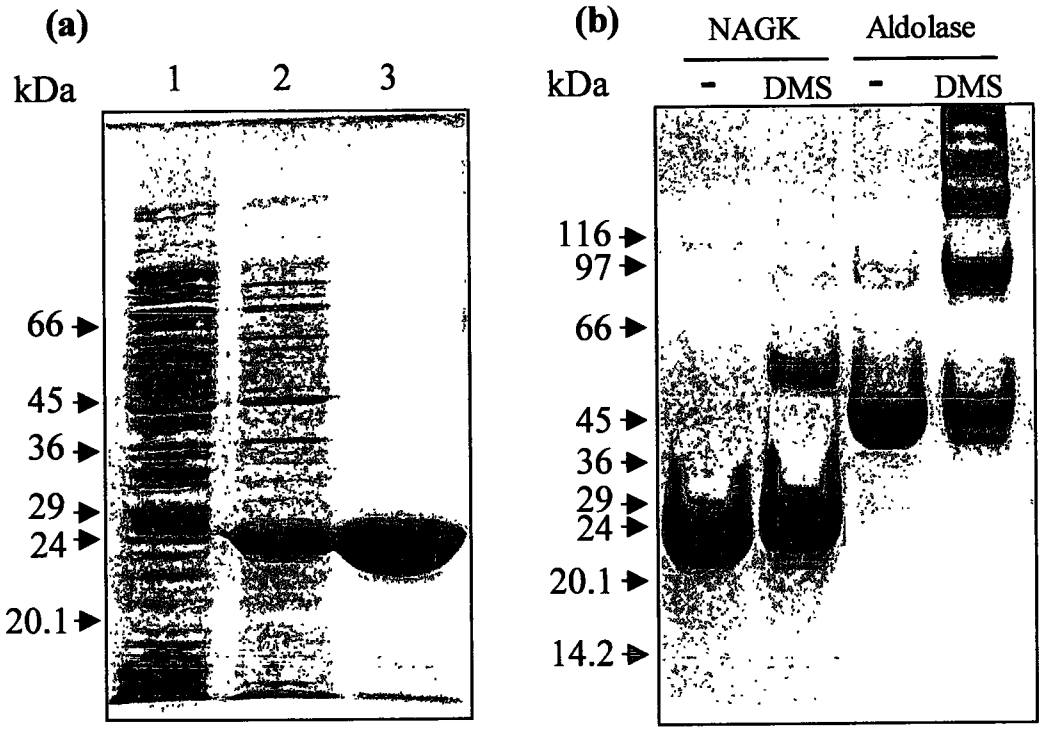
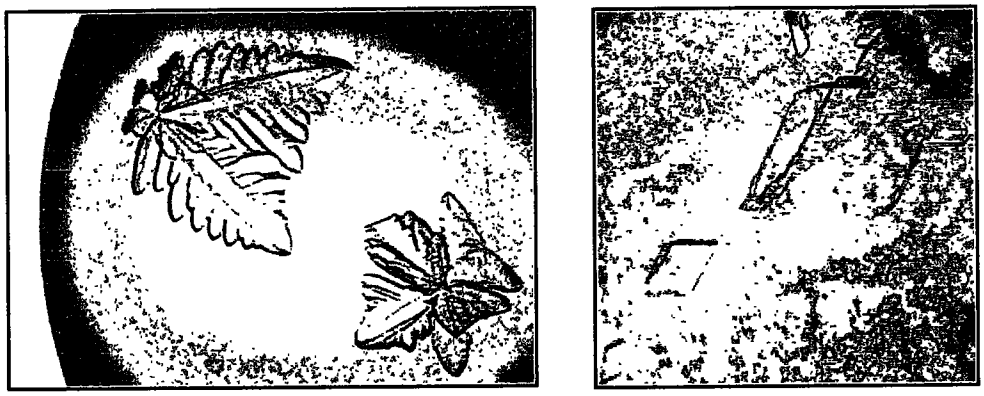


Figure 1. Expression, purification and cross-linking of NAGK. (a) SDS-PAGE of cell extracts of *E. coli* BL21(DE) cells transformed with plasmids pET-15°b (lane 1) or pNAGK24 (lane 2). Lane 3, purified enzyme. (b) Cross-linking with dimethylsuberimidate (DMS) using the method of Davies & Stark (1970) of 4 mg ml^{-1} NAGK or aldolase (from rabbit muscle; monomer mass, 40 kDa). The minus sign indicates omission of DMS. Arrows mark the migration position of protein standards of the indicated masses.

sulphate. Harvesting solutions were 36 % PEG 4K, 0.3 M ammonium acetate and 10 % ethylene glycol in 0.1 M sodium citrate pH 5.6 for crystals grown without substrates and 36 % PEG monomethylether 2K, 0.24 M ammonium sulphate and 5 % ethylene glycol in sodium acetate pH 4.6 for crystals grown with substrates. Crystals of about 0.3 mm in the longest dimensions were examined with a MAR Research image-plate area detector mounted on a Rigaku rotating copper-anode X-ray source operating at 40 kV and 100 mA. Data were collected at 100 K from crystals flash-cooled using the Oxford cryosystem. The data set for crystals grown with substrates was processed and scaled with



(a) (b)

Figure 2. Typical crystals of *E.coli* NAGK grown at 295 K (a) in the absence of ligands or (b) in the presence of 12 mM NAG, 15 mM $MgCl_2$ and 3 mM AMPPNP. In both cases, the largest dimension was approximately 0.6 mm.

DENZO and SCALEPACK (Otwinowski, 1993), whereas for crystals grown without substrates MOSFLM (Leslie, 1990) and SCALA (Evans, 1997) were used.

NAGK activity was assayed using the method of Haas & Leisinger (1975). One enzyme unit (16.7×10^{-9} mol s^{-1}) yields 1 μmol min^{-1} hydroxamate at 310 K. Protein was assayed using the method of Bradford (1976) with bovine serum albumin as standard. SDS-PAGE was carried out using the method of Laemmli (1970). Cross-linking with dimethyl suberimidate (DMS; Pierce) and SDS-PAGE of the covalent adducts was performed using the method of Davies & Stark (1970).

RESULTS AND DISCUSSION

The product generated by PCR-amplification of *E. coli* DNA was identical in size and DNA sequence the expectations for the NAGK gene (Parsot et al., 1988). Extracts of *E. coli*

BL21(DE3) transformed with an expression plasmid (pNAGK24) which contains as an insert the amplified DNA (see §2), but not extracts of cells transformed with the parental vector without the insert, exhibited NAGK activity (data not shown) and showed a prominent band on SDS-PAGE (Fig. 1a) migrating with the expected mass for NAGK (mass estimate from semilogarithmic plots of protein standards, 27 kDa; mass expected from the sequence, 27.15 kDa). The enzyme, purified in three steps from pNAGK24-transformed cells (Fig. 1a, lane 3), exhibited similar activity (1.06 μmol s^{-1} mg^{-1} protein) to the purest bacterial preparation reported thus far (Haas & Leisinger, 1975). Automated Edman degradation yielded the expected N-terminal sequence (MNPLIIK), except for the lack of Met1, which is removed postranslationally. Upon cross-linking with dimethylsuberimidate (DMS) a new band with the mass of a dimer was observed by SDS-PAGE (Fig. 1b), indicating that, as with *P. aeruginosa* NAGK (Haas & Leisinger, 1975), *E. coli* NAGK forms dimers. With the tetrameric enzyme aldolase, used as a control, the four expected bands were observed (Fig. 1b; Davies & Stark, 1970).

NAGK crystals obtained in the absence of substrates had fronded appearance (Fig. 2a), diffracting as monocrystals to at least 2 Å resolution with a conventional X-ray source. For practical reasons, a complete data set was only collected at 2.95 Å resolution (93.4 % completeness; R_{merge}= 12.1). The space group was hexagonal $P6_122$ or its enantiomorph $P6_522$, with unit-cell parameters a= b=78.6, and c= 278.0 Å. Packing-density considerations (Matthews, 1968) indicate that, for a monomer mass of 27 kDa, the unit cell could contain 24 monomers (V_m= 2.6 $Å^3$ Da^{-1}; solvent content 53 %), corresponding to two monomers per asymmetric unit.

In the presence of substrates, crystals with a rhomboidal shape (Fig. 2b) grew from an amorphous precipitate and diffracted to at least 1.8 Å, although a complete data set was collected at 2.3 Å (90.4 % completeness, R_{merge}= 11.5). The space group was $C222_1$, with unit-cell parameters a=60.0, b= 71.9 and c= 107.4 Å, and an estimated eight monomers per unit cell (V_m= 2.1 $Å^3$ Da^{-1}; solvent content, 42 %), or one monomer per asymmetric unit. If the molecule is organized as a dimer, then the molecular two-fold symmetry axis should be coincident with a crystallographic dyad axis. In fact, crystal soaking in solutions of K_2PtCl_6 yielded a Pt derivative with a single location of the metal on the crystallographic dyad X axis, suggesting that this axis could coincide with the molecular two-fold axis.

Molecular replacement with the program AMoRe (Navaza, 1994) using a polyalanine model of residues 2-105 and 160-310 of carbamate kinase (residues 106-159 form a flexible subdomain and were omitted; Marina et al., 1999) has thus far yielded suggestive but not

conclusive solutions. Further heavy-atom derivatives and the selenomethionine-substituted protein are presently being prepared.

This work was supported by grant DGES PM-97-0134-C02-01. We thank J.J. Calvete and A. Martínez for N-terminal and DNA sequencing, respectively. FG is a fellow of Fundación Ramón Areces, and SR-M and AM are pre-doctoral and post-doctoral fellows of the Generalitat Valenciana.

REFERENCES

Bradford, M.M. (1976). *Anal. Biochem.* **72**, 248-254.

Cunin, R., Glansdorff, N., Piérard, A. & Stalon, V. (1986) *Microbiol. Rev.* **50**, 314-352.

Davies, G.E. & Stark, G.R. (1970*). Proc. Natl. Acad. Sci. USA* **66**, 651-656.

Evans, P. R. (1997) *Joint CCP4 and ESF-EACBM Newsletters* **33**, 22-24.

Haas, D. & Leisinger, T. (1975). *Eur. J. Biochem.* **52**, 365-375

Jancarik, J. & Kim, S.H. (1991). *J. Appl. Cryst.* **24**, 409-411.

Laemmli, U. K. (1970). *Nature* **227**, 680-385

Leslie, A.G.W. (1990). *Crystallographic Computing*, Oxford University Press, Oxford.

Marina, A., Uriarte, M., Barcelona, B., Fresquet, V., Cervera, J. & Rubio, V. (1998). *Eur. J. Biochem*, **253**: 280-291.

Marina, A., Alzari, P.M., Bravo, J., Uriarte, M., Barcelona, B., Fita, I. & Rubio, V. (1999). *Protein Sci.* **8**, 1-7.

Matthew, B.W. (1968). *J. Mol. Biol.* **33**, 491-497.

Navaza, J.(1994). *Acta Cryst.* **A50***:* 157-163.

Otwinowski, Z. (1993). Oscillation data reduction program. In *Data collection and processing* (Sawyer, L., Isaccs, N. & Bailey, S., eds), pp 56-62, SERC Daresbury Laboratory, Warrington. UK.

Parsot, C., Boyen, A., Cohen, G.N. & Glansdorff, N. (1988). *Gene* **68**, 275-283.

Capítulo 2

Structure of acetylglutamate kinase, a key enzyme for arginine biosynthesis and a prototype for the amino acid kinase enzyme family, during catalysis

Trabajo publicado en

Structure **(2002) Vol. 10, Pags. 329-342**

Structure of Acetylglutamate Kinase, a Key Enzyme for Arginine Biosynthesis and a Prototype for the Amino Acid Kinase Enzyme Family, during Catalysis

Santiago Ramón-Maiques[1], Alberto Marina[1,3,4], Fernando Gil-Ortiz[1,4], Ignacio Fita[2] and Vicente Rubio[1]

[1]Instituto de Biomedicina de Valencia, Consejo Superior de Investigaciones Científicas (IBV-CSIC), C/ Jaime Roig 11, 46010-Valencia, Spain.

[2]Instituto de Biología Molecular de Barcelona (IBMB-CSIC), C/ Jordi Girona 18-21, 08034-Barcelona, Spain.

[3]Present address: Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032.

[4] These authors have contributed equally to this work.

SUMMARY

N-Acetyl-L-glutamate kinase (NAGK), a member of the amino acid kinase family, catalyzes the second and frequently controlling step of arginine synthesis. The *Escherichia coli* NAGK crystal structure to 1.5 Å resolution reveals a 258-residue subunit homodimer nucleated by a central 16-stranded molecular β sheet sandwiched between α helices. In each subunit, AMPPNP, as an αβγ-phosphate- Mg^{2+} complex, binds along the sheet C-edge, and N-acetyl-L-glutamate binds near the dyadic axis with its γ-COO^- aligned at short distance from the γ-phosphoryl, indicating associative phosphoryl transfer assisted by: (1) Mg^{2+} complexation; (2) the positive charges on Lys8, Lys217 and on two helix dipoles; and (3) with hidrogen bonding with the γ-phosphate. The structure resemblance with carbamate kinase and the alignment of the sequences suggest that NAGK is a structural and functional prototype for the amino acid kinase family, which differs from other acylphosphate-making devices represented by phosphoglycerate kinase, acetate kinase, and biotin carboxylase.

INTRODUCTION

Arginine is needed to make proteins and, together with its precursor ornithine, for making polyamines and urea, and it is a precursor of the important energy storage compounds creatine phosphate and arginine phosphate [1]. Microorganisms, plants and to a limited extent a number of

animals including mammals [2-6] synthesize arginine from ornithine by the reactions of the urea cycle [7] and derive the ornithine from glutamate. Ornithine synthesis in microorganisms and plants generally uses a route [2, 3] involving N-acetylated intermediates (Figure 1) in which glutamate is first N-acetylated, its γ-COOH group is then successively phosphorylated, reductively dephosphorylated and aminated, and the resulting N-acetyl-L-ornithine is deacylated either hydrolytically as in *Escherichia coli* or by transacetylation to glutamate, as in *Pseudomonas aeruginosa* [2]. In the latter case N-acetyl-L-glutamate (NAG) is regenerated and NAG phosphorylation, catalyzed by NAG kinase (NAGK) becomes the main controlling step of the route, being feed-back inhibited by the end product arginine [2, 8]. A different controlling role of NAGK is observed in *Saccharomyces cerevisiae*, an organism in which although N-acetylornithine is deacylated hydrolytically, NAGK forms a complex with NAG synthetase that is essential for the activity of the synthetase [9].

Animals make arginine using similar reactions to those of microorganisms and plants but do not utilize N-acetylated derivatives (Figure 1) [10, 11], and, in fact, there appears to be exquisite specificity for the N-acetylated or non-acetylated intermediates in one or the other pathway [8, 11, 12] providing a differential trait for interfering selectively with microorganism or plant arginine biosynthesis.

We have chosen for structural study NAGK, the enzyme that phosphorylates NAG in the bacterial route of ornithine synthesis. Little structural information existed for this enzyme other than it is a homodimer of a 27-29 kDa subunit [13, 14] and that in *Pseudomonas aeruginosa* it can associate into higher oligomers [13]. This enzyme may be an appropriate target for selective inhibition of arginine biosynthesis because it is not found in animals [11] and catalyzes an obligatory, and in many cases controlling step in the organisms that make arginine via NAG [2, 8]. In contrast, NAG synthetase would appear a less suitable target because in those organisms that recycle the acetyl group in the route of ornithine synthesis it only plays a purely anaplerotic role [2], and, in addition, NAG synthetase exists in animals that make urea, where it generates the NAG that is needed as an essential activator of carbamoyl phosphate synthetase I, the first enzyme of the urea cycle [15]. Thus, inhibition of NAG synthetase would result in hyperammonemia and abolition of arginine and urea synthesis in animals [16].

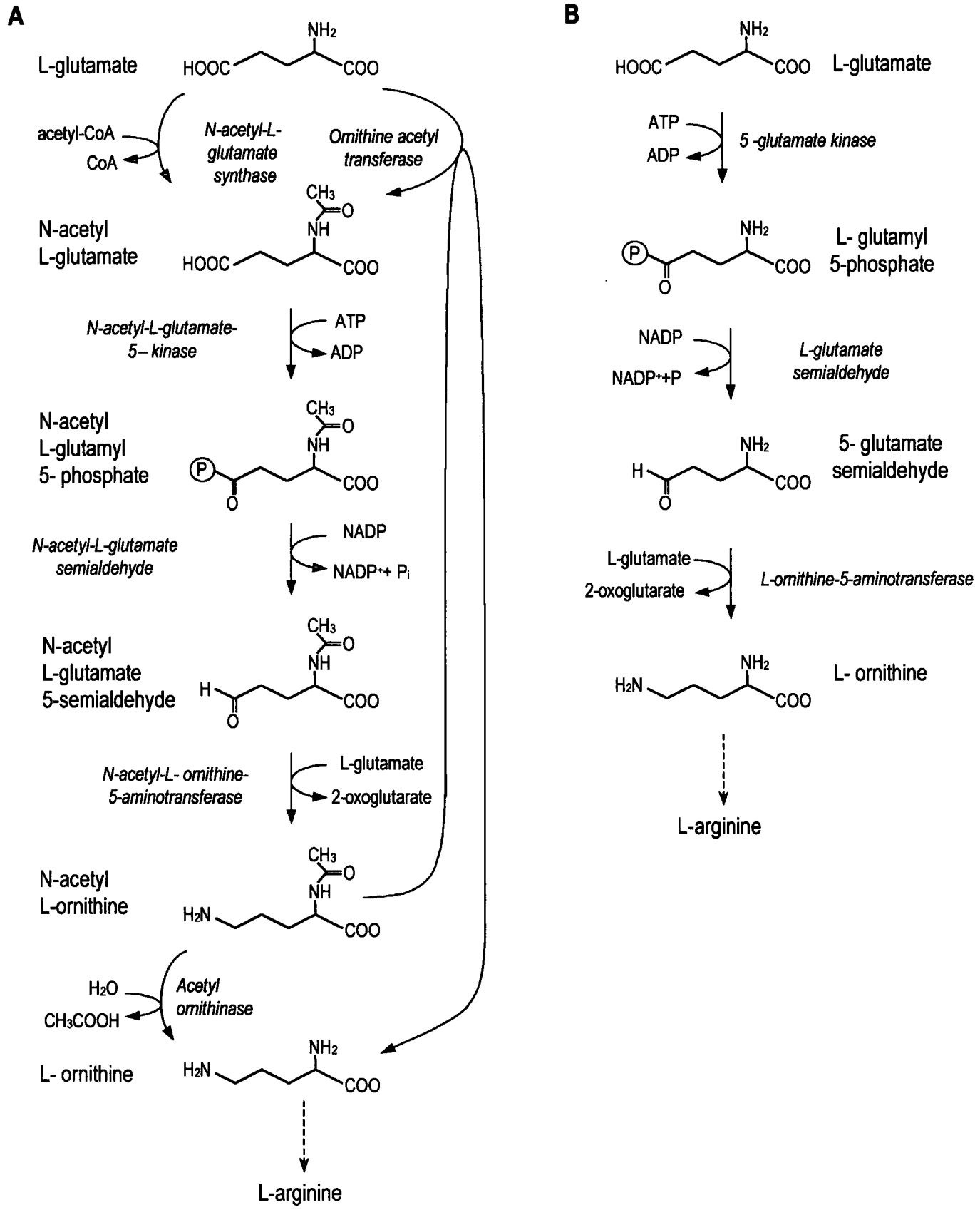


Figure 1. Acetylated (A) and Non-acetylated (B) Routes of Ornithine Biosynthesis.

We report here the structure at atomic resolution, determined using X-ray crystallography and selenomethionine (SeMet) labeling, of *Escherichia coli* NAGK complexed with the substrate NAG and with the inert nucleotide substrate analog AMPPNP. The structure shows that NAGK is a homodimer and reveals an organization nucleated by an open, mainly parallel β sheet that propagates across the dimer interface and spans the entire molecule, and which is surrounded on both sides by α helices. By detailing the structure of the ternary complex, the present results also clarify the mode of binding of the substrates in NAGK, and allow to propose a specific catalytic mechanism. Comparison with other enzymes of known structure demonstrates a strong resemblance of NAGK with carbamate kinase (CK) [17,18]. Thus, the present structure appears to be a basic fold among enzymes that make acylphosphate bonds. Given the sequence and reactional similarities, this fold is likely to be shared also by the enzymes glutamate-5 kinase and aspartokinase, which are of crucial importance for the biosynthesis of several amino acids [1, 10, 12, 19], and also by bacterial UMP kinase [20]. The information provides also insight into the structure, binding and catalysis in this enzyme family.

RESULTS AND DISCUSSION

Overall fold

MAD phasing of SeMet-substituted *E. coli* NAGK (Table 1) allowed the bulding of a refined model to 1.85 Å resolution of the crystalline ternary complex with NAG and AMPPNP. This model was used to solve by molecular replacement the structure of the same complex of the non-SeMet substituted enzyme at 1.5 Å resolution. This structure, agreeing very well (rmsd of 0.231 Å for all Cα atoms) with the model of the Se-Met enzyme and exhibiting excellent stereochemistry (no residues in the disallowed region and only one residue in the generously allowed region of the Ramachandran plot [21]), encompasses residues 1-258 as well as NAG and AMPPNP and will be described here.

As expected [13, 14], NAGK is a homodimer. The NAGK subunit exhibits an open α-β fold and is composed of 16 β strands, 8 α helices, and connecting loops and turns (Figure 2), being nucleated by an eight-stranded mainly parallel β sheet (at increasing distance from the dyadic axis, β5, β8, β2, β1, β11, β15, β16, β14, Figure 2b) sandwiched between a layer of four α helices (αD, αE, αG and αF) and another layer of three α helices (αC, αA and αH). Although α C is the longest helix (28 residues), it is broken at the middle by a wider turn that accommodates an extra residue (Ala81) and that has an abnormal hydrogen bond structure.

Table 1. Data Collection, Phasing and Refinement Statistics[a]

Data Collection and Phasing	SeMet			Native
	λ1 Edge	λ2 Peak	λ3 Remote	
Beamline	DESY X31	DESY X31	DESY X31	ESRF ID14-1
Wavelength (Å)	0.9770	0.9765	0.94	1.0
Resolution range (Å)	19.87-2.15	29.8-1.82	19.85-2.0	53.71-1.50
(outer shell)	(2.26-2.14)	(1.92-1.82)	(2.11-2.0)	(1.58-1.50)
Reflections (Total/ Unique)	60,371/12,703	106,899/20,642	75,090/15,674	196,777/37,500
I/σ_I	9.6 (3.4)	11.0 (2.5)	8.1 (2.5)	9.2 (2.1)
R_{sym}[b] (%)	7.4 (22.1)	5.9 (28.1)	8.5 (28.9)	4.2 (34.9)
Completeness (%)	98.9 (93.6)	98.8 (97.8)	96.8 (96.0)	99.9 (99.8)
FOM[c]		0.59 (0.40)		-

Refinement	SeMet	Native
Resolution range(Å)	29.88-1.85	53.71-1.50
Reflections (work/test)	18,682/1,005	35,599/1,861
R factor[d] (work/test) (%)	19.87/22.69	20.88/21.28
Average B ($Å^2$)		
Protein atoms	19.00	28.46
N-acetyl-L-glutamate	22.52	25.50
AMPPNP	44.2	34.89
Water	29.55	39.81
Magnesium	64.9	40.86
Number of atoms	2,147	2,147
Number of:		
Protein atoms	1,904	1,904
N-acetyl-L-glutamate molecules	1	1
AMPPNP molecules	1	1
Water molecules	198	198
Magnesium ions	1	1
Rmsd bond (Å)	0.005	0.012
Rmsd angle (°)	1.26	2.08
Ramachandran plot[e]		
Most favored (%)	93.2	93.6
Additional allowed (%)	6.3	5.9
Generously allowed (%)	0.5	0.5
Disallowed (%)	0	0

[a] Values in parenthesis are data for the highest resolution shell.
[b] $R_{sym} = \Sigma|I - \langle I \rangle|/\Sigma I$, where I is the observed intensity, and <I> is the average intensity of multiple observations of symmetry-related reflections.
[c] Figure of Merit calculated with program SOLVE.
[d] R factor = $\Sigma_{hkl}||F_{obs}|-|F_{calc}||/\Sigma_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are the observated and calculated structure factors, respectively.
[e] Calculated using PROCHECK [21].

However, this does not result in kinking of this helix. Helix B does not sandwich the β sheet, extending from the C-edge of β2 as an oblique prolongation of this strand in the direction of the intersubunit interface and connecting with the N-terminus of helix C through a large loop (residues 57-68, Figure 2c). This loop includes a β hairpin (β3-β4) and, in the view of Figure 2a, hangs over the β sheet C-edge. Another two large loops (residues 104-129 and 147-159, Figure 2c) each of which contains a β hairpin (β6-β7 and β9-β10) protrude over the β sheet C-edge

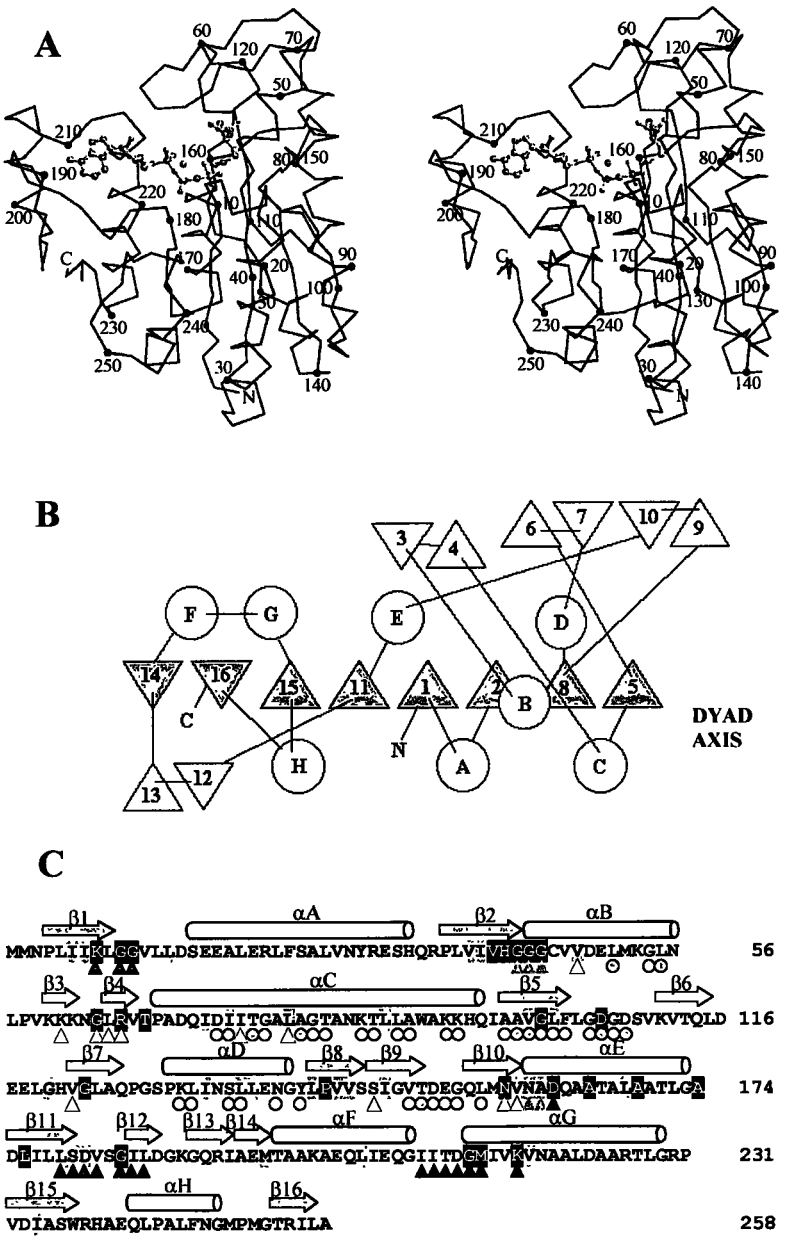


Figure 2. The NAGK Subunit.

(A) Stereoview of the Cα-trace. AMPPNP and NAG are shown in ball and stick, in gray.

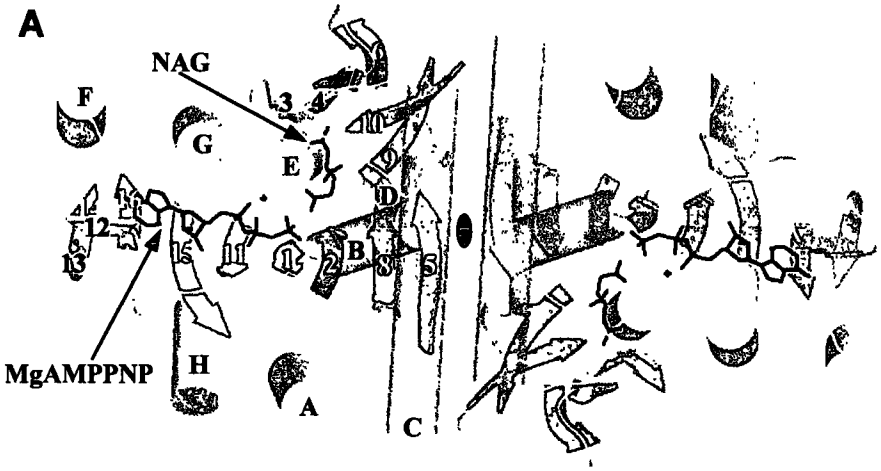
(B) Subunit topology. The main β sheet is shadowed. The arrow indicates the dyadic axis.

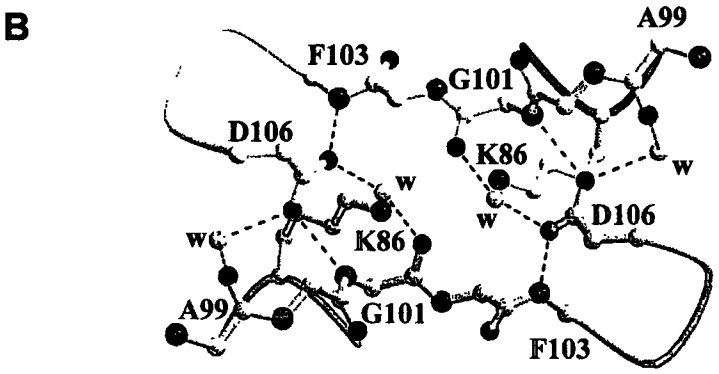
(C) Correspondence between the amino acid sequence and the secondary structure. The main β sheet is shadowed. Triangles under the sequence denote residues having decreased accessibility upon the binding of NAG (open), AMPPNP (closed), or both (shadowed). Shadowed circles denote decreased accessibility upon dimer formation. The black and the grey sequence backgrounds indicate, respectively, conservation or conservative replacement (I-L-M-V, A-G, D-N, T-S) in at least 22 of 24 genuine or putative NAGK sequences (Swissprot accession numbers, P11445, 9951641, O67848, O28988, Q07905, P36840, P57157, Q59281, O08320, Q60382, O26285, P94989, P31595, P73326, Q9X2A4, O50147, Q9YBY9, Q9V1I5, Q9SCL7, Q9PIR8, Q9PEM7, Q9LCS6, Q9L1A3, Q9HTN2 and Q9KNT7).

near the dimer interface, emerging from β5 and β8 and connecting back with αD and αE, respectively. Strands β6, β7, β9 and β10 form a small β sheet via parallel connections between β7 and β10 (Figure 2b). These three loops, together with β strands 1, 2, 5 and 8 and helices A-E constitute a compact N-terminal lobe (N-lobe, residues 1-173) that binds NAG and forms the dimer interface. The remainder of the polypeptide chain (residues 174-258) forms a smaller, C-terminal lobe (the C-lobe), that is composed of the final four strands of the main β sheet (β11 and 14-16), helices F-H, and two large loops that connect helices F and G (residues 209-212) and β11 to β14 (residues 184-193), including in the latter loop β hairpin β12-β13. The C-lobe binds the ADP moiety of the nucleotide (Figure 2a).

The dimer

The dimer, of 90 x 50 x 45 Å, is nucleated by the main β sheet (Figure 3a), which is continued across the dimer interface by non-canonical antiparallel connections between β5 from the two subunits. The hydrogen bonds formed between β5 and β5' are mediated by intervening water molecules and the β-COO^- of Asp106 (Figure 3b). In this way an open, 16-stranded, molecular β sheet is formed that spans the dimer from end to end. The mainly flat interface (Figure 3c) of 1279 $Å^2$ (determined as the inaccessible surface to a probe of radius 1.4 Å [22]) represents 11 % of the surface of each subunit. It is formed exclusively by residues of the N-lobe of each subunit (Figure 2c), in a number of 38, with a predominance of non-polar residues (22; 58 %) and atoms (71 %), and involves strand β5 and helices C and D, forming a cross-grid with the same elements of the other subunit, and helix B and strand β9 (Figure 3c). Helix C (residues 69-96) provides 13 of the 38 interfacial residues, including Ile75 and Leu89, which anchor hydrophobically αC in the other subunit; Asp74 which makes main chain and side chain direct and water-mediated hydrogen bonds with the β-hydroxyls of Thr83 and Thr87, from helix C of the other subunit; and lysines 86 and 93, that through their ε-NH_3^+ are bridged to aspartates 106 (through a water molecule, Figure 3b), 108 and 152 of the other subunit. The involvement of Asp106 in these interactions and in the hydrogen bond network linking strand β5 from the two subunits (Figure 3b) explain the conservation of this residue among NAGKs . The hydrogen bond network between the β5 strands cuts approximately at the middle a hydrophobic patch (Figure 3c) that interacts with the equivalent patch from the other subunit and that is formed by the side chains of Val100, Leu102, Phe103, from β5, Leu89 and the carbon side-chain of Lys86,


A
NAG
F
G
E
D
B
MgAMPPNP
H
A
C


B
A99
F103
G101
D106
K86
w
w
w
w
D106
K86
A99
G101
F103

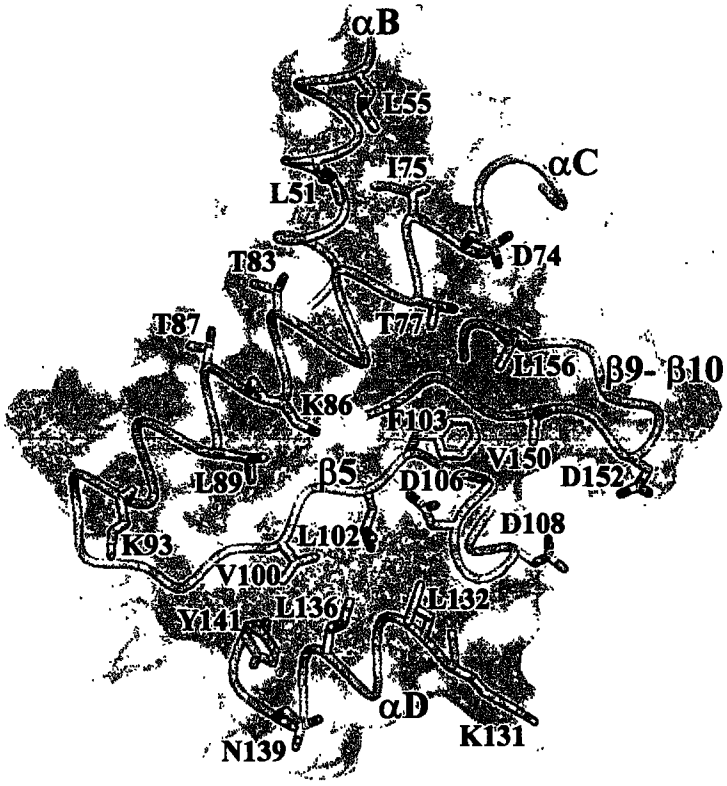

C
αB
L55
I75
αC
L51
D74
T83
T87
T77
L156
β9- β10
K86
F103
V150
D152
β5
L89
D106
D108
K93
L102
V100
L136
L132
Y141
αD
K131
N139

Figure 3. The NAGK dimer.
(A) Cartoon representation of the dimer with the dyadic axis (black ellipse) perpendicular to the paper and the main β sheet (colored green for one monomer and blue for the other) with its C-edge pointing towards the viewer. Substrates are shown in black, in bond representation. The small β sheet is colored orange, the β hairpins are yellow and α helices are gray in both subunits. Helices B and C are semitransparent to allow viewing the structure behind them. β strands and helices are identified with numbers and letters in one subunit.
(B) Hydrogen bonds mediated by strand β5, Asp106 and Lys86. Except for Asp106 and Lys86, side chains have been omitted for clarity. Different colors are used for the two subunits. Water molecules are light blue, and broken red lines denote hydrogen bonds.
(C) Semi-transparent representation of the dimerization surface (calculated with MSMS [57]) of one subunit, viewed from the other subunit, with its non-polar atoms colored red. Secondary stucture elements involved in intersubunit interactions and important residues of the interface are represented and identified. Prepared with DINO [58].

from helix C, Leu132 and Leu136 from helix D and the nearby Tyr141, and Val150 and Leu 156 from β9 and β10, respectively. The loop system of the N-lobe is anchored in the interface through the connection between β5 and β6 (involving Asp106) and through the tip of the β9-β10 hairpin (involving the contacts between Asp152 and Lys93, and hydrophobic contacts of nearby residues). The intersubunit surface is extended excentrically over the C-edge of the main β sheet (Figure 3c) by a hydrophobic patch formed by Leu51 and Leu55 from αB and Ile75 from αC, which interacts with the corresponding patch of the other subunit.

The extension and elaborate organization of the surface, which involves central secondary structure elements of each subunit and which provides many anchoring points of varied chemical nature, including six highly conserved residues (Figure 2c), supports the importance of dimer formation in NAGK, although the lack of evidence for cooperative kinetics with this enzyme [8, 23], would restrict the role of dimer formation to a purely structural one, such as the provision of an scaffold for anchoring the complex loop system that form the substrate sites in each subunit.

Binding of NAG

Despite the important roles of NAG (arginine precursor, *nod* signal in plants [24], carbamoyl phosphate synthetase and phosphate-dependent glutaminase activator [15, 16, 25, 26], brain metabolite [27]), there was no structural information on NAG binding to a protein. Figures 2a and 4a show NAG bound in the N-lobe of NAGK, in a pocket formed by the three loops of this lobe and by αB and the N-terminal half of αC, over the C-edge of strands β1 and β2 of the main β sheet. The pocket is separated from the dimer interface by a wall having no holes or

tunnels, that is formed by the β9-β10 hairpin and by helices C and B. The pocket opening is in the opposite wall, oriented towards the C-lobe, and is limited (Figure 4a) by the β2-αB and β10-αE connections and by the β3-β4 hairpin, which forms a lid over bound NAG. The side chains of Lys61 and Arg66 extend this lid towards αB, with which Arg66 is connected through hydrogen bonding and salt bridging of its guanidinium group with the side chain of Asp49. NAG is bound within this pocket in an extended, low energy conformation [28], with all the non-hydrogen atoms except the α and γ carboxylates approximately coplanar, and with the carboxylates emerging from the opposite faces of the plane of the molecule at angles of approximately 60° (in

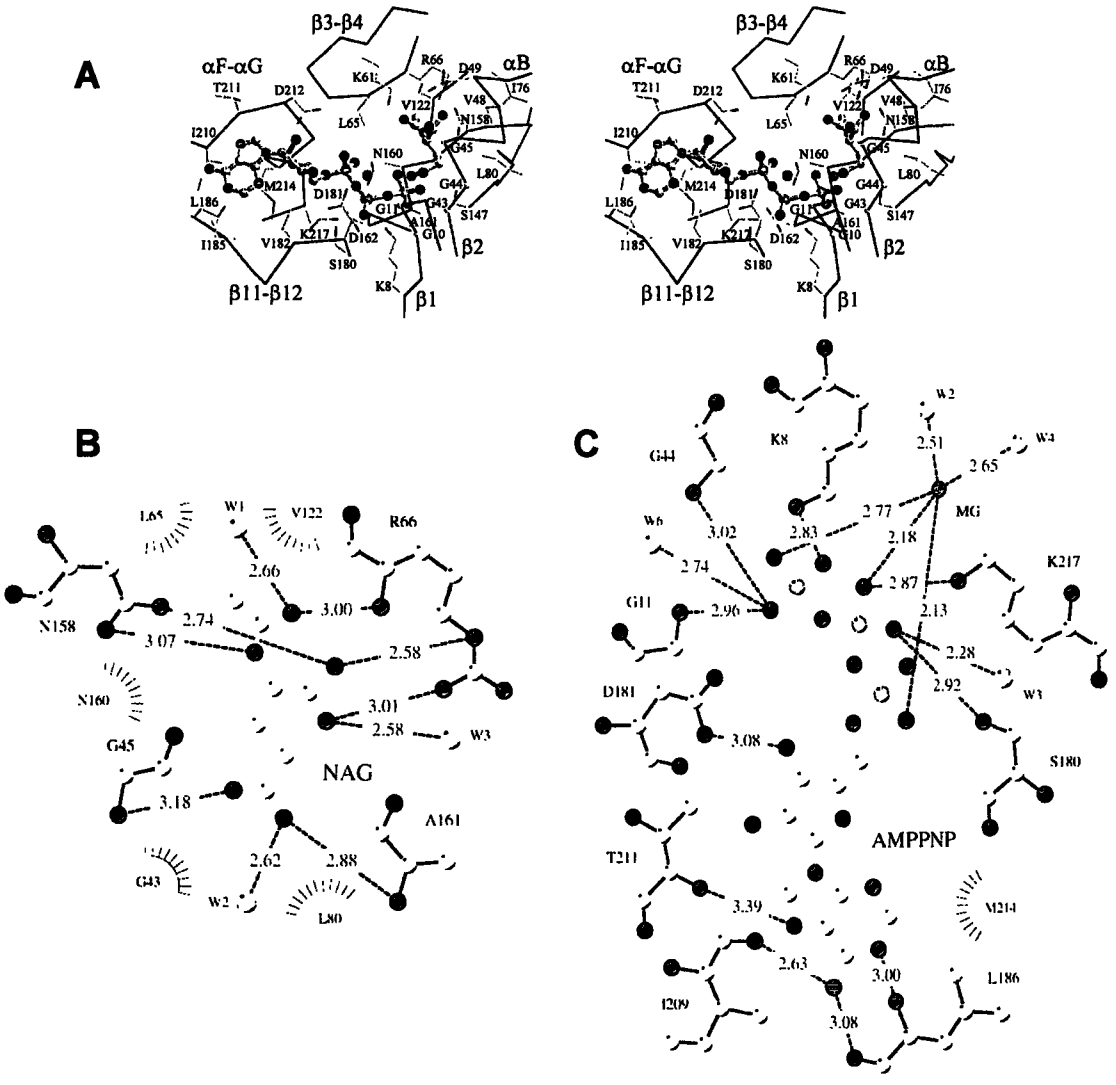


Figure 4. Binding of the Substrates to NAGK.
(A) Stereoview of the Cα-trace of the binding sites. AMPPNP and NAG are shown in ball and stick representation, and colored. Amino acid side chains are in color, in thinner trace.
(B and C) Schematic plots (drawn with LIGPLOT [59]) of the interactions between the protein and NAG (b) or AMPPNP (c). The letter w denotes a water molecule. Distances are in Å.

the line joining the α and γ NAG carbons). The NAG γ-COO^- sits at the entry of the NAG pocket, sandwiched between the β10αE and β2αB connections, fixed in position for attack on the nucleotide γ-phosphate by hydrogen bonds linking one of its O atoms to the N atom of Gly45 and linking the other to both the N atom of Ala161 and to a water molecule of the Mg^{2+}coordination sphere (Figure 4b). The NAG α-COO^- sits deep in the pocket, docked on the positive surface provided by the guanidinium group of Arg66 and making hydrogen bonds with the δ-N of Asn158, and with a water (Figure 4b) that is bonded to the ε-N of Lys61. Repulsive interactions between the guanidinium group of Arg66 and the positive charge of the α-NH_3^+ of L-glutamate could explain the inability of the enzyme to use glutamate [23]. The central atoms of the NAG molecule interact extensively with β10, whereas the N-acetyl group points towards β3-β4 and fits tightly into the hydrophobic cavity formed by Leu65, Val122 and the hydrocarbon parts of Arg66 and Asn160 (Figure 4a). The limited size of this cavity explains the inability of NAGK to use glutamate derivatives with large N-acyl groups and the ability to use N-carbamoyl-L-glutamate and N-formyl-L-glutamate [8, 23], which have similar or smaller groups that the acetyl group. Hydrogen bonds are formed (Figure 4b) between the N and O atoms of the acetamido group of NAG and the side chain O and main chain N of Asn158 and Arg66, respectively, and also between the acetyl O atom and a water molecule that bridges it to both Gly64 and the δ-N of Asn160. These bonds are expected also to contribute to the specificity of NAGK for NAG. The extended conformation of bound NAG is likely to account for the inability of NAGK to use as substrate N-acetyl-L-aspartate [8], in which the β-COO^- would be misplaced for attack on the ATP γ-phosphate.

As might be expected, given their high functional value, some residues forming the NAG pocket are highly conserved (Figure 2c), including Gly64, Arg66 and Asn158 and two polypeptide segments at the β2-αB and β10-αE junctions that constitute genuine NAGK sequence motifs: ^{39}VIV**HGGG**xxV and ^{158}**NVNAD** (where bold type denotes conservation, V or I can be any of V, L or I, and x indicates any residue).

Nucleotide binding

AMPPNP binds extended in a groove formed along the main β sheet between the two large loops of the C-lobe, the C-edge of β11, β1 and β2, and helices G and E, thus spanning from the C-lobe, were the adenine sits near the outward tip, to the opening of the NAG pocket, where lies the γ-phosphate (Figures 2a, 3a and 4a). The adenine ring, in the rare syn configuration [29],

is contoured by the large loops of the C-lobe thanks to the drastic turns allowed by Gly184 and Gly213 (both highly conserved, Figure 2c). The purine lies flat over the side chains of the conserved hydrophobic residues Ile185, Ile210 and specially Met214 (Figure 4a), with its opposite face exposed except for peripheral van der Waals contacts with Leu186, Thr211 and Asp212 (the residue found in the generously allowed Ramachandran region, possibly because of substrate-induced stress). Hydrogen bonds formed with the adenine N atoms, 6-N with the O atoms of Ile209 and Leu186, N1 with the NH of Leu186, and possibly (distance, 3.4 Å) N7 with the NH of Thr211 (Figure 4c), explain the inability of NAGK to use GTP or ITP [8, 23]. The ribose moiety binds in the narrowest part of the nucleotide groove (Figure 4a), with a face of the ring making contacts with the side-chains of Met214 and Val182, the other face shielded in part by the side chain of Asp212, and the endocyclic 4'O atom pointing towards helix G. The 2'- and 3'-OH are largely exposed, although the latter is hydrogen-bonded to the β-COO^- of Asp181 (a highly conserved residue having double conformation of which only one conformation binds to the 3'-OH). The lack of protein contacts of the 2'-OH explains the utilization by NAGK of dATP [8].

The extended polyphosphate does not cross over the main β sheet C-edge and makes protein contacts only through the β and γ phosphates. However, the three phosphates are complexed through non-bridging O atoms with the Mg^{2+} (a rare, but previously reported in other enzymes [30, 31] mode of coordination). The metal sits near the β-phosphate and is coordinated also with two water molecules (Figures 4a, 4c and 5a), one of which is hydrogen bonded to the NAG γ-COO^- (see above) and the other to the N atom of Gly213 and the β-COO^- of Asp162. Asp162 may be crucial, since it is absolutely conserved (Figure 2c) and since its β-COO^- is interposed between (Figure 4a) and hydrogen bonded to the ε-NH_3^+ groups of conserved Lys217 and Lys8, which, together with this aspartate and with Ser180, form the floor of the nucleotide groove under the β phosphate. Lys217 and Lys8 are hydrogen bonded to the Mg^{2+}-complexed O atom of the β-phosphate and to a non-bridging O atom of the γ phosphate, respectively (Figure 4c). In this way Asp162 may fix in appropriate position the two lysine side-chains that provide positive surfaces for binding the β and γ phosphates. The non-bridging O atom of the β-phosphate that is not coordinated with Mg^{2+} makes hydrogen bonds with the OH group of Ser180 (in one of its two conformations; Ser180 exhibits double conformation) and with a water molecule (Figure 4c) that is hydrogen bonded also to, the ε-NH_3^+ group of Lys8 and the main-chain N and O atoms of Gly10 and Leu 179, respectively. The γ-phosphate binds near the β1-αA and β2-αB junctions,

both of which exhibit conserved glycine-containing sequences (Figure 2c) that do not conform to the P-loop or to other known nucleotide-binding motifs [32], but that constitute true NAGK signatures: ^{6}I(I/f)**KxGG** (bold-type, full conservation; the I may be replaced by L or V or, when indicated with an "f", rarely by F; x, any residue) at the β1-αA junction, and the ^{39}VIV**HGGG**xxV described in the previous section. The non-bridging O atom of the γ-phosphate that is not bound to Mg^{2+} or to Lys8 is hydrogen bonded to the N atoms of Gly11 and Gly44, from these two signatures, and to a water molecule that is hydrogen bonded to the main-chain O and N atoms of Gly11 and of Gly46, respectively (Figure 4c). Considerable neutralization of the negative potential of the γ-phosphate must be effected by the positive end of the α helix B dipole, since the N-terminus of this helix is nearby and points towards the γ-phosphate (Figures 3a, 4a and 5a), and possibly also by the corresponding dipole of helix E, which is also near this phosphate.

Phosphoryl group transfer

The good quality of the map for the NAG γ-COO^- and the AMPPNP γ-phosphate, and the high resolution (Figure 5a), allow the unambiguous identification of the attacking O atom of the NAG γ-COO^-, given its position and its short distance from the γ-phosphate. In the model shown in Figure 5a the angle $O_{NAG}P_{\gamma}N_{\beta\gamma}$ (O_{NAG}, attacking O atom; $N_{\beta\gamma}$, bridging N atom in AMPPNP) has 155°, corresponding to a deviation of only 12.5° from ideality of the apical bonds in a potential trigonal bipiramid transition state for an in-line attack [33]. The P_{γ}-O_{NAG} distance of 2.8 Å is 2.1 Å too short for a dissociative mechanism, corresponding to fractional bond formation of 0.016 between the O_{NAG} atom and the P_{γ}, supporting the associativity of the phosphoryl transfer [34]. The bond number must be higher for the transition state, since the distance from the planar transition state P atom (assuming its localization at the center and within the plane of the three equatorial O atoms) and the attacking and leaving groups can be estimated as 2.31 and 2.15 Å, respectively, corresponding to 0.108 and 0.199 bond formation, and thus the mechanism would be, in the terminology of Mildvan [34], at least 10.8 % associative. A remarkable experimental finding of the present data is the observation of continuity in the electron density between the terminal phosphate of AMPPNP and the attacking O atom of the NAG γ-COO^- (Figure 5a) that must reflect bond formation between the attacking

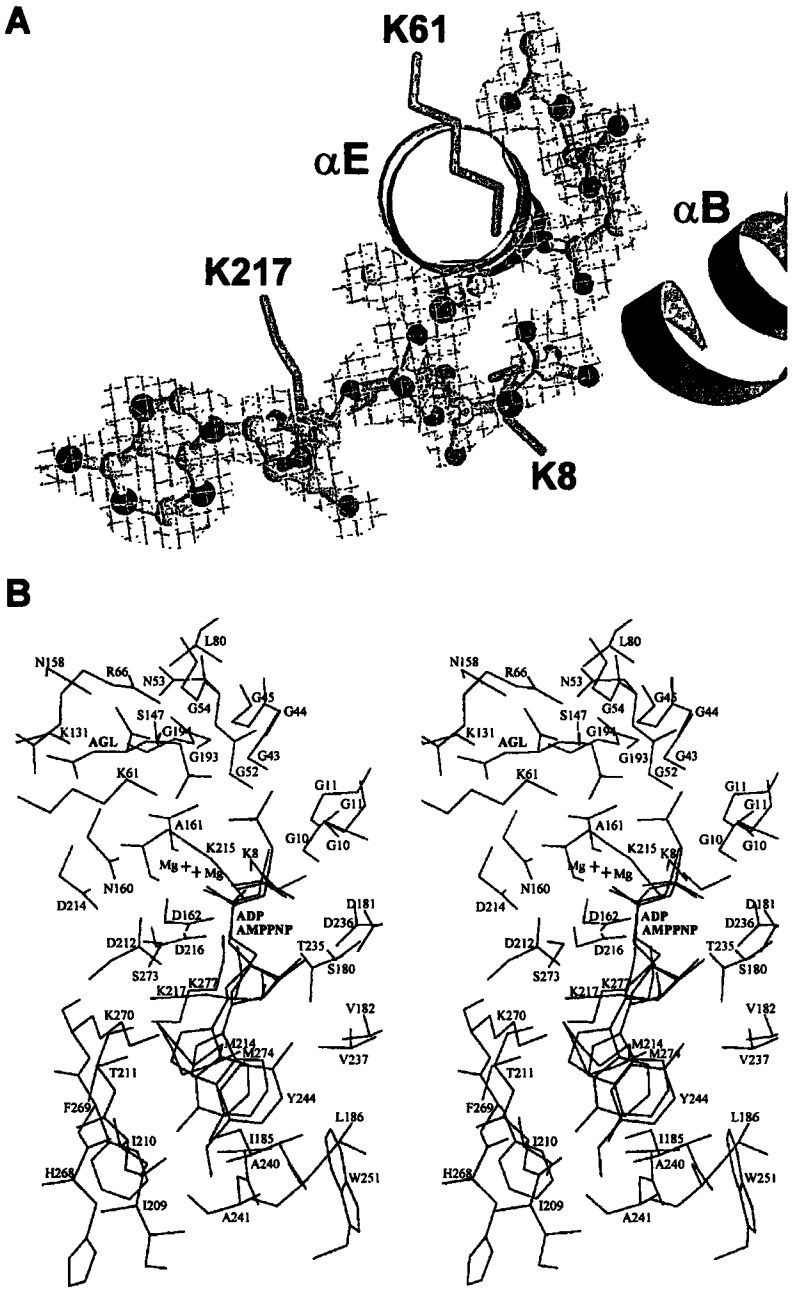


Figure 5. $2F_{obs}$-F_{calc} density map and model of MgAMPPNP and NAG bound to the enzyme, and comparison with ADP binding by CK-like CPS.

(A) $2F_{obs}$-F_{calc} density map contoured at 0.8 σ for bound MgAMPPNP (including in light blue two water molecules coordinated to Mg) and NAG. Helices B and E and the side chains of lysines 8, 61 and 217 are shown. Drawn with BOBSCRIPT [60].

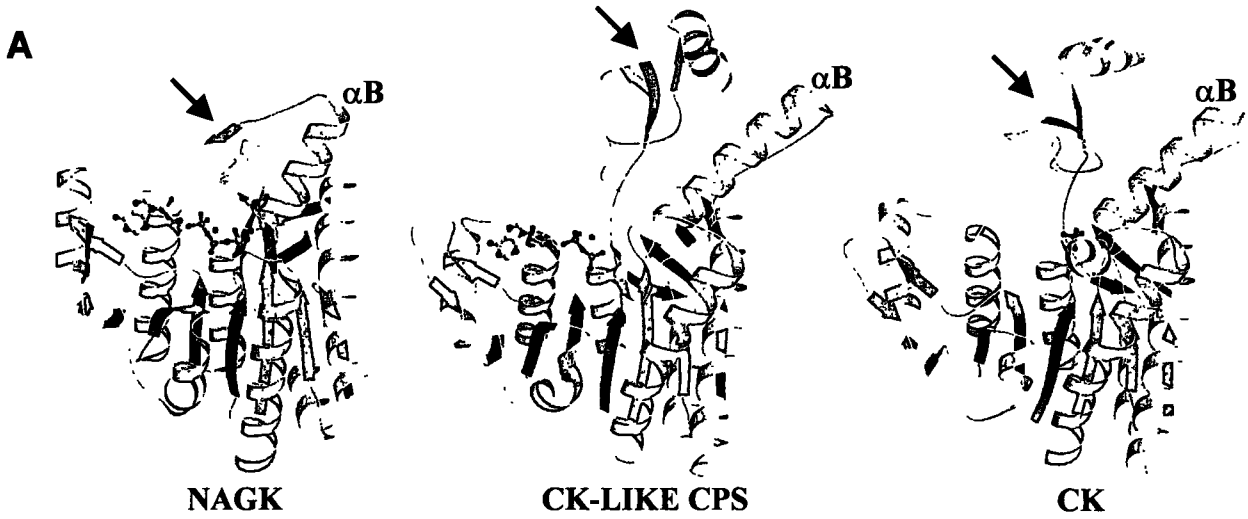
(B) Stereoview of bond representation of bound MgAMPPNP, NAG and nearby NAGK groups (in black) superimposed with bound ADP and nearby groups in CK-like CPS from *Pyrococcus furiosus* (in red; Protein Data Bank accession number 1e19).

O atom and the Pγ, and thus the present structure corresponds to an intermediate stage in the reactional path between substrates and products, representing a genuine snapshot of the enzyme in the act of catalyzing the formation of the new P-O bond. This would appear compatible with the inertness of AMPPNP as a substrate of kinases [35], since the effect of the bridging N atom of AMPPNP must be to hamper N-P bond breakage rather than to prevent new bond formation at the γ-P, in agreement with the present observations.

The organization of groups around the reacting AMPPNP γ-phosphate appears well suited for associative phosphoryl group transfer (Figures 4a and 5). The complexed Mg^{2+} ion, the positive end of the helix B dipole, and perhaps also of the nearby helix αE, the positive charge of the side-chain of Lys8 and the three donor hydrogen bonds from Gly11, Gly44 and a water molecule should withdraw electrons away from the Pγ atom, increasing its susceptibility to nucleophilic attack, and should shield and stabilize the equatorial negative charges of the pentavalent phophorus transition state. The position of the NAG γ-COO^-, fixed by the hydrogen bonds to the N atoms of Gly45 and Ala161, assures the proper stereochemistry of the reacting substrate groups for effective attack. The ε-NH_3^+ of Lys217 is positioned to help the development of a negative charge in the leaving group. Lys61 points towards the γ-phosphate but is too far from it (approximately 5 Å from its O atoms) for neutralization of the negative charge in the transition state (Figures 4a and 5b), and thus its role is at best uncertain, although in carbamate kinase (see below) a similarly placed lysine may be involved. Given the charged nature at neutral pH of the attacking and leaving groups, there is no need neither structural evidence for acid-base catalysis.

NAGK, carbamate kinase (CK) and carbamate kinase-like carbamoyl phosphate synthetase (CK-like CPS) define a characteristic fold involved in acylphosphate formation.

The comparison of the structure of NAGK with the reported structures [17, 18] for the CK from *Enterococcus faecalis* with a bound sulphate, and for the CK-like CPS of *Pyrococcus furiosus* with bound MgADP, confirms and extends the similarity proposed on the basis of the finding of 18-22% sequence identity [36] between NAGK and these other -COO^- phosphorylating enzymes. The three enzymes are homodimers nucleated (Figure 6a) by a central β sheet of 8 elements sandwiched between two layers of α helices in equal numbers and topologies [17, 18]. The superposition of the polypeptide chains of NAGK and CK-like CPS


A
αB
αB
αB
NAGK
CK-LIKE CPS
CK

B

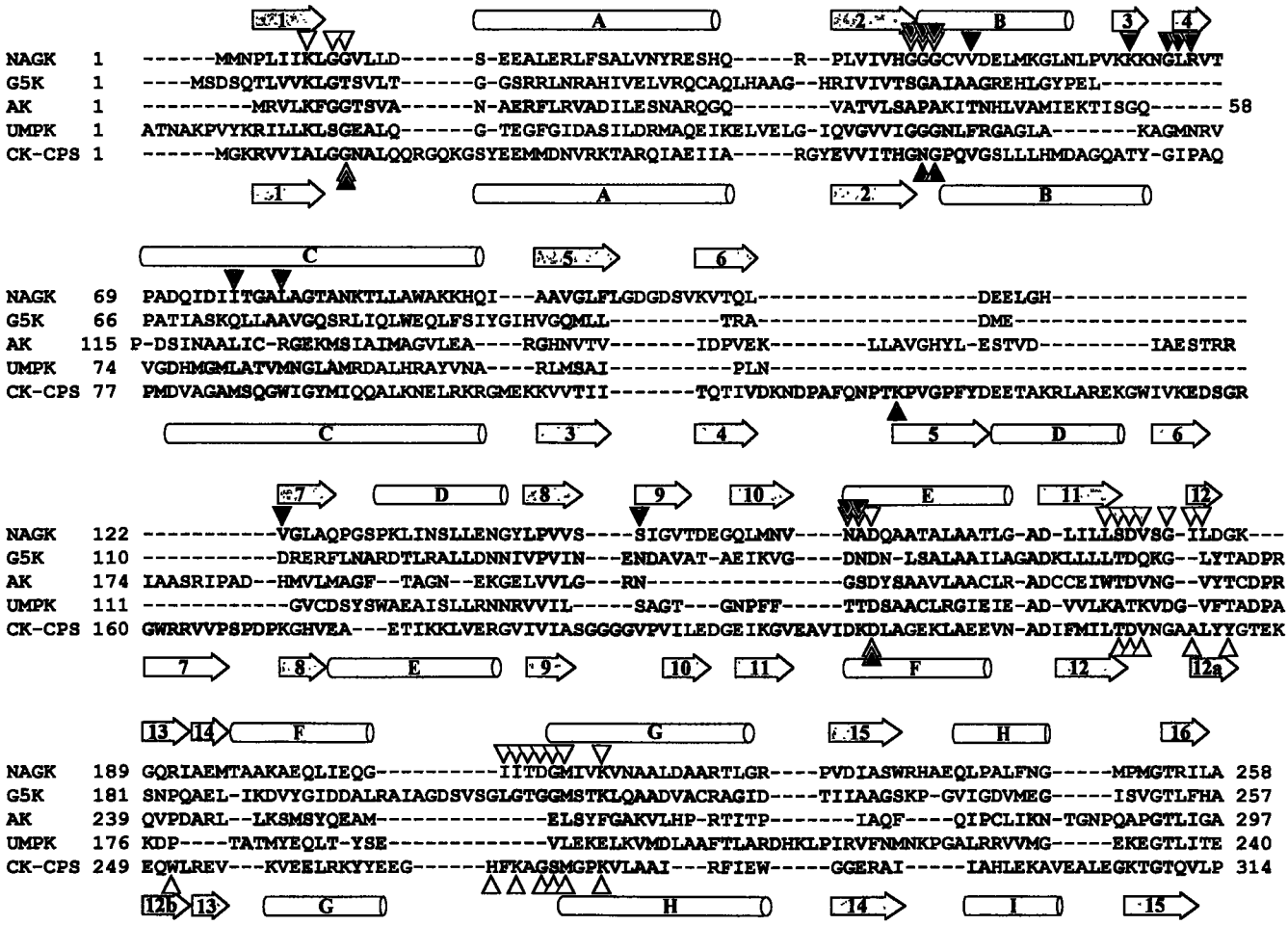

NAGK 1 ------MMNPLIIKLGGVLLD------S-EEALERLFSALVNYRESHQ-----R--PLVIVHGGGCVVDELMKGLNLPVKKKNGLRVT
G5K 1 ----MSDSQTLVVKLGTSVLT------G-GSRRLNRAHIVELVRQCAQLHAAG--HRIVIVTSGAIAAGREHLGYPEL----------
AK 1 ---------MRVLKFGGTSVA------N-AERFLRVADILESNARQGQ--------VATVLSAPAKITNHLVAMIEKTISGQ------ 58
UMPK 1 ATNAKPVYKRILLKLSGEALQ------G-TEGFGIDASILDRMAQEIKELVELG-IQVGVVIGGGNLFRGAGLA-------KAGMNRV
CK-CPS 1 ------MGKRVVIALGGNALQQRGQKGSYEEMMDNVRKTARQIAEIIA-----RGYEVVITHGNGPQVGSLLLHMDAGQATY-GIPAQ
NAGK 69 PADQIDIITGALAGTANKTLLAWAKKHQI---AAVGLFLGDGDSVKVTQL----------DEELGH----------
G5K 66 PATIASKQLLAAVGQSRLIQLWEQLFSIYGIHVGQMLL---------TRA----------DME----------
AK 115 P-DSINAALIC-RGEKMSIAIMAGVLEA----RGHNVTV-------IDPVEK--------LLAVGHYL-ESTVD---------IAESTRR
UMPK 74 VGDHMGMLATVMNGLAMRDALHRAYVNA----RLMSAI----------PLN----------
CK-CPS 77 PMDVAGAMSQGWIGYMIQQALKNELRKRGMEKKVVTII-------TQTIVDKNDPAFQNPTKPVGPFYDEETAKRLAREKGWIVKEDSGR
NAGK 122 ----------VGLAQPGSPKLINSLLENGYLPVVS----SIGVTDEGQLMNV----NADQAATALAATLG-AD-LILLSDVSG-ILDGK---
G5K 110 ----------DRERFLNARDTLRALLDNNIVPVIN---ENDAVAT-AEIKVG----DNDN-LSALAAILAGADKLLLLTDQKG--LYTADPR
AK 174 IAASRIPAD--HMVLMAGF--TAGN--EKGELVVLG---RN----------GSDYSAAVLAACLR-ADCCEIWTDVNG--VYTCDPR
UMPK 111 ----------GVCDSYSWAEAISLLRNNRVVIL-----SAGT---GNPFF-----TTDSAACLRGIEIE-AD-VVLKATKVDG-VFTADPA
CK-CPS 160 GWRRVVPSPDPKGHVEA---ETIKKLVERGVIVIASGGGGVPVILEDGEIKGVEAVIDKDLAGEKLAEEVN-ADIFMILTDVNGAALYYGTEK
NAGK 189 GQRIAEMTAAKAEQLIEQG----------IITDGMIVKVNAALDAARTLGR----PVDIASWRHAEQLPALFNG-----MPMGTRILA 258
G5K 181 SNPQAEL-IKDVYGIDDALRAIAGDSVSGLGTGGMSTKLQAADVACRAGID----TIIAAGSKP-GVIGDVMEG-----ISVGTLFHA 257
AK 239 QVPDARL--LKSMSYQEAM----------ELSYFGAKVLHP-RTITP-------IAQF----QIPCLIKN-TGNPQAPGTLIGA 297
UMPK 176 KDP----TATMYEQLT-YSE----------VLEKELKVMDLAAFTLARDHKLPIRVFNMNKPGALRRVVMG-----EKEGTLITE 240
CK-CPS 249 EQWLREV---KVEELRKYYEEG------HFKAGSMGPKVLAAI---RFIEW-----GGERAI-----IAHLEKAVEALEGKTGTQVLP 314

Figure 6. Comparison of AminoAcid Kinase Pfam Family Members.
(A) Ribbon diagrams of the monomers of NAGK with bound MgAMPPNP and NAG, of CK-like CPS from *P. furiosus* with bound MgADP (PDB 1e19) and of CK from *Enterococcus faecalis* with bound sulphate (PDB 1b7b). Ligands are represented as ball and stick models. An arrow signals the peripheral domain of CKs and the corresponding β3-β4 hairpin of NAGK. Helix B is also identified.
(B) Alignment of representative amino acid kinase Pfam family members (PF00696; http://www.sanger.ac.uk/cgi-bin/Pfam): *E. coli* NAGK, glutamate 5 kinase (G5K), aspartokinase (AK) and UMP kinase (UMPK) and *P. furiosus* CK-like CPS (CK-CPS). The Pfam alignment was modified to maximize coincidence of secondary structure elements of NAGK and CK-CPS (arrows and cylinders above and below the corresponding sequences; blue and pink arrows, main and small β sheets, respectively; orange arrows, β12-β13 hairpin or equivalent in CK-CPS; colored in green are the β3-β4 hairpin of NAGK and the peripheral domain of CK-CPS), and to maximize the stacking of regions of high conservation (highlighted with a yellow backgound) within each enzyme class. Some important residues that are conserved among the different enzyme classes are colored red. An internal region of *E. coli* AK (residues 59-114) is missing in many AKs and is not aligned. Residues involved in the binding of the nucleotide are indicated with orange triangles, those involved in binding NAG or sulphate with blue triangles.

gives a rms deviation of 2.17 Å for 141 α-carbon atoms. A major difference is the replacement of the peripheral subdomain that in CKs hangs over the C-edge of the main β-sheet [17, 18] by the topologically non-equivalent β3-β4 hairpin that in NAGK forms a lid over the NAG site (Figure 6a). Another difference involves the organization of the small β sheet of 4 elements, where the parallel contacts are formed between the second and the fourth strands (β7 and β10, ordered according to the sequence) and the first and third strands (β4 and β10 for CK-like CPS) in NAGK and CK, respectively. A third difference involves the intersubunit connections mediated by β5 (or β3 in CKs), which are canonical for antiparallel chains in CKs [17, 18], but not in NAGK (see above). In addition to these differences, helix B is in NAGK substantially shorter and projects less across the dimer interface than in CK (Figure 6a). In all other respects, the similarities between NAGK and CK or CK-like CPS predominate: strands β1, β2, β5, β6, β10, β11, β14 and β15 (numbering of NAGK) have identical length; β7, β8, β9, β12 and β13 differ by only one residue; and α helices except helix B differ in length by less than 4 residues (Figure 6b). Loops and turns exhibit the largest differences, being generally longer in CKs. Structure-based alignment with NAGK (Figure 6b) reveals in CK-like CPS a 46-residue insertion, corresponding to the entire peripheral subdomain, and 24 % identity, with 6 gaps per 100 residues, in the shared region. Very good similarity exists in the region of the NAGK signature sequences, which among CKs are highly conserved but are somewhat different from those in NAGKs: ^{6}VIALGG,

^{49}IxHGNGPQVG and ^{212}**VIDKD** (bold type, full conservation; other capitals, conservative replacement or high but incomplete conservation; x stands for any residue) instead of the respective NAGK signatures at β1-αA, β2-αB, and β10-αE.

The similarities extend to the active center, judged by the superimposition of NAGK with the MgADP site and surrounding regions of CK-like CPS (Figure 5b) that reveals a similar overall location and conformation of the bound nucleotide and of the Mg^{2+}, and general coincidence of the polypeptide segments surrounding it, including the conservation of the methionine (Met274 of CK-like CPS) on which the adenine lies, of Asp181 and Val182 (Asp236 and Val237 of CK-like CPS) that interact with the ribose, and of Asp162 and Lys217 (Asp216 and Lys277 of CK-like CPS) in the floor of the groove near the β phosphate, whereas Ser180 of NAGK is replaced conservatively by Thr235. The β1αA and β2αB junctions appear in CK to be involved also in the binding of the reacting groups, given their localization and the observation of interaction with the bound sulphate of the N atoms of Gly11, Asn51 and Gly52 of enterococcal CK (Gly11, Asn53 and Gly54 of CK-like CPS, corresponding to Gly11, Gly44 and Gly45 of NAGK). The COO^- group of carbamate may bind in an equivalent position to that of the NAG γ-COO^-, although there is too much room in CK-like CPS to explain the high specificity of CKs for carbamate [36]. However, a movement of approximation towards the active center of the peripheral subdomain of CK was revealed when the structures of the CK-like CPS and of enterococcal CK having a bound sulphate were compared [18].

This movement would result in the filling of the excess space by the side chain of Lys128 (Lys131 of CK-like CPS), a residue that was found making a hydrogen bond through its ε-NH_3^+ group with the bound sulphate [17] and that may be involved in neutralizing the negative charge of the transferred phosphoryl group in the transition state.

Nevertheless, there are also some differences between NAGK and CKs in the nucleotide-binding site (Figure 5b). CK-like CPS has considerably higher apparent affinity for the nucleotide [37] than NAGK [8, 23], and in line with this, in CK-like CPS the adenine is enclosed in a pocket formed by Tyr244, which is stacked over the purine ring and that is hydrogen bonded to the 2'OH of the ribose. In NAGK the adenine is exposed, the ribose makes no hydrogen bonds through its 2'OH and there is no equivalent tyrosine. Another difference affects a crucial residue, Lys8 of NAGK which interacts with the γ-phosphate and appears to be centrally involved in the catalysis of the phosphoryl transfer. There is no sequence-equivalent lysine in CKs (Figure 6b), but a conserved lysine emerging from a distant site in the amino acid sequence was found binding to the

sulphate in enterococcal CK (Lys209 in this CK, Lys215 in CK-like CPS) [17] and the ε-NH_3^+ of this lysine is found in the same place as that of Lys8 of NAGK (Figure 5b).

NAGK characterizes the structure of the amino acid kinase Pfam family.

Based on the amino acid sequence and reactional similarities, we proposed [14, 36] that NAGK and glutamate-5-kinase may resemble CK, and these enzymes, together with aspartokinase and bacterial UMP kinase, are grouped in the Pfam database (http://www.sanger.ac.uk/cgi-bin/Pfam) as the "amino acid kinase" family (PF00696). The structure and conservation-guided alignment shown in Figure 6b suggests that all these enzymes share the same basic structure characteristic of NAGK and CK, since the 8 strands of the main β sheet (except perhaps strand β14, that may be missing in UMPK) and the surrounding α helices appear conserved. Since this conservation includes the interfacial elements αC and β5, all these enzymes may form dimers similar to those of NAGK and CK, in agreement with experimental data for aspartokinases II and III [19] and *Pseudomonas aeruginosa* glutamate 5 kinase [12]. Since helix B appears implicated in the catalysis of phosphoryl group transfer (see above), its constancy in all these enzymes as well as the conservation of the β1-αA and β2-αB junctions, which are involved in the positioning of the reactive groups, supports a similar catalytic mechanism for the entire family. The alignment agrees with the expectation of differences in the mode of binding of the different substrates used by these enzymes (carbamate, NAG, glutamate, aspartate, UMP). Thus, the β3-β4 hairpin which in NAGK is involved in NAG binding appears to be present among the other enzymes only in NAGK and UMP kinase, and the alternative peripheral subdomain found in CK and apparently involved in carbamate binding [17, 18] may be present only in aspartokinase, whereas no hairpin or peripheral subdomain appear present in glutamate-5-kinase. Further, the β9-β10 hairpin, which in NAGK is also involved in NAG binding, appears missing in aspartokinase. These differences support the binding of the enzyme-specific substrate in the same place in all these enzymes. The alignment supports also the binding of ATP in essentially the same way and site in all these enzymes, since a number of residues involved in the binding of this nucleotide are conserved, such as the adenine-stacking Met214 (numbering of NAGK), which is conserved or replaced by Leu, the ribose 3'OH binding Asp181, conserved except in UMP kinase, or the polyphosphate site residues Asp162, which is fully conserved and which when replaced in UMP kinase by Asn caused inactivation [38], or Lys8,

fully conserved or replaced in CK by Lys215, and Ser180, either conserved or replaced conservatively by threonine in all the family members except UMP kinase.

How many molecular machines to synthesize acylphosphates?

The present structural results, the comparisons with the structure of CK and the sequence comparisons with the other members of the amino acid kinase Pfam family puts on firm ground a type of structural device for making acylphosphates that is epitomized by NAGK. Indeed, the present report provides the first snapshot with a high degree of resolution and detail of an acylphosphate-making enzyme in the process of catalyzing the reaction. However, not all the enzymes that make an acylphosphate share the NAGK fold. Acetate kinase, which exhibits significant sequence identity with formiate kinase and butyrate kinase but not with the members of the amino acid kinase family, represents another type of fold for catalyzing acylphosphate formation, as shown by the recent structure of its ADP complex that revealed that this enzyme presents the ASKHA (acetate and sugar kinases/Hsc70/actin) superfamily fold [39], differing widely in architecture and mode of nucleotide binding with NAGK. Another classical acylphosphate forming enzyme, phosphoglycerate kinase, for which the 3-D structure is known for many years [40], is composed of two independently folded domains of similar size, one of them binding the nucleotide and the other the 3-phosphoglycerate. The extended two-domain structure is associated with large-scale hinge bending conformational changes and domain closure; indeed, even in the most closed structure that was studied [30] domain closure was incomplete, since there was still a water molecule interposed between the reacting groups of the two substrates. Therefore, among the enzymes catalyzing the formation of an acylphosphate only in the case of NAGK the productive ternary complex has been characterized structurally in detail. In any case there are ample differences in architecture, domain folding and way of nucleotide binding between phosphoglycerate kinase, NAGK and acetate kinase, and thus phosphoglycerate kinase represents another type of structural device for acylphosphate synthesis. These three types of structure would account for all of the components of the 2.7.2. group of the Enzyme Commission, which assembles those enzymes that transfer a phosphoryl group to a carboxylate. However, carbamoyl phosphate synthetase [41], biotin-dependent carboxylases [42], D-ala-D-ala ligase [43] and glutathione synthetase [44] catalyze reactions in which a step is the phosphorylation by ATP of bicarbonate or a carboxylate, and even one of these enzymes, carbamoyl phosphate synthetase, catalyzes as a part of its mechanism the phosphorylation of carbamate, the same reaction that is catalyzed by CK (reviewed in [17]). However, the structures

of the phosphorylation components of these enzymes [43-46] do not resemble CK or NAGK, nor acetate kinase or phosphoglycerate kinase, representing a type of structure exemplified by biotin carboxylase, consisting of three domains and characterized by the type of nucleotide binding site called the grasp site [32, 44]. Thus, there are at least four types of structural devices for making acyl phosphates, having as paradigms NAGK, acetate kinase, phosphoglycerate kinase and biotin carboxylase, and the more complete structural characterization thus far of the central catalytic complex for any of them is reported here for NAGK.

BIOLOGICAL IMPLICATIONS

The amino acid arginine is an essential component of living beings, since it is a protein amino acid, a precursor of the energy storage compounds arginine phosphate and creatine phosphate, an intermediate in the production of urea, and, together with its precursor non-protein amino acid ornithine, a precursor of the polyamines. Plants, most microorganisms and many animals, including mammals, make arginine "de novo", invariably from ornithine which is synthesized from glutamate. However, there are two mechanisms to make ornithine, one in microorganisms and plants, in which glutamate is N-acetylated prior to phosphorylation, reductive dephosphorylation and transamination at its side-chain carboxyl group, ending with the removal of the N-acetyl group, and another that is characteristic of animals, in which the same steps take place but the glutamate derivatives are not acetylated. This provides an important difference that opens the way for selective inhibition of ornithine and arginine biosynthesis in microorganisms and plants. We have studied the structure of the enzyme catalyzing the second step of the N-acetylated route, N-acetyl-L-glutamate kinase (NAGK), because in contrast to N-acetyl-L-glutamate (NAG) synthetase, that exists in animals since NAG is an essential activator of ureagenesis, it is not found in animals and is in many organisms the controlling step of arginine biosynthesis.

Our studies on crystalline recombinant NAGK from *Escherichia coli* complexed with NAG and with the ATP analog AMPPNP have revealed that NAGK is a homodimer of a subunit having a central 8-stranded mainly parallel open β sheet, sandwiched between layers of 3 and 4 α helices. The sheet becomes continuous across the dimer interface, giving a 16-stranded molecular β sheet. An independent active center exists in each subunit, on the C-edge of this sheet, formed by an additional α helix and a system of three large loops, each containing a β-hairpin, and a smaller loop. The structure reported here provides a snapshot of NAGK catalyzing phosphoryl

group transfer by an associative in line mechanism, and reveals the implication in the catalysis of two conserved lysines and the positive charge at the N-end of at least one α helix. The comparison with other proteins demonstrates extensive structural similarity with the enzyme carbamate kinase, and allows to propose that the structure and mode of substrate binding and catalysis found in NAGK is shared also by carbamate kinase, glutamate-5-kinase, aspartokinase and bacterial UMP kinase, all of which catalyze the phosphorylation of a -COO^- or a phosphate and which constitute the amino acid kinase family of the Pfam database. Comparison with other enzymes that phosphorylate a COO^- has allowed to group them into four structural groups exemplified by NAGK, acetate kinase, phosphoglycerate kinase and biotin carboxylase.

EXPERIMENTAL PROCEDURES

Enzyme expression, SeMet-substitution, purification and crystallization

Overexpression and purification of plasmid-encoded *E. coli* NAGK were carried out as described [14], except for the use, when indicated, of SeMet-substituted enzyme [47]. Mass spectrometry (MALDI-TOF using the Voyager DE-Pro instrument of Applied Biosystems) demonstrated a mass excess consistent with the replacement of six methionines of the NAGK monomer by SeMet. The preparation of crystals of approximately 0.5 mm in the largest dimension in the presence of 24 mM NAG, 6 mM AMPPNP and 30 mM $MgCl_2$, using the vapor diffusion method was already described [14].

Data Collection and Processing

A crystal of the Se-Met substituted enzyme harvested in cryobuffer as described [14] was diffracted at 100 K at the three wavelengths indicated in Table 1, collecting data up to 1.85 Å resolution using an image-plate detector from Mar Research. Data were processed using MOSFLM, SCALA and TRUNCATE [48]. Crystals belong to space group $C222_1$ and present cell dimensions a=59.4 Å, b=71.7 Å and c=107.3 Å. A crystal of the native enzyme (no SeMet substituted) grown exactly as above was soaked 1 min in the cryobuffer supplemented with 12 mM NAG, 3 mM AMPPNP, 7.5 $MgCl_2$ and 10 mM KCl, and was then flash-freezed in the 100 K cryo-stream, followed by data collection to 1.5 Å resolution at beamline ID14-1 at the ESRF, Grenoble, using a MarCCD detector. Processing, space group and cell dimensions were identical to those above for the SeMet derivative.

Phasing, Model Building, and Refinement

In the crystals of the SeMet NAGK one protein chain was found in the asymmetric unit and the positions of 6 Se atoms were determined using SOLVE [49] and refined with SHARP [50]. Initial MAD phases were improved by density modification using histogram matching and solvent flattening as implemented in program DM [48]. A first model including 226 of the 258 residues in the polypeptide chain was traced and built using O [51]. The high quality of the map allowed the identification of 153 side chains for a total of 258 residues. Following completion of the model building process, automatic refinement was carried out, using program REFMAC [52], alternating with graphic model-building sessions with program O. B-factors restrain was gradually released as refinement progressed. All the diffraction data were used throughout the refinement process except the 5% randomly selected data for calculating the R_{free}. Solvent positions were automatically assigned using programs PEAKMAX and WATPEAK [48]. The final model (R value, 19.9 %; Rfree, 22.7 %) corresponds to a monomer including residues 1 to 258, one bound NAG and AMPPNP molecules, and 199 water molecules. Structure analysis with PROCHECK [21] indicated that all main chain torsion angles fall within the allowed regions of the Ramachandran plot (Table 1).

A partially refined model of the SeMet derivative structure was used for phasing the native enzyme crystal data by molecular replacement, using AMoRe [53]. A few cycles of model building and refinement, using program CNS [54] were enough to construct the model at 1.50 Å resolution. The high resolution data allowed the characterization of all the protein residues, one MgAMPPNP and one NAG molecules bound to the enzyme, and 198 water molecules. The final model has an R value of 21.3% (R_{free} 24.1%) and good stereochemistry (Table 1).

Calculations and illustrations

Superposition of structures was carried out with the lsq-option of program O using default parameters [51], and with program LSQKAB [48]. Buried surface areas were calculated using program NACCESS [22]. Except when indicated, molecular models were drawn with MOLSCRIPT [55] and RASTER3D [56].

ACKNOWLEDGMENTS

We acknowledge DESY (Hamburg) and ESRF (Grenoble) outstations for financial assistance and support for data collection, J.J. Calvete for mass spectrometry, and S. Hubbard for

use of NACCESS. The work was supported by Grants Rayos X of Fundación Ramón Areces and GV01-259 of the Generalitat Valenciana. S.R.-M. and F.G.-O. were fellows of the Generalitat Valenciana and of the Fundación Ramón Areces, respectively.

REFERENCES

1. Bender, D.A. (1975). Amino Acid Metabolism. (London: John Wiley and Sons).
2. Cunin, R., Glansdorff, N., Piérard, A. and Stalon, V. (1986). Biosynthesis and metabolism of arginine in bacteria. Microbiol. Rev. *50*, 314-352.
3. Shargool, P.D., Jain, J.C. and McKay, G. (1988). Ornithine biosynthesis, and arginine biosynthesis and degradation in plant cells. Phytochemistry *6*, 1571-1574.
4. Visek, W.J. (1986). Arginine needs, physiological state and usual diets. A reevaluation. J. Nutr. *116*, 36-46.
5. Synderman, S.E., Boyer, A. and Holt, E. (1959). The arginine requirement of the infant. Am. J. Dis. Child. *97*, 192-195.
6. Bell, J.M. and John, A.M. (1981). Amino acid requirements of growing mice: arginine, lysine, tryptophan and phenylalanine. J. Nutr. *111*, 525-530.
7. Krebs, H.A. (1976). The discovery of the ornithine cycle. In The Urea Cycle. S Grisolia, R Báguena and F Mayor, eds. (New York: John Wiley and Sons), pp. 1-12.
8. Haas, D. and Leisinger, T. (1975). N-acetylglutamate 5-phosphotransferase of *Pseudomonas aeruginosa*. Catalytic and regulatory properties. Eur. J. Biochem. *52*, 365-375.
9. Abadjieva, A., Pauwels, K., Hilven, P. and Crabeel, M. (2001). A new yeast metabolon involving at least the two first enzymes of arginine biosynthesis: acetylglutamate synthase activity requires complex formation with acetylglutamate kinase. J. Biol. Chem. *276*, 42869-42880.
10. Jones, M.E. (1983). Catalysis of the urea cycle. Trans. NY Acad. Sci. *41*, 77-82.
11. Alonso, E. and Rubio, V. (1989). Participation of ornithine aminotransferase in the synthesis and catabolism of ornithine in mice. Studies using gabaculine and arginine deprivation. Biochem. J. *259*, 131-138.
12. Krishna, R.V. and Leisinger, T. (1979). Partial purification and characterization of γ-glutamyl kinase. Biochem. J. *181*, 215-222.

13. Haas, D. and Leisinger, T. (1975). N-Acetylglutamate 5-phosphotransferase of *Pseudomonas aeruginosa*. Purification and ligand-directed association-dissociation. Eur. J. Biochem. *52*, 365-375.

14. Gil, F., Ramón-Maiques, S., Marina, A., Fita, I. and Rubio, V. (1999). N-acetyl-L-glutamate kinase from *Escherichia coli*: cloning of the gene, purification and crystallization of the recombinant enzyme and preliminary X-ray analysis of the free and ligand-bound forms. Acta Crystallogr. D *55*, 1350-1352.

15. Shigesada, K. and Tatibana, M. (1978). N-acetylglutamate synthetase from rat-liver mitochondria. Partial purification and catalytic properties. Eur. J. Biochem. *84*, 285-291.

16. Bachmann, C., Krähenbühl, S., Colombo, J.P., Schubiger, G., Jaggi, K.H. and Toenz, O. (1981). N-Acetylglutamate synthetase deficiency: a disorder of ammonia detoxication. N. Engl. J. Med. *304*, 543.

17. Marina, A., Alzari, P.M, Bravo, J., Uriarte, M., Barcelona, B., Fita, I. and Rubio, V. (1999). Carbamate kinase: new structural machinery for making carbamoyl phosphate, the common precursor of pyrimidines and arginine. Protein Sci. *8*, 934-940.

18. Ramón-Maiques, S., Marina, A., Uriarte, M., Fita, I. and Rubio, V. (2000). The 1.5Å resolution crystal structure of the carbamate kinase-like carbamoyl phosphate synthetase from the hyperthermophilic archaeon *Pyrococcus furiosus*, bound to ADP, confirms that this thermostable enzyme is a carbamate kinase, and provides insight into substrate binding and stability in carbamate kinases. J. Mol. Biol. *299*, 463-476.

19. Scapin, G. and Blanchard, J.S. (1998). Enzymology of bacterial lysine biosynthesis. Adv. Enzymol. Relat. Areas Mol. Biol. *72*, 279-324.

20. Serina, L. Blondin, C., Krin, E., Sismeiro, O., Danchin, A., Sakamoto, H., Gilles, A.M., and Bârzu, O. (1995) *Escherichia coli* UMP-kinase. a member of the aspartokinase family, is a hexamer regulated by guanine nucleotides and UTP. Biochemistry *34*, 5066-5074.

21. Laskowsky, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallog. 26, 283-291.

22. Hubbard, S.J., and Thornton, J.M. (1993). NACCESS, Computer Program, Department of Biochemistry and Molecular Biology, University College London.

23. Denes, G. (1973). N-Acetylglutamate-5-phosphotransferase. In The Enzymes, 3rd Edition, vol. IX-B, PD Boyer, ed (New York: Academic Press) pp. 511-520.

24. Philip-Hollingsworth, S., Hollingsworth, R.I., and Dazzo, F.B. (1991). N-Acetylglutamic acid: an extracellular nod signal of *Rhizobium trifolii* ANU843 that induces root hair branching and nodule-like primordia in white clover roots. J. Biol. Chem. *266*, 16854-16858.

25. Rubio, V., Britton, H.G., and Grisolia, S. (1983). Mitochondrial carbamoyl phosphate synthetase activity in the absence of N-acetyl-L-glutamate. Mechanism of activation by this cofactor. Eur. J. Biochem. *134*, 337-343.

26. Blackburn, E.H., Hird, F.J., and Jones, I.K. (1972). Metabolism of glutamine and ammonia in rat liver: the effects of N-acetylglutamate and phosphate. Arch. Biochem. Biophys. *152*, 265-271.

27. Alonso, E., Garcia-Perez, M.A., Bueso, J., and Rubio, V. (1991). N-acetyl-L-glutamate in brain: assay, levels, and regional and subcellular distribution. Neurochem. Res. *16*, 787-794.

28. Britton, H.G., Garcia-España, A., Goya, P., Rozas, I., and Rubio, V. (1990). A structure-reactivity study of the binding of acetylglutamate to carbamoyl phosphate synthetase I. Eur. J. Biochem. *22*, 47-53.

29. Saenger, W. (1984). Principles of Nucleic Acid Structure (New York: Springer-Verlag).

30. Auerbach, G., Huber, R., Grattinger, M., Zaiss, K., Schurig, H., Jaenicke., and Jacob, U. (1997). Closed structure of phosphoglycerate kinase from *Thermotoga maritima* reveals the catalytic mechanism and determinants of thermal stability. Structure *5*, 1475-1483.

31. Larsen, T.M., Benning, M.M., Rayment, I., Reed, G.H. (1998). Structure of the bis(Mg2+)-ATP-oxalate complex of the rabbit muscle pyruvate kinase at 2.1 Å resolution: ATP binding over a barrel. Biochemistry *37*, 6247-6255.

32. Vetter, I.R., and Wittinghofer, A. (1999). Nucleoside triphosphate-binding proteins: different scaffolds to achieve phosphoryl transfer. Q. Rev. Biophys. *32*, 1-56.

33. Knowles, J.R. (1980). Enzyme-catalyzed phosphoryl transfer reactions. Annu. Rev. Biochem. *49*, 877-919.

34. Mildvan, A.S. (1997). Mechanisms of signaling and related enzymes. Proteins: Struc. Func. Gen. *29*, 401-416.

35. Yount, R.G., Babcock, D., Ballantyne, W., and Ojala, D. (1971). Adenylyl imidodiphosphate, an adenosine triphosphate analog containing a P-N-P linkage. Biochemistry *10*, 2484-2489.

36. Marina, A., Uriarte, M., Barcelona, B., Fresquet, V., Cervera, J., and Rubio, V. (1998). Carbamate kinase from *Enterococcus faecalis* and *Enterococcus faecium*. Cloning of the genes, studies on the enzyme expressed in *Escherichia coli*, and sequence similarity with N-acetyl-L-glutamate kinase. Eur. J. Biochem. *253*, 280-291.

37. Uriarte, M., Marina, A., Ramón-Maiques, S., Fita, I., and Rubio, V. (1999). The carbamoyl phosphate synthetase of *Pyrococcus furiosus* is enzymologically and structurally a carbamate kinase. J. Biol. Chem. *274*, 16295-16303.

38. Bucurenci, N., Serina, L., Zaharia, C., Landais, S., Danchin, A., and Bârzu, O. (1998). Mutational analysis of UMP kinase from *Escherichia coli*. J. Bacteriol. *180*, 473-477.

39. Buss, K.A., Cooper, D.R., Ingram-Smith, C., Ferry, J.G., Sanders, D.A., and Hasson, M.S. (2001). Urkinase: structure of acetate kinase, a member of the ASKHA superfamily of phosphotransferases. J. Bacteriol. *183*, 680-686.

40. Banks, R.D., Blake, C.C., Evans, P.R., Haser, R., Rice, D.W., Hardy, G.W., Merret, M., and Phillips, A.W. (1979). Sequence, structure and activity of phosphoglycerate kinase: a possible hinge-bending enzyme. Nature *279*, 773-777.

41. Rubio, V., Britton, H.G., Grisolía, S., Sproat, B.S., and Lowe, G. (1981). Mechanism of activation of bicarbonate ion by mitochondrial carbamoyl phosphate synthetase: formation of enzyme-bound adenosine diphosphate from the adenosine triphosphate that yields inorganic phosphate. Biochemistry *20*, 1969-1974.

42. Climent, I., and Rubio, V. (1986). ATPase activity of biotin carboxylase provides evidence for initial activation of HCO_3^- by ATP in the carboxylation of biotin. Arch. Biochem. Biophys. *251*, 465-470.

43. Fan, C., Moews, P.C., Walsh, C.T., and Knox, J.R. (1994). Vancomycin resistance: structure of D-Alanine:D-Alanine ligase at 2.3 Å resolution. Science *266*, 439-443.

44. Yamaguchi, H., Kato, H., Hata, Y., Nishioka, T., Kimura, A., Oda, J., and Katsube, Y. (1993). Three-dimensional structure of the glutathione synthetase from *Escherichia coli* B at 2.0 Å resolution. J. Mol. Biol. *229*, 1083-1100.

45. Waldrop, G.L., Rayment, I., and Holden, H.M. (1994). Three-dimensional structure of the biotin carboxylase subunit of acetyl-CoA carboxylase. Biochemistry *33*, 10249-10256.

46. Thoden, J.B., Holden, H., Wesenberg, G., Raushel, F.M., and Rayment, I. (1997). Structure of carbamoyl phosphate synthetase: a journey of 96 Å from substrate to products. Biochemistry *36*, 6305-6316.

47. Budisha, N., Steipe, B., Denange, P., Eckerskorn, C., Kellermann, J., and Huber, R. (1995). High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in *Escherichia coli*. Eur. J. Biochem. *230*, 788-796.

48. CCP4 (Collaborative Computational Project 4) (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallog. D, *50*, 760-763.

49. Terwilliger, T.C., and Berendzen, J. (1999). Automated structure solution for MIR and MAD. Acta Crystallogr. D *55*, 849-861.

50. De la Fortelle, E., and Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for the multiple isomorphous replacement and multiwavelength anomalous diffraction methods. Methods Enzymol. *276*, 472-494.

51. Jones, T.A., Zou, J., Cowan, S., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110-119.

52. Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D *53*, 240-255.

53. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. Acta Crystallogr. A *50*, 157-163.

54. Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr.D *54*, 905-921.

55. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. *24*, 946-950.

56. Merritt, E.A., and Murphy, M.E.P. (1994). Raster3D version 2.0. A program for photorealistic molecular graphics. Acta Crystallogr D *50*, 869-873.

57. Sanner, M.F., Olson, A.J., and Spehner, J.C. (1996). Reduced surface: an efficient way to compute molecular surfaces. Biopolymers 38, 305-320.

58. Phillippsen, A. (1998). Dino, a visualization system for structural data. http//www.bioz.unibas.ch/~xray/dino.

59. Wallace, A.C., Laskowski, R.A., and Thornton, J.M. (1995). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng. *8*, 127-134.

60. Esnouf, R.M. (1999). Further additions to MolScript version 1.4, including reading and contouring of electron-density maps. Acta Crystallogr. D *55*, 938-940.

ACCESSION NUMBERS

The NAGK coordinates have been deposited in the Protein Databank under the accession codes 1gs5 for the native enzyme and 1gsj for SeMet-substituted NAGK.

Capítulo 3

A crystallographic glimpse of a nucleotide triphosphate (AMPPNP) bound to a protein surface. External and internal AMPPNP molecules in crystalline N-acetyl-L-glutamate kinase

Trabajo publicado en

Acta Crystallographica Section D **(2002) D58, Pags. 1892-1895.**

A crystallographic glimpse of a nucleotide triphosphate (AMPPNP) bound to a protein surface. External and internal AMPPNP molecules in crystalline N-acetyl-L-glutamate kinase

Fernando Gil-Ortiz[1], Ignacio Fita[2], Santiago Ramón-Maiques[1], Alberto Marina[1,†] and Vicente Rubio[1]

[1]Instituto de Biomedicina de Valencia, Consejo Superior de Investigaciones Científicas (IBV-CSIC), C/ Jaime Roig 11, 46010-Valencia, Spain.

[2]Instituto de Biología Molecular de Barcelona (IBMB-CSIC), C/ Jordi Girona 18-21, 08034-Barcelona, Spain.

[†]Present address: Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA

ABSTRACT

A large volume of unexplained electron density in the crystal of N-acetyl-L-glutamate kinase (NAGK) is now interpreted as an external, very extended, metal-free AMPPNP molecule that occupies two alternative positions and that makes contacts with the protein exclusively through its γ-imidophosphate. This external nucleotide is compared with the active-site nucleotide, and the reasons for its extended shape, lack of complexed metal and peripheral binding are analyzed. Further, the possibility that this bystander AMPPNP is waiting to occupy the active center is discussed.

INTRODUCTION

The crystal structure (Ramón-Maiques *et al.*, 2002) to 1.5 Å of the complex with N-acetyl-L-glutamate (NAG) and MgAMPPNP of *Escherichia coli* N-acetyl-L-glutamate kinase (NAGK), the enzyme that catalyzes the second and frequently controlling step of the route of arginine biosynthesis (Cunin *et al.*, 1986), revealed that NAGK is a homodimer nucleated by a molecular 16-stranded open β sheet sandwiched between α-helices. In each subunit, an N-terminal lobe binds NAG and forms the dimer interface, and the C-lobe binds the ADP moiety of

MgAMPPNP. The nucleotide lies extended along the sheet C-edge, with the γ-phosphoryl appropriately oriented for transfer to the NAG γ-carboxylate.

An unresolved aspect of the NAGK crystal structure was a large volume of electron density localized over the dyadic axis of the NAGK homodimer that could not be interpreted on the basis of the NAGK polypeptide chain model. This electron density is shown here to arise from an external AMPPNP molecule that occupies two alternative positions around a crystal symmetry two-fold axis coincident with the molecular dyad axis. The position of this nucleotide made it conceivable that it was waiting to occupy the nucleotide site at the active center, but the chemical characteristics of this nucleotide differ from those of the active-center-bound nucleotide, rendering this possibility unlikely. The reasons for the conformational stability of this nucleotide and its fixed localization in the protein crystal, as well as for its lack of Mg^{2+} complexation despite the abundance of $MgCl_2$ in the crystal mother liquor are also discussed.

EXPERIMENTAL METHODS

NAGK expression, purification and crystallization at pH 4.6 as the ternary complex with AMPPNP and NAG, and its structure determination at 100 K, to 1.5 Å resolution, at beamline ID14-1 at the ESRF, Grenoble have previously been described (Gil *et al.*, 1999; Ramón-Maiques *et al.*, 2002). The crystals (space group $C222_1$; unit-cell parameters a=59.4, b=71.7, c=107.3 Å), contain one protein subunit in the asymmetric unit with about 50% solvent. The final model (PDB code 1gs5) of the ternary NAGK-AMPPNP-NAG complex, (with agreement values of R = 20.88 %; R_{free} = 21.28 % and good stereochemistry) corresponds to a monomer including all protein residues (Met1-Pro258), 198 water molecules and one bound molecule each of NAG and MgAMPPNP. In addition, a large volume of electron density located in the vicinity of the crystallographic twofold axis coincident with the molecular dyad axis (Fig. 1a) is now interpreted as an AMPPNP molecule. A few cycles of model building with O (Jones *et al.*, 1991) and of refinement with REFMAC (Murshudov *et al.*, 1997), were carried out to construct the model for this extra AMPPNP. Bond lengths and angles for this modeled nucleotide fall in all cases within the expected ranges for AMPPNP (Saenger 1984). Difference and omit Fourier electron-density maps (Fig. 1b) were computed to cross-check the reliability of the proposed interpretation, particularly taking into account possible phasing artifacts arising from the proximity of the crystallographic symmetry axis. Figures were drawn with *MOLSCRIPT* (Kraulis, 1991),

BOBSCRIPT (Esnouf, 1999), *Raster*3D (Merritt & Murphy, 1994), and *DINO* (Phillippsen, 1998).

RESULTS

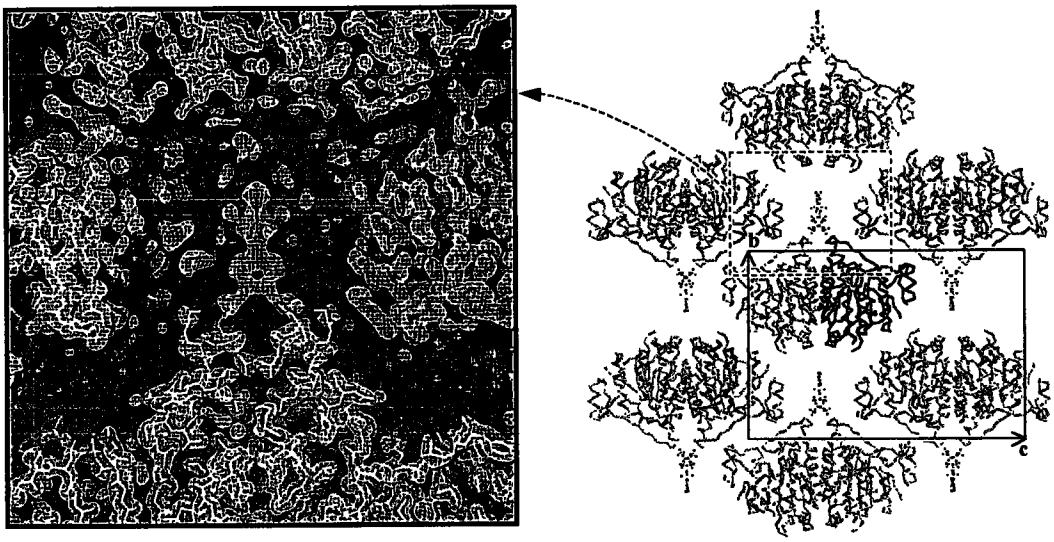
An AMPPNP molecule explains best the external density

A large volume (~150 Å^3) of electron density, having in projection a frog-like appearance (Fig. 1a), located between different enzyme molecules at the crystallographic two fold axis corresponding to the dimer dyadic axis, contacts NAGK through the lower part of the density (the frog legs), at the connection of helix αB with strand β3 (Fig. 1a). No such density has been observed in isomorphous crystals of the enzyme grown in the presence of ADP (data not shown). An external AMPPNP molecule having two symmetric conformations, each with 50% occupancy, accounts for this electron density (Ghosh *et al.*, 2000), which is of comparable density to the protein, and for the continuity of the density across the two fold symmetry axis (Fig. 1b). The adenine and the ribose, overlapping with the symmetrically related alternative position, fit the frog head and body and the two alternative positions of the AMPPNP fit within the rear frog legs. Some of the atoms are situated near the surface of the density map (Fig. 1b; in blue). These correspond to electronically light atoms that do not overlap with the symmetry-related molecule. The difference Fourier F_o-F_c map computed in the presence of this AMPPNP confirms the quality of the fit (Fig. 1b; in red). The only positive densities in the difference map are the two symmetrical peaks that evoke the frog front legs, possibly corresponding to either solvent molecules or to an alternative conformation of the adenosine moiety having low occupancy. R values remain essentially unchanged with the introduction of the model.

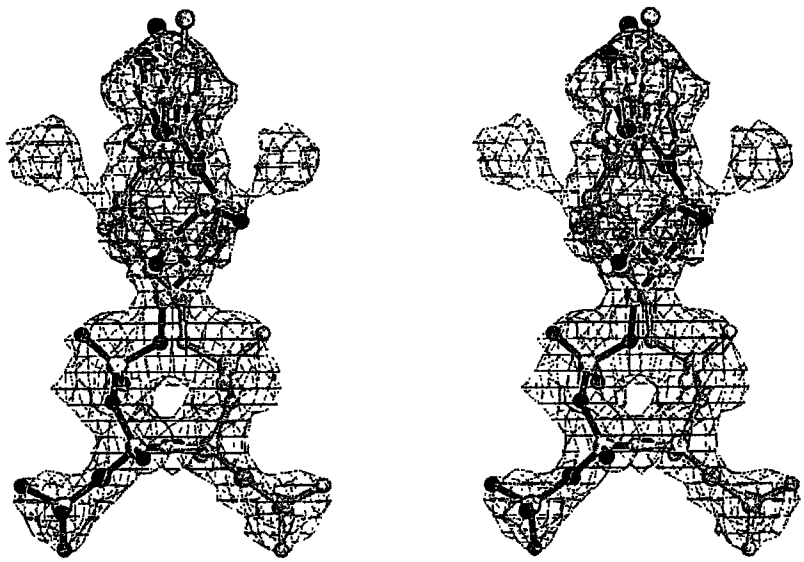
Conformation of the external AMPPNP, compared with that of active-site AMPPNP

The external AMPPNP molecule (Figs. 1b and 1c) presents good stereochemistry (Saenger, 1984) and has a conformation expected to be highly stable in solution. The adenine ring is *anti* (χ=140°), the ribose exhibits a C3'-*endo* pucker (pseudorotation angle values, P = 23°), and the γ torsion angle (C4′-C5′ bond) is in the limit for *–syn clinal* (-97°), allowing the staggering of C4' and C5' substituents and appropriately orienting the 3' and 5' O ribose atoms for formation of an internal canonical hydrogen bond (Fig. 1c) that may contribute importantly to the stability of the nucleotide extended form. Torsion angle β (C5′-O5′ bond) is antiperiplanar (154°), and the α and β phosphates are staggered, while the β and γ phosphates approach the

a)



b)



c)

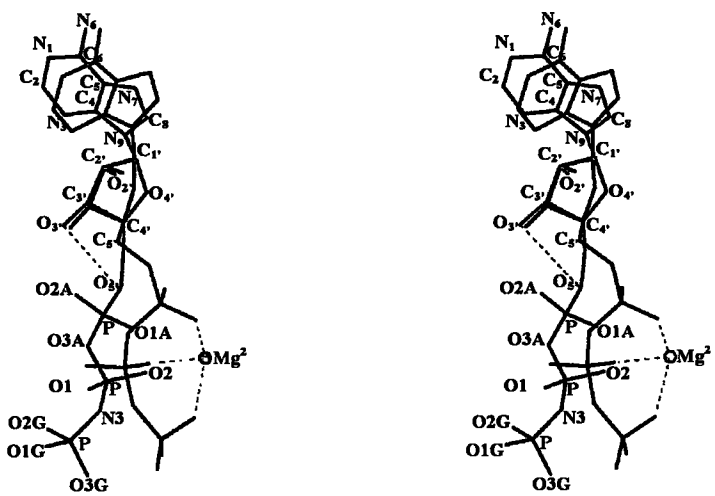

N6
N1
C6
C2
C5
C4
N7
N3
C8
N9
C1'
C2'
O2'
C3'
O4'
O3'
C4'
C5'
O3A
O2A
P
O1A
OMg2
O3A
O1
P
O2
O1
N3
O2G
O1G
P
O3G

Figure 1. (a) Left panel: $2F_o$-F_c map showing the unexplained electron density, having in projection a frog-like appearance, among four different enzyme molecules (bonds modeled in different colors) corresponding to the region enclosed in the broken-line rectangle of the right panel. This panel represents the crystal packing of the protein (in backbone representation) and also, in color, the external AMPPNP modeled in the previously unexplained density. One enzyme dimer is also colored within the unit cell, represented as a red rectangle. The *a* axis is perpendicular to the paper and thus is not shown. (b) Stereoview of F_o-F_c frog-like electron density, obtained in the absence (in blue) or in the presence (in red) of the AMPPNP model. The AMPPNP model is shown in ball-and-stick representation, colored in one of the two alternative conformations, and grey in the other conformation. (c) Stereo comparison of the internal (in black) and external (in orange) AMPPNP, in bond representation. Atoms nomenclature is also indicated. The intramolecular hydrogen bond is shown as a broken line in the external nucleotide. In the internal nucleotide, a Mg^{2+} ion is bound to the polyphosphate moiety in a tridentated coordination (broken black lines).

stable eclipsed conformation with no short contacts, possibly resulting from interactions with the protein. No electron density attributable to a complexed Mg^{2+} was observed in maps computed either without or with the nucleotide (Fig. 1b). At the 1.5 Å resolution of the present data, if Mg^{2+} were complexed, it would have been observed, as it was the case for the internally bound nucleotide (Ramón-Maiques *et al.*, 2002). At the pH of 4.6 and the concentrations of 3 mM AMPPNP and 15 mM $MgCl_2$ used in the crystallization drop, ~2/3 of the AMPPNP should not be complexed to Mg^{2+} (apparent stability constant for the MgAMPPNP complex at pH 4.6 is $K'_{pH4.6}$=35 M^{-1}; calculated from $K'_{pH4.6}$= $K'_{pH8.5}/\delta$ (O'Sullivan & Smithers, 1979); where $\delta = [1+10^{(pKa-4.6)}]/[1+10^{(pKa-8.5)}]$, assuming a pK_a of 7.7 for the γ-phosphate and a $K'_{pH8.5}$= 38,200 M^{-1} for AMPPNP (Yount *et al.*, 1971)).

The AMPPNP molecule bound at the active site is also highly extended (Ramón-Maiques *et al.*, 2002), but it differs importantly from the external AMPPNP molecule (Fig. 1c; Table 1). The adenine is in the *syn* configuration (χ=63°), the ribose exhibits a C4′-*exo* pucker (pseudorotation angle P = 66°) and the torsion angle γ is -antiperiplanar (-159°), allowing staggering of the C4′-C5′ substituents but preventing internal hydrogen-bond formation between the 3' and 5' O atoms. A Mg^{2+} ion is complexed with a non-bridging oxygen from the three phosphates in Λ, *exo* conformation (Saenger 1984), accounting for the eclipsed short contact conformation of Pα and Pβ and for the stable eclipsed conformation with no short contacts around the N atom bridging Pβ and Pγ.

Table 1. *Conformation of the two independent AMPPNP molecules bound to NAGK*

Torsion angle[b]	Defining atoms[c]	Torsion angle value (°)[a] External AMPPNP	Active-site AMPPNP
	O3G, P_γ, N3B, P_β	-154	-180
	P_γ, N3B, P_β, O3A	-94	-123
	N3B, P_β, O3A, P_α	-94	103
	P_β, O3A, P_α, O_5	-81	103
α	O3A, P_α, $O_{5'}$, $C_{5'}$	-156	59
β	P_α, $O_{5'}$, $C_{5'}$, $C_{4'}$	154	75
γ	$O_{5'}$, $C_{5'}$, $C_{4'}$, $C_{3'}$	-97	-159
δ	$C_{5'}$,$C_{4'}$, $C_{3'}$, $O_{3'}$	75	86
χ	$O_{4'}$, $C_{1'}$, N_9, C_4	140 (*anti*)	63 (*syn*)
ν_0	$C_{4'}$, $O_{4'}$,$C_{1'}$, $C_{2'}$	-0	-32
ν_1	$O_{4'}$, $C_{1'}$,.$C_{2'}$, $C_{3'}$	-27	3
ν_2	$C_{1'}$, $C_{2'}$, $C_{3'}$, $C_{4'}$	41	24
ν_3	$C_{2'}$, $C_{3'}$, $C_{4'}$, $O_{4'}$	-42	-41
ν_4	$C_{3'}$, $C_{4'}$, $O_{4'}$, $C_{1'}$	28	48
P[d]	Ribose conformation	23 (C3′-*endo*)	66 (C4′-*exo*)[b]

[a]The *cis* conformation for successive bonds is given a zero value and positive and negative values are right and left-handed rotations, respectively.
[b]Nomenclature according to Arnott & Huckins (1969).
[c]Atom nomenclature as defined in Fig. 1(c).
[d]Pseudorotation angle (*P*) is calculated (Altona & Sundaralingam, 1972) from the endocyclic sugar torsion angles as $\tan P = [(\nu_4 + \nu_1) - (\nu_3 + \nu_0)]/[2 \times \nu_2 \times (\sin(36°) + \sin(72°))]$.

Protein contacts of the external AMPPNP molecule

The interactions between the external nucleotide and the protein are mediated by the γ-imidophosphate (Fig. 2b): the bridging N atom (a hydrogen donor in this case) and the γ-phosphoryl O3G atom form hydrogen bonds with the Asn56 γ-amidic O and N atoms, respectively; the O3G atom, which is predominantly protonated at pH 4.6 (pK_a= 7.7; Yount *et al.*, 1971) is the donor in a hydrogen bond with the O atom of Gly54; the deprotonated O2G is a hydrogen acceptor and the countercharge for the N^ε atom of Lys53. The neutralization by the latter interaction of the negative charge on O2G should be critical to prevent repulsion with the negative C-end of the dipole at helix *B*.

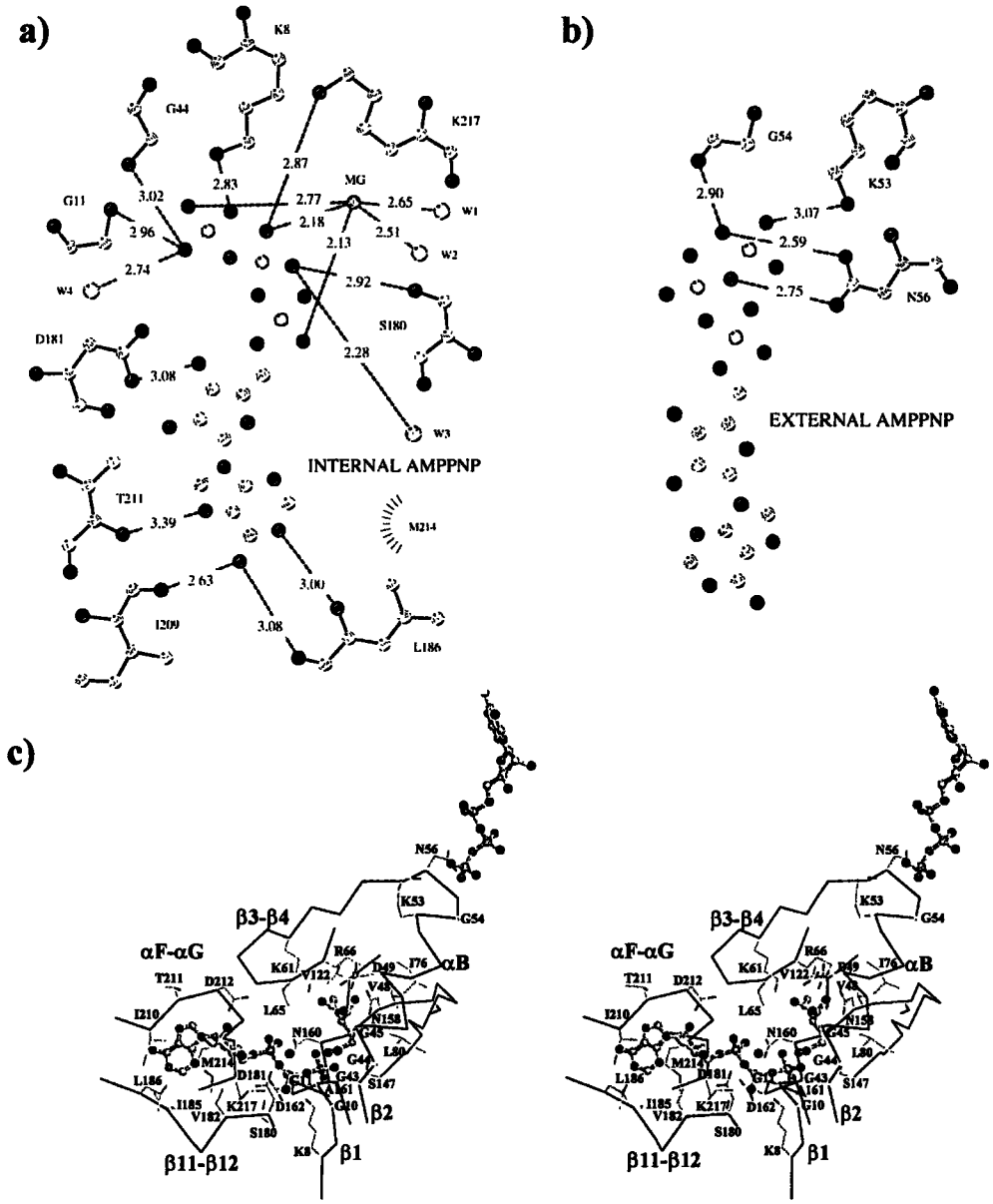


Figure 2. (a) and (b) Schematic plots drawn with LIGPLOT (Wallace *et al.*, 1995) of the interactions between the protein and the internal or external AMPPNP. Single-letter amino-acid code is used for protein residues. W denotes water molecules. MG, denotes Mg^{2+}. Distances are in Å. (c) Stereoview of the C^{α} trace of the binding sites. AMPPNP and NAG are shown in ball-and-stick representation, and colored. Amino-acid side chains are in color in a thinner trace.

DISCUSSION

The extended configuration, orientation and localization of the external nucleotide relative to the active center and the possibility of defining a direct path of approach without visible hindrances, along the groove between helix *αB* and strand *β3* (Fig. 2c), might suggest that the external nucleotide is waiting to occupy the active center. However, the active center and the

peripheral site exhibit strong differential selectivity for the Mg-complexed and Mg-free forms of the nucleotide, and there are important differences in conformation between the nucleotides bound at the two sites. Further, two of the four interactions between the peripheral nucleotide and the protein would not take place with ATP at neutral pH, since the hydrogen bonds involving the $\beta\gamma$ bridging N atom and the protonated O3G atom cannot be formed in the latter case, and thus ATP may simply not bind peripherally at neutral pH. These facts cast strong doubts about the possibility of the external nucleotide being a precursor of the active-site nucleotide. In any case, the mainly free crystalline external AMPPNP offers a rare glimpse of this nucleotide in the solution, as it approaches the enzyme.

CONCLUSIONS

In the crystal structure of NAGK at 1.5 Å resolution, a large volume of electron density has been interpreted as an extended AMPPNP molecule presenting two alternative dispositions around the two fold axis, with about half occupancy each. Unlike most protein-bound ATP molecules (Vetter & Wittinghofer, 1999), this AMPPNP is not metal-complexed and is mostly external to the protein, interacting with the latter only through its γ-imidophosphate. This external AMPPNP, although not appearing to be involved in the enzyme reaction, provides a glimpse of how free nucleotides can approach protein surfaces.

We thank the ESRF, Grenoble for the time at beamline ID14-1. We thank R. G. Yount for advice concerning AMPPNP properties. This work was supported by grants Rayos X of Fundación Ramón Areces and GV01-259 from the Generalitat Valenciana. F.G-O. and S.R-M. were fellows of the Fundación Ramón Areces and of the Generalitat Valenciana, respectively.

REFERENCES

Altona, C. & Sundaralingam, M. (1972). *J. Am. Chem. Soc.* **94**, 8205-8212

Arnott, S. & Hukins, D.W.L. (1969). *Nature* **224**, 886-888.

Cunin, R., Glansdorff, N., Piérard, A. & Stalon, V. (1986). *Microbiol. Rev.* **50**, 314-352.

Esnouf, R.M. (1999). *Acta Crystallogr. D* **55**, 938-940.

Gil, F., Ramón-Maiques, S., Marina, A., Fita, I. & Rubio, V. (1999). *Acta Crystallogr. D* **55**, 1350-1352.

Ghosh, M., Meerts, I.A., Cook, A., Bergman, A., Brouwer, A. & Johnson L.N. (2000). *Acta Crystallogr. D* **56,** 1085-1095.

Jones, T.A., Zou, J., Cowan, S. & Kjeldgaard, M. (1991). *Acta Crystallogr. A* **47,** 110-119.

Kraulis, P.J. (1991). *J. Appl. Crystallogr.***24,** 946-950.

Merritt, E.A. & Murphy, M.E.P. (1994). *Acta Crystallogr. D* **50,** 869-873.

Murshudov, G.N., Vagin, A.A. & Dodson, E.J. (1997). *Acta Crystallogr. D* **53,** 240-255.

O'Sullivan, W.J. & Smithers, G.W. (1979). *Methods Enzymol.* **63,** 294-336.

Phillippsen, A. (1998). http//www.bioz.unibas.ch/~xray/dino

Ramón-Maiques, S., Marina, A., Gil-Ortiz, F., Fita, I. & Rubio, V. (2002). *Structure* **10,** 329-342.

Saenger, W. (1984). *Principles of nucleic acid structure.* Springer-Verlag, New York.

Vetter, I.R. & Wittinghofer, A. (1999). *Quart. Rev. Biophysics* **32,** 1-56.

Wallace, A.C., Laskowski, R.A. & Thorton, J.M. (1995). *Protein Eng.* **8,** 127-134.

Yount, R.G., Babcock, D., Ballantyne, W. & Ojala, D. (1971). *Biochemistry* **10,** 2484-2489.

Capítulo 4

The course of phosphorus in the reaction of N-acetyl-L-glutamate kinase, determined from the structures of crystalline complexes, including a complex with an AlF_4^- transition state mimic

Trabajo aceptado en

Journal of Molecular Biology

The course of phosphorus in the reaction of N-acetyl-L-glutamate kinase, determined from the structures of crystalline complexes, including a complex with an AlF_4^- transition state mimic

Fernando Gil-Ortiz[1,3], Santiago Ramón-Maiques[1,3], Ignacio Fita[2] and Vicente Rubio[1]

Affiliations:

[1]Department of Genomics and Proteomics, Instituto de Biomedicina de Valencia, Consejo Superior de Investigaciones Científicas (IBV-CSIC), C/ Jaime Roig 11, 46010-Valencia, Spain.

[2]Instituto de Biología Molecular de Barcelona (IBMB-CSIC), C/ Jordi Girona 18-21, 08034-Barcelona, Spain.

[3] These authors have contributed equally to this work.

Running title: The course of phosphorus in the reaction of NAGK

SUMMARY

N-Acetyl-L-glutamate kinase (NAGK), the structural paradigm of the enzymes of the amino acid kinase family, catalyzes the phosphorylation of the γ-COO^- group of N-acetyl-L-glutamate (NAG) by ATP. We determine here the crystal structures of NAGK complexes with MgADP, NAG and the transition-state analog AlF_4^-; with MgADP and NAG; and with ADP and SO_4^{2-}. Comparison of these structures with that of the MgAMPPNP-NAG complex allows to delineate three successive steps during phosphoryl transfer: at the beginning, when the attacking and leaving O atoms and the P atom are imperfectly aligned and the distance between attacking O atom and the P atom is 2.8 Å; midway, at the bipyramidal intermediate, with nearly perfect alignment and a distance of 2.3 Å; and, when the transfer is completed. The transfer occurs in line and is strongly associative, with Lys8 and Lys217 stabilizing the transition state and the leaving group, respectively, and with Lys61, in contrast with an earlier proposal, not being involved. Three water molecules found in all the complexes play, together with Asp162 and the Mg, crucial structural roles. Two glycine-rich loops (β1-αA and β2-αB) are also very important, moving in the different complexes in concert with the ligands, to which they are hydrogen-bonded, either locking them in place for reaction or stabilizing the transition state. The active

site is too narrow to accommodate the substrates without compressing the reacting groups, and this compressive strain appears a crucial component of the catalytic mechanism of NAGK, and possibly of other enzymes of the amino acid kinase family such as carbamate kinase. Initial binding of the two substrates would require a different enzyme conformation with a wider active site, and the energy of substrate binding would be used to change the conformation of the active center, causing substrate strain towards the transition state.

Key words: Acetylglutamate kinase, amino acid kinase, phosphoryl group transfer, arginine metabolism, X-ray diffraction.

INTRODUCTION

The amino acid kinase enzyme family (Pfam[1] group PF00696; http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00696) groups on the basis of sequence homology the enzymes carbamate kinase, N-acetyl-L-glutamate kinase (NAGK), gamma-glutamyl kinase, and the major domain of the enzyme aspartokinase. All of these enzymes catalyze essentially the same type of reaction, the reversible transfer of the γ-phosphoryl group of ATP to a carboxylate or carbamate group, resulting in the synthesis of an acylphosphate anhydride. Two additional enzymes catalyzing other types of reactions, the bacterial UMP kinase and acetylglutamate synthetase, are also recognized as members of the family, although only the former of these enzymes uses ATP and makes an anhydride, in this case a phosphate-phosphate anhydride.

No enzyme of the amino acid kinase family had been characterized structurally before 1999, when we reported the structure at 2.8 Å resolution of *Enterococcus faecalis* carbamate kinase[2], showing that it presents with a new type of homodimeric open alpha/beta/alpha sandwich architecture. Essentially the same structure was demonstrated later on at 1.5 Å resolution in a thermophilic carbamate kinase from *Pyrococcus furiosus*[3] that had been characterized initially as a carbamoyl phosphate synthetase.[4-6] More recently, the structure of a third enzyme of this family, the NAGK from *Escherichia coli*, has been determined[7] also at 1.5 Å resolution, and has conformed essentially with the same topology and architecture as carbamate kinase. In fact, the combination of the structural data with sequence alignments has allowed to propose[7] on firm basis that all the members of the amino acid kinase family share the same polypeptide fold and are likely to bind their substrates and to catalyze their reactions similarly.

Structural information on how are the substrates bound in this enzyme family was first obtained with the *P. furiosus* carbamate kinase,[3] to which MgADP was bound extended along the C-edge of the main beta sheet, with the polyphosphate chain pointing in the direction of the dyadic axis of the dimer. More information was obtained when the structure of NAGK was determined,[7] since this structure corresponded to the ternary complex of the enzyme with the substrate N-acetyl-L-glutamate (NAG) and with the inert MgATP analog MgAMPPNP. In this complex, the ADP moiety of the nucleotide binds essentially as the MgADP in the *P. furiosus* carbamate kinase,[3] whereas the γ-phosphate points towards the γ-carboxylate of NAG in a way suitable for phosphoryl group transfer. The distance between the attacking O atom and the γ-P is 2.8 Å, and there is intervening electron density between these atoms, indicating that in this complex there is already some degree of bond formation between the O and P atoms. The structure of this complex has allowed to predict the localization of the site for the second product (or substrate, depending on the direction of the reversible reaction), carbamoylphosphate, in carbamate kinases,[8] substantiating the view that a sulphate ion that was found[2] binding in the structure of *E. faecalis* carbamate kinase represents the phosphate group of the carbamoyl phosphate. Thus, the combination of the data obtained with NAGK and with the two carbamate kinases of known structure appears to provide a fairly good representation of how substrates may bind and how phosphoryl group transfer may be initiated in any of the enzymes of the amino acid kinase family.

Nevertheless, much more information is needed to characterize phosphoryl group transfer in NAGK. For example, there are small but substantial chemical and structural differences between AMPPNP and ATP[9, 10] and it is uncertain how faithfully the structure of the NAGK ternary complex incorporating AMPPNP reflects the structure of the genuine complex with ATP. Furthermore, since the MgAMPPNP-NAG complex is an approximation to an early stage in the phosphoryl group transfer process, snapshots of more advanced stages are required if a complete description of the process has to be obtained. Such full characterization would be highly desirable, particularly since NAGK may represent a paradigm for the amino acid kinase family, given its similitudes with the other members of the family,[1,7] the large amount of structural information gathered on NAGK,[7] the fact that this enzyme presents the more basic architecture within the family both in terms of oligomeric organization[7] and of domain organization (a single consensus amino acid kinase domain[1]), and the existence, as with other members of the family,[11,12] of forms of NAGK that are feed-back inhibited by the final product (arginine[13,14]) of the pathway of which NAGK catalyzes an early step (although the NAGK from *E. coli* is not inhibited by arginine[15]).

the pathway of which NAGK catalyzes an early step (although the NAGK from *E. coli* is not inhibited by arginine[15]). Indeed, a thorough understanding of the mechanism of the NAGK reaction and catalysis may be of paramount importance to understand how feed-back inhibition is effected on the enzyme itself and, perhaps, on other members of the amino acid kinase family.

With the goal of characterizing fully the processes of phosphoryl group transfer and of catalysis in NAGK, we now determine using X-ray diffraction the structure of crystalline *E. coli* NAGK complexed with MgADP, NAG and an interposed planar AlF_4^- that approximates the phosphoryl group midway through the phosphoryl transfer process.[16,17] We also determine the structures of two additional crystalline NAGK complexes, one of them with MgADP and NAG but without an interposed AlF_4^-, and another with ADP and a sulphate but without NAG. The structural comparisons between these complexes and with the previously reported MgAMPPNP-NAG complex sheds light on the course of the phosphoryl group in the reaction, from ground state to products, and reveals movements of groups in the active center associated with the catalysis of the phosphoryl transfer. To our knowledge, this is the best structural characterization of the process of phosphoryl group transfer and catalysis among any of the enzymes of the EC group 2.7.2, which catalyze the synthesis of acylphosphates.[18-20] The present information sets the basic traits for understanding the reaction in the remaining enzymes of the amino acid kinase family.

A surprising finding in the previously reported NAGK structure[7,10] was the observation of an external, extended, Mg-free AMPPNP molecule, bouncing on to the surface of the NAGK molecule at a position that would suggest that it was waiting to occupy the active center as soon as the latter would be evacuated by the reaction products. The present structures do not reveal an external nucleotide, substantiating the view that this nucleotide would not be a normal element of the enzyme reaction, being found solely when using AMPPNP, and having no functional significance.

RESULTS

Overall fold and its changes

Three crystalline complexes of NAGK have been studied here (Table 1), one with MgADP, AlF_4^- and NAG, another with MgADP and NAG, and the third with ADP (in partial occupancy) and a bound sulphate but without Mg or NAG. The crystals of these three complexes and of the previously studied MgAMPPNP-NAG complex[7] are isomorphous, what has allowed

Table 1. Data collection and refinement statistics[a]

	MgADP- AlF_4^--NAG	MgADP-NAG	ADP-SO_4^{2-}
		Data Collection	
Beamline	BW7B (DESY)	ID14-2 (ESRF)	X11 (DESY)
Wavelength (Å)	0.8424	0.9340	0.9092
Resolution range (Å)	30.02-1.91	19.43-1.90	41.89-1.90
(Highest resolution shell)	(2.02-1.91)	(2.00-1.90)	(2.00-1.90)
Reflections (Total/ Unique)	89,855/18,025	154,986/18,636	100,167/18,284
I/σ_I	14.9 (5.3)	7.3 (1.3)	9.7 (2.1)
R_{sym}[b] (%)	3.7 (14.4)	6.8 (46.9)	7.5 (35.7)
Completeness (%)	99.0 (95.0)	99.9 (99.6)	99.9 (100.0)
		Refinement	
Resolution range (Å)	30.00-1.91	19.43-1.90	41.89-1.90
Reflections (work/test)	17,143/867	17,671/944	17,359/897
R factor[c] (work/test) (%)	18.58/22.92	20.00/22.95	19.48/22.11
Average B (Å^2)			
Protein	20.6	25.6	18.8
ADP	42.3	26.5	35.8
NAG	17.6	19.3	-
Water	31.8	32.4	27.1
Sulphate	-	-	32.9
Magnesium	55.3	23.6	-
Acetate	-	46.0	36.7
Number of atoms	2,160	2,090	2,077
Number of:			
Protein atoms	1,904	1,904	1,904
ADP molecules	1	1	1
NAG molecules	1	1	-
Water molecules	210	141	137
Sulphate ions	-	-	1
Magnesium ions	1	1	-
Acetate ions	-	1	1
Rmsd bond (Å)	0.005	0.005	0.005
Rmsd angle (°)	1.25	1.50	1.20
Ramachandran plot[d]			
Most favored (%)	94.5	94.5	94.5
Additional allowed (%)	5.0	5.0	5.5
Generously allowed (%)	0.5	0.5	0.0
Disallowed (%)	0.0	0.0	0.0

[a] Values in parenthesis are data for the highest resolution shell.

[b] $R_{sym} = \Sigma|I - \langle I \rangle|/\Sigma I$, where I is the observed intensity, and $\langle I \rangle$ is the average intensity of multiple observations of symmetry-related reflections.

[c] R factor $= \Sigma_{hkl}||F_{obs}|-|F_{calc}||/\Sigma_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are the observated and calculated structure factors, respectively.

[d] Calculated using PROCHECK.[31]

the use of the coordinates of the MgAMPPNP-NAG complex for phase determination, resulting in well defined electron density maps for the entire NAGK polypeptide (residues 1 to 258) and for the ligands, allowing us to construct 1.9 Å-resolution molecular models for the three complexes (Table 1).

The enzyme exhibits in the three complexes the same dimeric architecture (Fig. 1a) and overall polypeptide fold (Fig. 1b) as in the MgAMPPNP-NAG complex reported previously,[7] with good superposition of all the Cα atoms (Fig. 1b; rmsd values, 0.187, 0.383 and 0.476 Å for the superpositions of the MgADP-AlF_4^--NAG, MgADP-NAG and ADP-SO_4^{2-} complexes with the MgAMPPNP-NAG complex, respectively). The largest differences are observed between the MgADP-NAG and ADP-SO_4^{2-} complexes (rmsd for the superposition of all the Cα atoms of these two complexes, 0.643 Å), with the largest movements affecting the β1-αA loop (^{10}GGVLLD^{15}) and the connection between β2 and αB (^{43}GGG^{45}). These regions are hydrogen bonded to the nucleotide γ-phosphate and to the attacking O atom of the NAG γ-COO^-, strongly suggesting that they play key roles in the transfer of the phosphoryl group. In the MgADP-NAG and ADP-SO_4^{2-} complexes the β1-αA loop moves laterally (Fig. 1c) in opposite directions up to 2.26 Å (Cα displacement between extreme positions) and the N atom of Gly11 makes in the two complexes a hydrogen bond with the β-phosphate of ADP (see below, Fig. 5), whereas in the MgADP-AlF_4^--NAG and MgAMPPNP-NAG complexes the β1-αA loop occupies an position between those of the other two complexes (Fig. 1c) and the N atom of Gly11 is bonded, respectively, to a F of AlF_4^- and to the γ-phosphate of AMPPNP (see below, Fig. 5). The loop movement in the MgADP-NAG complex causes the reorientation and a displacement amounting to 3 Å (Fig. 2) of the interacting side chains of Leu13 (a β1-αA loop residue) and Trp237, with respect to their positions in the other complexes. The side-chain of Leu13 emerges from the N-lobe, and the side-chain of Trp237 emerges from the C-lobe of the enzyme (the N- and C- lobes host the NAG and ADP sites, respectively). Thus, these side-chains connect the two enzyme lobes. The β2-αB connection is reorganized in the ADP-SO_4^{2-} complex in the shape of a meander (Fig. 1c) in which Gly44 and Gly45 form the two opposite curves, with displacements of their respective Cα atoms of 1.69 and 2.84 Å, relative to the MgAMPPNP-NAG complex, thus accommodating the sulphate, and making a hydrogen bond with it through the N atom of Gly 45 (see below, Fig. 5). Helix B also changes somewhat its tilt, and the β3-β4 hairpin that follows this helix and that roofs the NAG site moves somewhat (maximal Cα displacement, 1.61 Å) in the ADP-SO_4^{2-} complex (Fig. 1b).

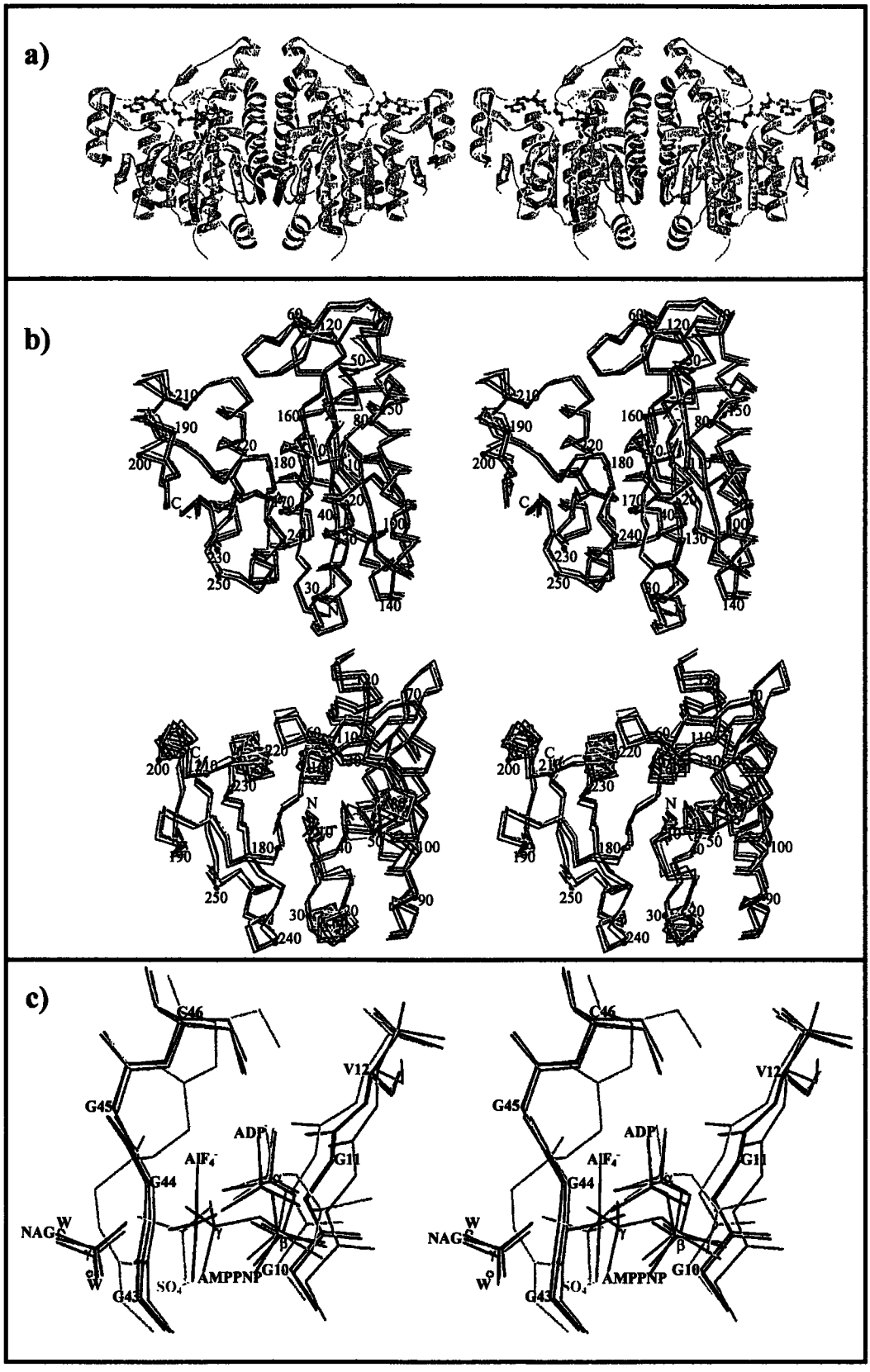

a)
b)
c)
G46
G45
G44
G43
V12
G11
G10
ADP
AlF4-
AMPPNP
SO4
NAG
W

Figure 1. Stereoviews of the *E. coli* NAGK fold (a,b) and detail of the active center (c). a) Ribbon representation of the complex with MgADP, AlF_4^- and NAG, to illustrate the dimeric architecture that is shared by all NAGK complexes. Ligands are represented in ball-and-stick, with the C, O, N, P, Al and F atoms colored in yellow, red, blue, green, grey and orange, respectively. b) Two perpendicular views of the Cα traces (with numbering every 10 residues) of the NAGK monomer in the complexes with MgAMPPNP-NAG (black), MgADP-AlF_4^--NAG (red), MgADP-NAG (green) and ADP-SO_4^{2-} (cyan). c) Detailed bond representation of the β1-αA and β2-αB connections of NAGK in the different complexes, and of the nearby groups of the ligands: the γ-COO^- of NAG (or, in the complex lacking NAG, two water molecules drawn as spheres), the polyphosphate chain of the nucleotide and the AlF_4^- and SO_4^{2-} ions. Color coding is as in panel b).

NAG binding

In the MgADP-NAG and MgADP-AlF_4^--NAG complexes NAG is bound in the same way as in the previously studied MgAMPPNP-NAG complex,[7] as reflected in the low rmsd values (0.073 to 0.131 Å) obtained by superposing the C, N and O atoms of the NAG bound in these complexes, and as also reflected in the nearly identical positions of the amino acid residues that conform the NAG pocket (Fig. 3a and 3b). No NAG is bound in the ADP-SO_4^{2-} complex (Fig. 3c), but its site is filled with one acetate (0.1 M acetate was present in the solutions) that replaces the acetamido group of NAG, and with five water molecules, three of which sit approximately at the positions of the two O atoms of the α-carboxylate and of the non-attacking O atom of the γ-carboxylate of NAG. The attacking O atom of the γ-COO^- of NAG is replaced by

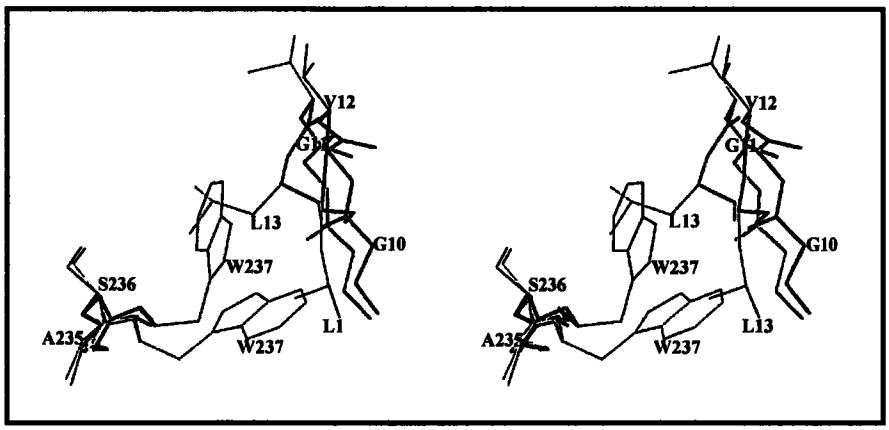


Figure 2. Stereoview bond representation of the reorientation of the side chains of Leu13 and Trp237 in the MgADP-NAG complex (green), compared with their orientation in the other complexes (illustrated for the MgAMPPNP-NAG complex, black). The main chain is shown in thick trace.

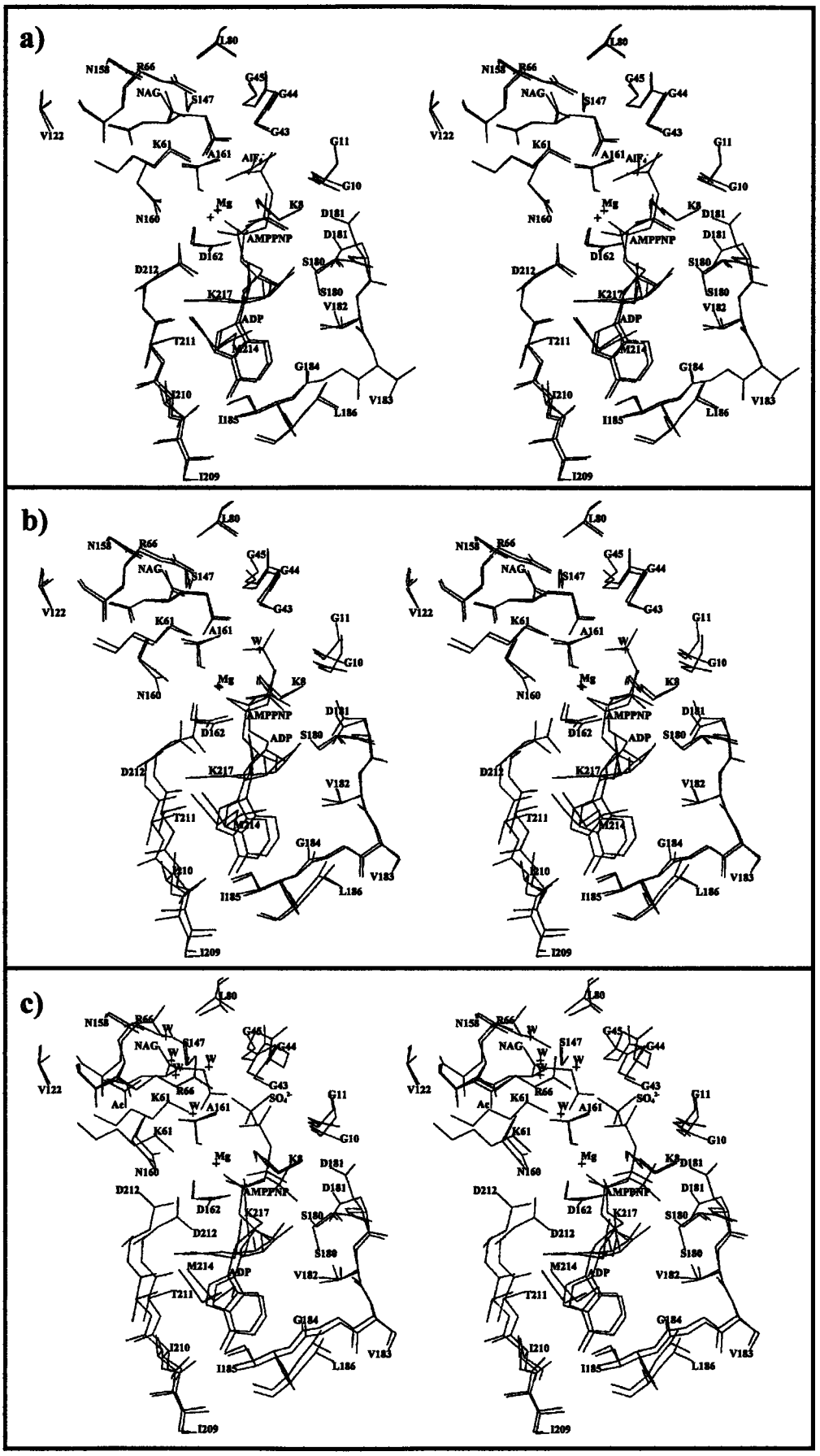

a)
b)
c)
L80
N158
R66
NAG
S147
G45
G44
G43
V122
K61
A161
G11
G10
N160
Mg
K9
D181
AMPPNP
D162
S180
D212
K217
V182
ADP
T211
M214
G184
I210
I185
L186
V183
I209
W
Ac
SO4 2-

Figure 3. Stereoview of the superimposition of the substrate-binding region and of the bound ligands in the MgAMPPNP-NAG complex (in black) and in each one of the other three complexes (colored): a) MgADP-AlF_4^--NAG; b) MgADP-NAG; c) ADP-SO_4^{2-}; in the latter complex an acetate (Ac) and five water molecules (crosses labeled W) occupy the NAG site and are shown. When present, the Mg ion is shown as a cross labeled Mg.

one O atom of the bound SO_4^{2-} (Fig. 3c). The NAG site is somewhat distorted relative to the NAG-containing complexes, because of the movements described above of the β2-αB connection, helix B and hairpin β3-β4. The side chain of Lys61, emerging from the β3-β4 hairpin, moves towards Asp212 across the interlobar divide, and the side chain of Arg66, also emerging from the β3-β4 hairpin, exhibits double conformation (the two conformations are illustrated in Fig. 3c).

ADP binding

The orientation of ADP is grossly the same in all the complexes (Fig. 3). The best superposition of both the nucleotide and the nucleotide site is observed for the MgAMPPNP-NAG and MgADP-AlF_4^--NAG complexes. The only substantial differences between these affect the prevailing conformer (represented in Fig. 3a) of the side chains of Ser180 and Asp181. In the MgADP-NAG and ADP-SO_4^{2-} complexes the positions of the AMP moiety, of the lining protein residues 202-212, and of the adenine-stacking Met214 are displaced, relative to the positions in the MgAMPPNP-NAG complex, by approximately 1 Å in opposite directions, respectively (Fig. 3b and 3c). In the ADP-SO_4^{2-} complex the AMP-encircling residues 184-186 also move in parallel with the AMP moiety. The positions of the β phosphate, α-β bridging O atom, and Mg are in the same place in all the complexes except the ADP-SO_4^{2-} complex (Fig. 3c). In contrast with the other complexes, the ADP-SO_4^{2-} complex lacks Mg, and the β-phosphate of ADP adopts in this complex a widely different orientation than in the other complexes. Correspondingly, the Ser180, Asp181 and Asp212 conformers fit the new orientation of the β-phosphate. Asp212 moves in the direction of Lys61, forming with it a salt bridge across the interlobar divide.

AlF_4^- mimics the phosphoryl group at the transition state.

Electron density of essentially planar appearance and approximately square shape was found between the ADP β-phosphate and the NAG γ-carboxylate in the complex prepared in the presence of $AlCl_3$ and NaF (Fig. 4a). The dimensions and the shape of this electron density are

conditions, AlF_4^- rather than AlF_3 is expected to be formed[17] from $AlCl_3$ and NaF. The center of this electron density, corresponding to the Al atom, falls nearly exactly in the straight line that joins the attacking O atom of NAG and the closest O atom of the ADP β-phosphate (O_{NAG}-Al-O_{ADP} angle, 176°), at distances of 2.33 and 1.88 Å from these atoms, respectively (Fig. 4a). Thus, the Al atom is more in line and more evenly centered between the attacking and leaving atoms than the γP atom of the MgAMPPNP-NAG complex (O_{NAG}-P_γ-N_{AMPPNP} angle, 155°; P_γ-O_{NAG} and N_{AMPPNP}-P_γ distances, 2.8 and 1.66 Å, respectively), and the distances between the attacking and leaving O atoms are similar in the two complexes (best estimates, 4.21 and 4.36 Å, respectively) (Figs. 4a and 4b). These characteristics approximate even more the MgADP-AlF_4^--NAG complex than the MgAMPPNP-NAG complex to the expected transition state for in-line phosphoryl group transfer in the NAGK reaction, rendering likely that the interactions of the ligands with the enzyme approximate largely those of the genuine transition state.

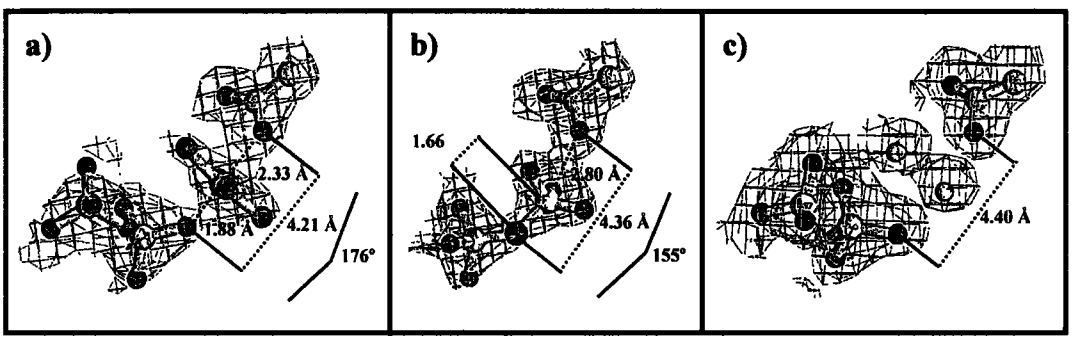


Figure 4. $2F_{obs}$-F_{calc} maps containing ball-and-stick models of the groups involved in phosphoryl group transfer in the complexes with MgADP-AlF_4^--NAG (a), MgAMPPNP-NAG (b) and MgADP-NAG (c). In the MgADP-NAG complex two water molecules occupying an intermediate position between ADP and NAG are also represented as cyan spheres. Panels b-d give also the indicated interatomic distances and angles.

In the MgAMPPNP-NAG complex[7] (Fig. 5a) the γ phosphate is fixed in place by the complete and partial bonding of the γP to the bridging N and attacking O_{NAG}, respectively, and by the interactions of the O atoms of the γ-phosphate. One O atom is hydrogen-bonded to the ζN of Lys8, another O is complexed to the Mg, and the third O atom makes hydrogen bonds with the N atoms of Gly11 and Gly44 and with a water molecule. In the MgADP-AlF_4^--NAG complex (Fig. 5b), the Al atom is octahedrally coordinated with the four equatorial F atoms and with the two apical attacking and leaving O atoms of NAG and ADP, and only one F atom is exposed to the free solvent. The other three F atoms play nearly exactly the same roles as the three O atoms of the AMPPNP γ-phosphate, with only minor differences that should generally result in similar

the free solvent. The other three F atoms play nearly exactly the same roles as the three O atoms of the AMPPNP γ-phosphate, with only minor differences that should generally result in similar or even stronger binding than for AMPPNP, given the somewhat shorter distances of some of the bonds (for example, 2.67 versus 2.83 Å to the ζN of Lys8, 2.53 versus 2.77 Å to the Mg, and 2.86 versus 2.96 to the N atom of Gly11).

All the interactions around the γ-phosphate found in the MgAMPPNP-NAG complex are preserved in the MgADP-AlF_4^--NAG complex (Fig. 5a and 5b). Thus, Mg, Asp162 and three fixed internal water molecules (shown in Fig. 5 as W1, W2 and W3) are key organizing elements of the active center, keeping in fixed appropriate places the positive groups of Lys8 and Lys217, which are likely to play crucial catalytic roles. In fact, the observation in the AlF_4^- complex (Fig. 5b) that a F atom is hydrogen-bonded to the central W2 molecule raises the possibility that an equatorial O atom of the transferred phosphoryl group is also hydrogen bonded to this water molecule, a possibility that was previously unrecognized, thus conferring to the W2 molecule an even more central role than anticipated already from its binding to the nucleotide-complexed Mg and to the NAG molecule.

Two water molecules replace the γ-phosphoryl group in the ADP-NAG complex.

In the MgADP-NAG complex the distance between the attacking O atom of NAG and the leaving O atom of ADP (4.4 Å, Fig. 4c) is essentially the same, and the disposition and interactions of the protein groups and of water molecules W1, W2 and W3 found in this region are also the same (Fig. 5c) as in the MgAMPPNP-NAG and MgADP-AlF_4^--NAG complexes. However, a new water molecule, W4, occupies a position approximately equivalent to that of the Mg-complexed O atom of the AMPPNP γ-phosphate, replacing it in the Mg coordination sphere and donating a hydrogen bond to the attacking O atom of NAG, and another water molecule (W5) not found in the other complexes approximately fits the position of the AMPPNP γ-phosphoryl O atom that is hydrogen bonded to the ζN atom of Lys8, replacing it in such hydrogen bond. W5 is also bonded to the nucleotide via a strong (distance, 2.55 Å) hydrogen bond with the leaving O atom of ADP. A third hydrogen bond links W5 to the attacking O atom of NAG. In this way, W4 and W5 fill in the MgADP-NAG complex most of the space of the missing γ-phosphate and either establish or mimic the connections of this phosphate. Only the γ-phosphoryl O atom that was hydrogen bonded to Gly11 and Gly44 is not replaced by a fixed water molecule in the MgADP-NAG complex. In this complex, Gly11 moves towards the ADP

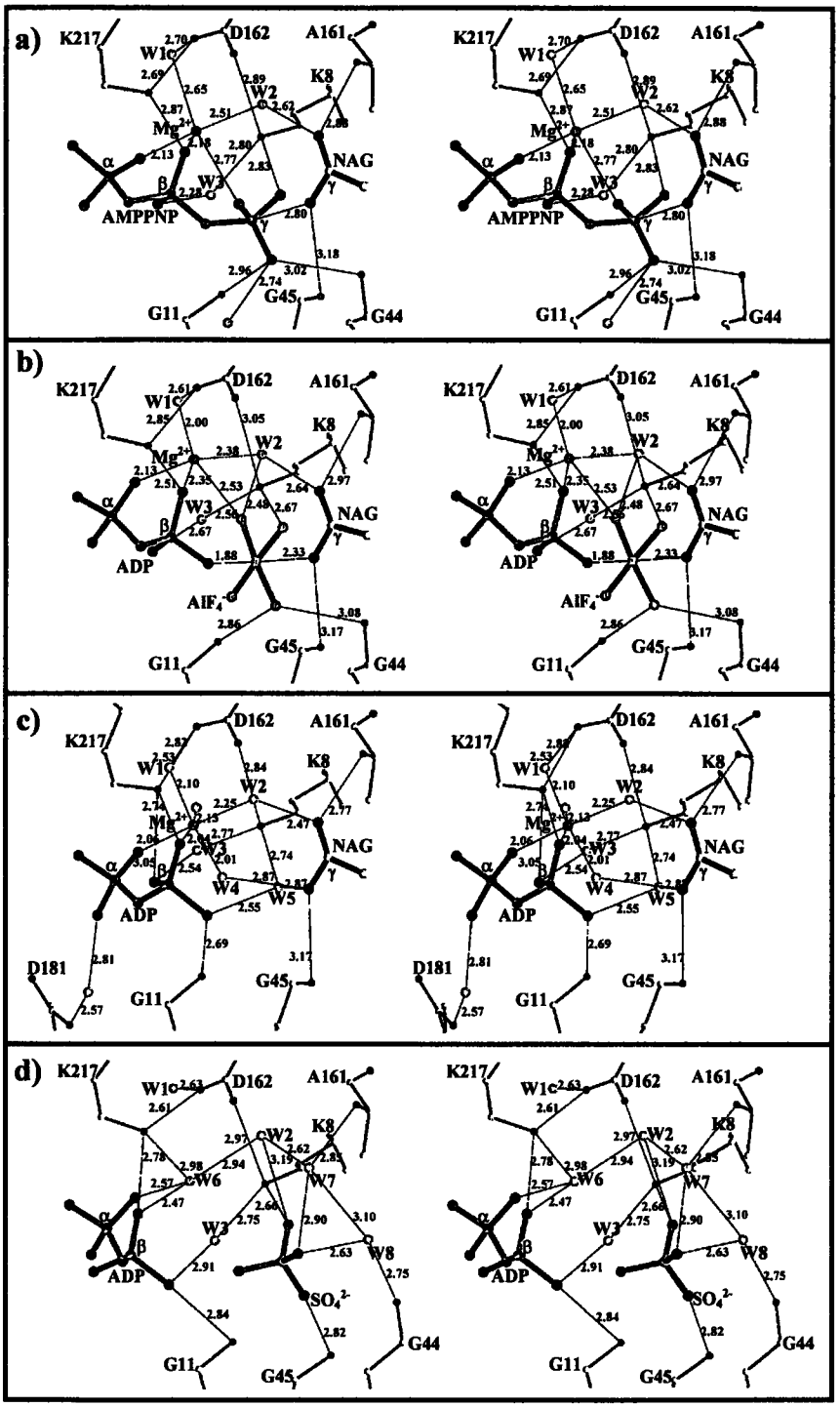


Figure 5. Stereoview ball-and-stick representation of the phosphoryl group transfer site in the NAGK complexes with MgAMPPNP-NAG (a), MgADP-AlF_4^--NAG (b), MgADP-NAG (c) and ADP-SO_4^{2-} (d). Of the ligands only the polyphosphate chain of the nucleotide and the γ-carboxylate of NAG are represented. Mg ion and water molecules are drawn as purple and cyan spheres, respectively. Nearby protein residues are shown in thinner trace. Hydrogen bonds and coordination bonds with Mg are shown as red broken lines, indicating the interatomic distances in Å. In panel a) the interatomic distance between the attacking O atom of NAG and the γP is represented with a blue broken line. AlF_4^- is shown in green.

β-phosphate and donates a strong hydrogen bond (distance, 2.69 Å) to the leaving O atom of this phosphate. A small shift in the position of the β-phosphate with respect to its position in AMPPNP allows the phosphate to make an additional hydrogen bond with the ζN atom of Lys217. Consequently, the β-phosphate within this complex is fixed by an extensive network of interactions that include the linkage of one non-bridging O atom to Gly11, of another non-bridging O to the ζN of Lys217 and to the organizing Mg, and of the third non-bridging O to the ζN atoms of Lys217 (direct hydrogen bond) and Lys8 (indirectly, through fixed water W3). An additional interaction that is not observed in the MgAMPPNP-NAG or MgADP-AlF_4^--NAG complexes links one non-bridging O atom of the ADP α-phosphate to the γ-COO^- of Asp181 via an intervening water molecule. These interactions might replicate those of ADP in the NAGK complex with the products.

Another differential characteristic of the MgADP-NAG complex with respect to other complexes is the perfection of the octahedral coordination sphere of the Mg and the presence of a water molecule occupying the sixth coordination position (Fig. 5c). In contrast, in the other complexes having bound Mg no ligand was observed at the sixth position, the geometry of coordination is distorted somewhat from perfectly octahedral, and the distances from the Mg to the coordinated groups are generally longer than in the MgADP-NAG complex (Fig. 5a and 5b), strongly suggesting that the coordination of Mg in these other complexes is strained.

A bound sulphate mimics the phosphate group of the product NAG-phosphate

In crystals soaked long periods in cryobuffer containing 0.25 M sulphate a large mass of electron density consistent with a sulphate ion was found between the ADP β-phosphate and the NAG site (which contained no bound NAG). In the structurally analogous enzyme carbamate kinase[2] a sulphate bound at a similar place was equated with the phosphate group of the product carbamoyl phosphate. The sulphate found in the ADP-SO_4^{2-} complex of NAGK (Figs. 1c, 3c and 5d) may also correspond to the phosphate group of the product NAG-phosphate (NAGP) since 1) it is closer to the NAG site than either the AMPPNP γ-phosphate or the AlF_4^- group; 2) one of its O atoms nearly occupies the site of the attacking O atom of NAG; and 3) the other three O atoms are disposed around the S atom as expected for the phosphoryl group after in-line transfer to NAG with inversion of configuration at phosphorus. Therefore, the ADP-SO_4^{2-} complex may shed light about the complex of the enzyme with the reaction products. For example, the β2-αB loop, which interacts with the sulphate, may accommodate in the genuine enzyme-products complex the

phosphate group of NAGP by adopting the meander shape and by hydrogen-bonding the phosphate via the N atom of Gly45, as observed with the sulphate (Figs. 1c and 5d). Similarly, in the genuine enzyme-products complex the β1-αB loop is likely to be hydrogen-bonded to the β-phosphate of ADP via the N atom of Gly11, since such a hydrogen-bond is observed in both the ADP-SO_4^{2-} and the MgADP-NAG complexes (Fig. 5c, 5d). However, the nucleotide product ADP is expected to be complexed to Mg, whereas in the ADP-SO_4^{2-} complex Mg is absent and the β-phosphate of ADP appears distorted. Consequently, the position of the β1-αA loop in the enzyme-products complex may be reflected better by the position of this loop in the MgADP-NAG complex than by that in the ADP-SO_4^{2-} complex (Fig. 1c).

As in the other complexes, the three fixed water molecules W1, W2 and W3, are also found in the ADP-SO_4^{2-} complex (Fig. 5d) and thus these waters appear constant elements of the active center. They exhibit the same interactions as in the other complexes, except with Mg, which is missing in the present complex, although W2 makes an equivalent connection with another water molecule, W6, that replaces the Mg and that also replicates the connections of the metal with the α and β phosphates of ADP, in addition to being hydrogen bound to Lys217. As already indicated, NAG is not present in this complex, but another water molecule, W7, replaces the non-attacking O atom of the NAG γ-COO^- and makes the same connections as this O atom, forming hydrogen bonds with W2 and with the N atom of Ala161. However, a hydrogen bond formed by W7 with the sulphate cannot replicate a similar bond in the genuine enzyme-product complex, since neither the carbonyl O atom that is replaced by W7 nor the bridging O atom of NAGP, corresponding to the O atom of the SO_4^{2-}, have a H atom to be shared. Finally, W7 is hydrogen bonded to another water molecule (W8) that bridges the sulphate ion with the N atom of Gly44, helping to fix the sulphate in place, but again it is uncertain whether an equivalent hydrogen bond exists in the genuine enzyme-product complex, since W8 would not be present in this complex, although perhaps Gly44 is hydrogen-bonded with the bridging O atom of the product NAGP.

None of the ADP-containing complexes has a peripherally bound nucleotide

In the structure of the NAGK complex with MgAMPPNP and NAG we observed a mass of extra electron density centered in the dyadic axis of the dimer, at the tip of helix B, outside the bulk of the protein.[10] We have shown this electron density to correspond to a Mg-free external AMPPNP molecule exhibiting double conformation and bouncing on the protein surface as if it

represented an initial stage in the approach of the nucleotide substrate to its final location in the enzyme active site. However, a detailed analysis of the interactions of this peripheral nucleotide with the enzyme raised important concerns about the functional significance of this binding, which our analysis[10] predicted that would be restricted to AMPPNP, not occurring with either the substrate ATP or the product ADP. We have been unable to obtain thus far complexes of NAGK with ATP, but the structures of the present complexes with ADP fulfill this prediction for the nucleotide product. As illustrated in Fig. 6 for the AlF_4^--containing complex, none of the three ADP complexes studied here exhibits any mass of electron density suspicious of concealing a peripherally bound nucleotide either were found in the MgAMPPNP-NAG complex or at other locations contacting the protein surface (not shown). The fact that the crystal group and packing of the present complexes is identical to the packing of the crystalline complex with MgAMPPNP-NAG also excludes that the presence of the external nucleotide were an essential requirement for crystal packing. Therefore, our results strengthen the view that the peripheral binding of a nucleotide to NAGK is not important either functionally or for the crystal structure, and may be considered a fortunate occurrence that has allowed the viewing of an extended nucleotide bouncing on a protein surface.

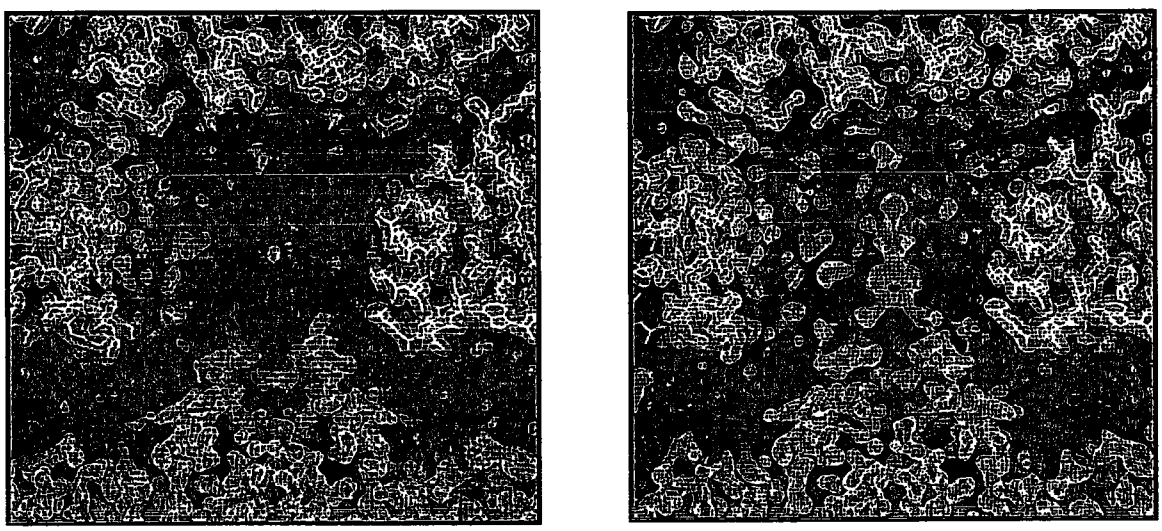


Figure 6. The $2F_{obs}$-F_{calc} maps of the ADP complexes (left panel, illustrated for the MgADP-AlF_4^--NAG complex) lack the electron density observed in the MgAMPPNP-NAG complex (right panel) that corresponds to an extended molecule of nucleotide bound peripherally. The two panels show the same region of the maps of the isomorphous crystals. This region is among four different enzyme molecules, and the density observed in the right panel contacts the top of an enzyme dimer (in the projection of Fig. 1a) at the intersubunit boundary.

DISCUSSION

The structure of the present complexes and of the MgAMPPNP-NAG complex[7] provides information about three steps along the process of phosphoryl group transfer catalyzed by NAGK. The MgAMPPNP-NAG complex reflects an early step, although not the earliest one, given the electron density and the short distance (2.8 Å, less than the sum of the van der Waals radii[7]) between the γ-P and the attacking O atom of NAG, evidencing already some degree of bond formation. The MgADP-AlF_4^--NAG complex mimics the transition state, approximately midway between substrates and products, given the even shorter distance (2.33 Å only) between the Al atom that corresponds to the transferred γP and the attacking O atom of NAG, and given the planarity of the AlF_4^- that is interposed between the attacking and leaving O atoms, with the Al atom in line with them, as expected for a mimic of the bipyramidal phosphorus transition state. The SO_4^{2-} in the ADP-SO_4^{2-} complex appears to represent the phosphate group of the product, whereas the MgADP in the MgADP-NAG complex is highly likely to correspond to the product nucleotide in the enzyme-products complex. Although the complexes studied only represent approximations rather than true snapshots of the actual phosphoryl group transfer process, the information that they convey is highly valuable to reconstruct the genuine process. Thus, AMPPNP differs slightly from ATP[9,10] in the angle and bond distances at the βP-γP bridge and it has a H atom linked to the bridging N atom, but it is unlikely that these differences should result in important distortions of the resulting complex, compared with the genuine MgATP-NAG complex. Similarly, the Al atom of the MgADP-AlF_4^--NAG complex is the center of a quadrangular bipyramid,[16] with F-Al-F angles of approximately 90°, whereas in the genuine transition state the P atom should center a triangular bipyramid with O-P-O angles for the equatorial oxygens of 120°.[21] However, as shown in the results section, three of the F atoms replicate the interactions of the non-bridging O atoms of the γ-phosphate of AMPPNP, and, therefore, the complex appears to approximate to a large extent the expected transition state. In any case, the informations provided by each of the complexes, combined, add up to give more strength to the conclusions about phosphoryl transfer in NAGK. For example, although in the MgAMPPNP-NAG complex the angle $N_{\beta\gamma}$-P_γ-O_{NAG} was only 155° rather than the 180° expected for in-line phosphoryl group transfer, there is no question that the transfer is in line, given the nearly perfect alignment of the attacking and leaving O atoms and the Al atom in the MgADP-AlF_4^--NAG complex. Similarly, the distance in the latter complex of 2.33 Å between the attacking O and the Al atom allows to calculate that the transfer has, in the terminology of Mildvan,[22] 10 %

associative character, confirming the previous estimate for the degree of associative character based on the structure of the MgAMPPNP-NAG complex.[7]

The observations made with the various complexes also provide very robust information concerning the roles of the different enzyme groups in the phosphoryl transfer process. The interactions of Lys217 with the β-phosphate of the nucleotide in all the complexes clearly support the involvement of this residue in assisting the development of a negative charge in the leaving group (the β-γ bridging O atom of ATP). Depending on the complex studied, Lys8 interacts with the γ-phosphate of AMPPNP, a F atom of AlF_4^- or the sulphate, indicating that the positive charge of Lys8 accompanies the γ-phosphoryl group throughout the entire phosphoryl transfer process, withdrawing electrons from the Pγ and shielding and stabilizing the equatorial negative charges in the pentavalent phosphorus transition state. In all the complexes studied the ζ-amino groups of Lys8 and Lys217 also interact with the β-COO^- of Asp162, and therefore this group plays a key role at all the stages of the reaction in the correct positioning of the catalytic groups of these lysine residues. The potential role proposed previously[7] for the partial positive charges at the N-ends of α-helices B and E in withdrawing negative charge from the transition state appears confirmed, since these helices are fixed in position throughout the entire phosphoryl transfer process near the transferred phosphoryl group or its mimics. In contrast, the previous proposal of a catalytic role for Lys61[7] is not supported by the present results, since although the position of the side chain of Lys61 varies in one of the complexes, in none of these Lys61 is close enough for effective negative charge dissipation at the transferred phosphoryl group. Three water molecules, W1, W2 and W3, are constant elements of the active site but, since the reaction does not demand acid-base catalysis,[7] the role of these water molecules cannot be to donate or to abstract protons. Therefore, these waters appear to play important roles in structuring the active site. W1 and W2 link the Mg ion with the side-chain of Asp162 and with the non-attacking O atom of the NAG γ-COO^-, respectively, whereas W3 links the ζ-amino group of catalytic Lys8 with the β-phosphate of the nucleotide. In this way, these water molecules connect either the site with the nucleotide or mutually the two substrates, helping to guarantee that the catalytic groups and the reacting substrates are favourably positioned for catalysis.

Another important observation derived from the study of the present complexes bears on the role of Gly11, Gly44 and Gly45. Gly11 and Gly44 are hydrogen bonded to the γ-phosphate of AMPPNP and to a F of AlF_4^-, and thus, they should withdraw electrons from the pentavalent phosphorus transition state. However, in the MgADP-NAG and ADP-SO_4^{2-} complexes Gly11 is

hydrogen-bonded to the ADP β-phosphate, and thus this hydrogen bond appears to help stabilize the negative charge that develops on the leaving β-phosphate, and Gly44 diverts away towards the NAGP site, and perhaps stabilizes this product, although this is not reflected in the present complexes. In the complexes having NAG Gly45 is hydrogen-bonded to the attacking O atom of this substrate, fixing it in place for attack, but in the sulphate-containing complex Gly45 makes a strong H bond with the sulphate, and therefore hydrogen-bonding of Gly45 is likely to help stabilize the product NAGP in the final stages of the transfer of the phosphoryl group.

A remarkable observation in the present studies is that, despite the differences observed between the various complexes, the active site has essentially the same dimensions in all cases and experiences little reshuffling during catalysis (Fig. 3), being too narrow to accommodate both MgATP and NAG without compressing the ATP γ-phosphate with the NAG γ-COO^-.[7] This is observed even with the MgAMPPNP-NAG complex, where the distance between the γ-P and the attacking O atom of NAG is less than the sum o the corresponding van der Waals radii.[7] These observations strongly suggest that the compressive strain imposed on the substrates plays a central role in the catalysis by this enzyme.[23] The complementarity between the NAGK site and the transition state would drive the reaction in the direction of the transition state, lowering the activation energy of the reaction.[24] Given the important structural similitudes between the active sites of NAGK and carbamate kinase[2,3,7] it is tempting to propose that strain may be a leading driving force in catalysis with all the enzymes of the amino acid kinase family that catalyze acylphosphate synthesis.

A problem posed by the narrowness of the active site to accommodate both substrates is to figure out how these substrates can be bound into the ternary complex, since if the site presented always with the shape observed here the binding of one substrate would be incompatible with the binding of the other substrate. It appears logical to propose that, in order to allow binding of both substrates, the substrate-free enzyme exhibits a more ample active center fitting both substrates, and that as the substrates bind a part of the binding energy is used to trigger a conformational change in the active center that gives the site its catalytically competent conformation causing substrate compression.[25] This view appears consistent with the observation of high K_m values of NAGK for its two substrates.[26] In any case, the fact that in the ADP-SO_4^{2-} complex, which has no NAG, the active center has approximately the same size and shape as in the other complexes (Fig. 3d) excludes the possibility that the binding of NAG were the trigger to induce the proposed change of the site to a narrow conformation. Thus, it appears more likely

that the energy for the conformational change is provided by the binding of the nucleotide. Determination of the 3-D structure of NAGK without the substrates or with the substrate NAG alone will be crucial to provide an experimental answer to this question.

The lack of NAG from the ADP-SO_4^{2-} complex is puzzling, and again demands the existence of different conformational forms of the enzyme that are presently unrecognized. The ADP-SO_4^{2-} complex was obtained when crystals prepared in the presence of MgADP and NAG were soaked at ambient temperature for a relatively long period (approximately 10 min) in a cryosolution rich in sulphate and devoid of any other ligands. The lack of Mg in this complex can be explained relatively easily since the metal can be accessed from the outside and may have been chelated out by the sulphate present in the cryosolution. After all, Mg-sulphate complexes are only ~10-fold less strong than MgADP complexes.[27] The different orientation of the β-phosphate of ADP in this complex, compared with the orientation in the MgADP-NAG complex, reflects the importance of the Mg as an organizing element. More difficult to visualize is the absence of NAG from this complex. NAG must have been a component of the crystalline complex before soaking the crystal in the cryobuffer, and the structure of the complex fails to reveal an obvious escape route for NAG from its binding site, since the NAG pocket remains closed and its opening is essentially blocked by the bound sulphate, which exhibits complete occupation. A plausible but presently unproven possibility is that the NAG site can exist in open and closed conformations, with the closed form predominating largely, but nevertheless, with the open conformation being sufficiently frequent, at 22°C, to allow the escape of the bound NAG over a period of 10 min. Overall, the lack of NAG from this complex and the narrowness of the site are important arguments in favor of the existence of important changes in the conformation of the active site, stressing the need for additional structural studies to define these changes.

MATERIALS AND METHODS

Crystallization and data collection

E. coli NAGK production from plasmid pNAGK24, and the purification of the enzyme and its crystallization in hanging drops by vapor diffusion at 294 K have been reported already.[15] Three types of crystalline complexes have been studied: with MgADP, NAG and AlF_4^-; with MgADP and NAG; and with ADP and SO_4^{2-}. The crystallization drops were prepared by mixing 1.5 µl of reservoir solution and 1.5 µl of a 10 mg/ml solution of NAGK in 6 mM of either sodium phosphate pH 7.0 (sulphate-containing complex) or Tris-HCl pH 7 (other two complexes), and

0.6 mM dithioerythritol, 20 mM KCl, 30 mM $MgCl_2$, 6 mM ADP, 24 mM NAG and, when preparing AlF_4^- complexes, 80 mM NaF and 10 mM $AlCl_3$. Crystals of rhomboidal habit and 0.3 mm maximal dimension developed within 1-2 weeks. The best crystals were obtained when the reservoir buffer contained 27-29 % polyethylene glycol monomethyl ether 2 K (from Hampton), sodium acetate 0.1 M pH 4.6 and either ammonium citrate 0.25-0.4 M (AlF_4^- crystals), or ammonium sulphate 0.15-0.3 M. Crystals were exposed briefly (<1 min) before diffraction to harvesting cryobuffers consisting in reservoir buffer enriched with 5-10 % polyethylene glycol monomethyl ether 2 K and supplemented with 5-15 % ethylene glycol. Only in the case of the sulphate-containing complex the crystal was passed through a graded series of solutions of increasing cryoprotectants concentrations having 0.24 M ammonium sulphate but no ADP or NAG, over a period totalizing approximately 10 min. Data were collected at 100 K (Oxford cryosystem) either at beamline BW7B (AlF_4^-) or X11 (sulphate complex) of the DESY synchrotron (Hamburg) or, for the complex without AlF_4^- or sulphate, at beamline ID14-2 of ESRF, Grenoble. Datasets at 1.9 Å resolution were obtained for all three types of crystal and were processed using MOSFLM[28], SCALA[28] and TRUNCATE[28] (Table 1). Crystals belong to space group $C222_1$ and present the same cell dimensions a=59 Å, b= 72 Å and c=107 Å, with one monomer in the asymmetric unit. These parameters are essentially the same found for the crystal complex of NAGK with MgAMPPNP and NAG.[7,15] The estimated solvent content of the crystals is 41 %, and the volume/mass ratio is 2.1 $Å^3.D^{-1}$.

Structure solution and refinement

The coordinates of NAGK with MgAMPPNP and NAG (PDB accesion code: 1gs5), minus ligands and minus water molecules were used for phase determination. Rigid body refinement in steps with increased resolution was followed by automatic refinement, using programs PROTIN[28] and REFMAC,[28] alternating with graphic manual model adjustment sessions with program O.[29] B-factor restrains were gradually released as refinement progressed. All the diffraction data were used throughout the refinement process except the 5 % randomly selected data for calculating R_{free}. At the middle of the refinement process energy minimization and individual B-factor refinement were applied using program CNS.[30] Solvent positions were automatically assigned using the option WATERPICK of CNS. The final models for the MgADP-AlF_4^--NAG and MgADP-NAG complexes (R value, 18.58 % and 20.00 % respectively; R_{free}, 22.92 % and 22.95 % respectively) correspond to one polypeptide chain (residues 1 to 258), one bound NAG and MgADP, 210 and 141 water molecules respectively, one AlF_4^- molecule in the

AlF_4^- complex and one acetate ion in the MgADP-NAG complex. The model of the ADP-SO_4^{2-} complex (R value, 19.48 %; R_{free}, 22.11 %) contains one sulphate ion, a Mg-free ADP with partial occupancy, one acetate ion and 137 water molecules. Structure analysis with PROCHECK[31] indicated good stereochemistry for all three models (Table 1).

Figures were drawn using programs MOLSCRIPT,[32] BOBSCRIPT,[33] RASTER3D[34] and DINO (Philippsen, A. (1998) *Dino, a visualization system for structural data.* http://cobra.mih.unibas.ch/dino).

Abbreviations used: NAGK, N-acetyl-L-glutamate kinase; NAG, N-acetyl-L-glutamate; NAGP, N-acetyl-L-glutamyl phosphate; rmsd, root mean square deviation.

ACKNOWLEDGEMENTS

Data collection was carried out at DESY (EMBL Hamburg Outstation) and ESRF (EMBL Grenoble) under the auspices of the EU. This work was supported by grants Rayos X of Fundación Ramón Areces and BMC2001-2182 of the Spanish Ministry of Science and Technology (MCYT). F. Gil-Ortiz was a predoctoral fellow of the Fundación Ramón Areces, and S. Ramón-Maiques a postgraduate I3P fellow of the Consejo Superior de Investigaciones Científicas (CSIC) and Bruker Española S.A.

REFERENCES

1. Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S. R., Griffiths-Jones, S., Howe, K. L., Marshall, M. & Sonnhammer, E. L. (2002). *Nucleic Acids Res.* **30**, 276-280.
2. Marina, A., Alzari, P. M., Bravo, J., Uriarte, M., Barcelona, B., Fita, I., & Rubio, V. (1999). Carbamate kinase: new structural machinery for making carbamoyl phosphate, the common precursor of pyrimidines and arginine. *Protein Sci.* **8**, 934-940.
3. Ramón-Maiques, S., Marina, A., Uriarte, M., Fita, I. & Rubio, V. (2000). The 1.5 Å resolution crystal structure of the carbamate kinase-like carbamoyl phosphate synthetase from the hyperthermophilic archaeon *Pyrococcus furiosus*, bound to ADP, confirms that this thermostable enzyme is a carbamate kinase, and provides insight into substrate binding and stability in carbamate kinases. *J. Mol. Biol.* **299**, 463-476.
4. Durbecq, V., Legrain, C., Roovers, M., Piérard, A. & Glansdorff, N. (1997). The carbamate kinase-like carbamoyl phosphate synthetase of the hyperthermophilic archaeon *Pyrococcus*

furiosus, a missing link in the evolution of carbamoyl phosphate biosynthesis. *Proc. Natl. Acad. Sci. USA,* **94**, 12803-12808.

5. Uriarte, M., Marina, A., Ramón-Maiques, S., Fita, I. & Rubio, V. (1999). The carbamoyl phosphate synthetase of *Pyrococcus furiosus* is enzymologically and structurally a carbamate kinase. *J. Biol. Chem.* **274**, 16295-16303.
6. Uriarte, M., Marina, A., Ramón-Maiques, S., Rubio, V., Durbecq, V., Legrain, C., & Glansdorff, N. (2001). Carbamoyl phosphate synthesis: carbamate kinase from *Pyrococcus furiosus*. *Methods Enzymol.* **331**, 236-247.
7. Ramón-Maiques, S., Marina, A., Gil-Ortiz, F., Fita, I. & Rubio, V. (2002). Structure of acetylglutamate kinase, a key enzyme for arginine biosynthesis and a prototype for the amino acid kinase enzyme family, during catalysis. *Structure,* **10**, 329-342.
8. Ramón-Maiques, S., Britton, H. G. & Rubio, V. (2002). Molecular physiology of phosphoryl group transfer from carbamoyl phosphate by a hyperthermophilic enzyme at low temperature. *Biochemistry,* **41**, 3916-3924.
9. Yount, R. G., Babcock, D., Ballantyne, W. & Ojala, D. (1971). Adenylyl imidodiphosphate, an adenosine triphosphate analog containing a P-N-P linkage. *Biochemistry,* **10**, 2484-2489.
10. Gil-Ortiz, F., Fita, I., Ramón-Maiques, S., Marina, A. & Rubio, V. (2002). A crystallographic glimpse of a nucleotide triphosphate (AMPPNP) bound to a protein surface: external and internal AMPPNP molecules in crystalline N-acetyl-L-glutamate kinase. *Acta Crystallog. sect. D,* **58**, 1892-1895.
11. Truffa-Bachi, P. (1973). Microbial aspartokinases. In *The Enzymes* (Boyer, P., ed.) 3rd edit., vol. 8, pp. 509-553, Academic Press, New York.
12. Baich, A. (1969). Proline synthesis in *Escherichia coli*. A proline-inhibitable glutamic acid kinase. *Biochim. Biophys. Acta,* **192**, 462-467.
13. Haas, D. & Leisinger, T. (1975). N-acetylglutamate 5-phosphotransferase of *Pseudomonas aeruginosa*. Catalytic and regulatory properties. *Eur. J. Biochem.* **52**, 377-383.
14. Fernández-Murga M. L., Ramón-Maiques, S., Gil-Ortiz, F., Fita, I. & Rubio, V. (2002). Towards structural understanding of feedback control of arginine biosynthesis: cloning and expression of the gene for the arginine-inhibited N-acetyl-L-glutamate kinase from *Pseudomonas aeruginosa*, purification and crystallization of the recombinant enzyme and preliminary X-ray studies. *Acta Crystallog. sect. D,* **58**, 1045-1047.

15. Gil, F., Ramón-Maiques, S., Marina, A., Fita, I. & Rubio, V. (1999). N-Acetyl-L-glutamate kinase from *Escherichia coli*: cloning of the gene, purification and crystallization of the recombinant enzyme and preliminary X-ray analysis of the free and ligand-bound forms. *Acta Crystallog. sect. D,* **55**, 1350-1352.

16. Tesmer, J. J., Berman, D. M., Gilman, A. G. & Sprang, S. R. (1997). Structure of RGS4 bound to AlF_4^--activated $G_{i\alpha1}$: stabilization of the transition state for GTP hydrolysis. *Cell*, **89**, 251-61.

17. Schlichting, I. & Reinstein, J. (1999). pH influences fluoride coordination number of the AlF_x phosphoryl transfer transition state analog. *Nature Struc. Biol.* **6**, 721-723.

18. Auerbach, G., Huber, R., Grattinger, M., Zaiss, K., Schurig, H., Jaenicke, R. & Jacob, U. (1997). Closed structure of phosphoglycerate kinase from *Thermotoga maritima* reveals the catalytic mechanism and determinants of thermal stability. *Structure,* **5**, 1475-1483.

19. Buss, K. A., Cooper, D. R., Ingram-Smith C., Ferry, J. G., Sanders, D. A. & Hasson, M. S. (2001). Urkinase: structure of acetate kinase, a member of the ASKHA superfamily of phosphotransferases. *J. Bacteriol.* **183**, 680-686.

20. Miles, R. D., Iyer, P. P. & Ferry, J. G. (2001). Site-directed mutational analysis of active site residues in the acetate kinase from *Methanosarcina thermophila. J. Biol. Chem.* **276**, 45059-45064.

21 Knowles, J. R (1980). Enzyme-catalyzed phosphoryl transfer reactions. *Annu. Rev. Biochem.* **49**, 877-919.

22. Mildvan, A. S. (1997). Mechanisms of signaling and related enzymes. *Proteins: Struc. Func. Gen.* **29**, 401-416.

23. Fersht, A. (1999). *Structure and mechanism in protein science. A guide to enzyme catalysis and protein folding.* W. H. Freeman & Co., New York.

24. Pauling, L. (1946). Molecular architecture and biological reactions. *Chem. Eng. News.* **24**, 1375-1377.

25. Jencks, W. P. (1975). Binding energy, especificity, and enzimic catalysis: The circe effect. *Adv. Enzymol. Relat. Areas Mol. Biol.* **43**, 219-410.

26. Vogel, H. J., & McLellan, W. L. (1970). N-Acetyl-γ-glutamokinase (*Escherichia coli*). *Methods Enzymol.* **17A**, 251-255.

27. Dawson, R. M. C., Elliott, D. C., Elliott, W. H. & Jones, K. M. (1986). *Data for biochemical research*, Oxford University Press Inc., New York.

28. Collaborative Computational Project No. 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D,* **50**, 760-763.

29. Jones, T. A., Zou, J., Cowan, S. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A,* **47**, 110-119.

30. Brünger, A. T., et al. & Warren, G.L. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D,* **54**, 905-921.

31. Laskowsky, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283-291.

32. Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946-950.

33. Esnouf, R. M. (1999). Further additions to MolScript version 1.4, including reading and contouring of electron-density maps. *Acta Crystallog. sect. D,* ***55***, 938-940.

34. Merritt, E. A. & Murphy, M. E. P. (1994). Raster3D version 2.0. A program for photorealistic molecular graphics. *Acta Crystallog. sect. D,* **50**, 869-873.

Capítulo 5

Towards structural understanding of feed-back control of arginine biosynthesis: cloning and expression of the gene for the arginine-inhibited N-acetyl-L-glutamate kinase from *Pseudomonas aeruginosa*, purification and crystallization of the recombinant enzyme and preliminary X-ray studies.

Trabajo publicado en

Acta Crystallographica Section D **(2002) D58, Pags. 1045-1047.**

Towards structural understanding of feed-back control of arginine biosynthesis: cloning and expression of the gene for the arginine-inhibited N-acetyl-L-glutamate kinase from *Pseudomonas aeruginosa*, purification and crystallization of the recombinant enzyme and preliminary X-ray studies

M. Leonor Fernández-Murga[a], Santiago Ramón-Maiques[a], Fernando Gil-Ortiz[a], Ignacio Fita[b] and Vicente Rubio[a]

[a]Instituto de Biomedicina de Valencia, Consejo Superior de Investigaciones Científicas (IBV-CSIC), C/ Jaime Roig 11, 46010-Valencia, Spain, and [b]Instituto de Biología Molecular de Barcelona (IBMB-CSIC), C/ Jordi Girona 18-21, 08034-Barcelona, Spain.

ABSTRACT

N-acetyl-L-glutamate kinase (NAGK) catalyzes the second step in the pathway of arginine biosynthesis in microorganisms and plants. In many species it is the pathway-controlling enzyme and is subject to feed-back inhibition by arginine. The gene for the best characterized arginine-inhibitable NAGK, from *Pseudomonas aeruginosa*, has been cloned in a pET22 plasmid and overexpressed in *Escherichia coli*. The enzyme was purified in three steps to 95 % purity and was shown by cross-linking to form dimers. It was crystallized by the hanging drop vapor diffusion method at 277 K in the presence of ADP, Mg and N-acetyl-L-glutamate. The crystallization solution contained 0.1 M Na cacodilate pH 6.5, 150-170 mM magnesium acetate, and 13 % polyethylene-glycol 8000. Prismatic crystals of approximately 0.5 mm maximum dimension diffract to 2.75 Å resolution and belong to space group P1 (unit cell parameters a= 71.86 Å, b= 98.78 Å, c= 162.9 Å, α= 91.49°, β= 92.03°, γ= 107.56°). Packing density considerations agree with 6-18 NAGK monomers for 79-36 % solvent content in the asymmetric unit. Self-rotation function calculations confirm the space group and suggest the presence of 3-7 dimers in the unit cell.

Synopsis: N-Acetyl-L-glutamate kinase of *P. aeruginosa* catalyses the controlling step of arginine biosynthesis and is feed-back inhibited by arginine. The gene for this enzyme was cloned and expressed at high levels in *E. coli*. Crystals of the purified recombinant enzyme, of space group P1, diffracted to a resolution of 2.75 Å.

INTRODUCTION

Microorganisms and plants synthesize arginine from glutamate. Glutamate is first N-acetylated, and then it is γ-phosphorylated by N-acetyl-L-glutamate kinase (NAGK) to give γ-phosphoryl N-acetyl-L-glutamate, which is converted in two steps to N-acetyl-L-ornithine (Cunin *et al.*, 1986; Shargool *et al.*, 1988). After deacetylation, the ornithine is finally converted to arginine. There are two variants of this pathway (Cunin *et al.*, 1986). In one, typified by the pathway in *Escherichia coli*, glutamate is acetylated by acetyl-CoA and acetylornithine is deacylated hydrolytically. In the other, typified by the pathway in *Pseudomonas aeruginosa*, the acetyl group is recycled by reversible transacetylation from acetylornithine to glutamate. Both of these pathways are subject to feed-back inhibition by arginine. In the *E. coli* pathway, inhibition occurs at the initial acetylation of glutamate. In the *P. aeruginosa* pathway, phosphorylation of N-acetylglutamate (NAG) by NAGK is inhibited by arginine (Cunin *et al.*, 1986; Haas & Leisinger, 1975b).

In our laboratory we have recently determined the 3-D structure of *E. coli* NAGK bound to NAG and to the ATP analog AMPPNP, having thus determined the mechanism of catalysis by this enzyme (Ramón-Maiques *et al.*, 2002). However, since *E. coli* NAGK is not a controlling enzyme and is not inhibited by arginine (Cunin *et al.*, 1986), the mechanism of allosteric regulation of NAGK by arginine is an important issue that remains to be clarified. The determination of the regulatory mechanism may also be relevant for understanding the inhibition by arginine of NAG synthetase in those organisms that do not recycle the acetyl group (Leisinger & Haas, 1975; Cunin *et al.*, 1986). A protein module conferring sensitivity to arginine might be incorporated either to NAG synthetase or to NAGK, depending on the organism and the metabolic organization of the route. To understand in physical terms the mechanism of inhibition of NAGK by arginine, structural studies have been initiated with the NAGK from *P. aeruginosa*. This enzyme is well characterized biochemically and functionally and is strongly inhibited by arginine (Haas & Leisinger, 1975a and 1975b). The crystallization and initial X-ray diffraction studies of this enzyme, overexpressed in *E. coli* from the cloned gene, are reported here. The crystals pose an interesting and difficult crystallographic problem, given their triclinic character (P1 group) and the large number of enzyme monomers in the unit cell.

EXPERIMENTAL

P. aeruginosa PAO1 (donated by Dr. Dieter Haas, University of Lausanne), the strain originally used to purify NAGK (Haas & Leisinger, 1975a), was the source of genomic DNA. Using this DNA as template, the putative *P. aeruginosa* NAGK gene (*argB*, PA5323, Pseudomonas genome project, http://www.pseudomonas.com) was amplified by PCR utilizing a high fidelity proofreading thermostable DNA polymerase (Deep Vent, from New England Biolabs) and the direct and reverse primers, 5′TCGGAGCTCC*AT*ATGACCCTGAGTCGCGATGACG3′ and 5′GCGCGCCG*AA*GC*TT*GCGATCAGTGACGCTTGCGGTTGCTGATCA3, respectively. These primers (derived from nucleotides 5993778-5993811 and complementary sequence for 5994670-5994713 of the *P. aeruginosa* genome) encompass the beginning and the end of the ORF (the coding sequences are underlined) and short flanking genomic *P. aeruginosa* sequences. They include mutations (shown in italic) to introduce *Nde*I (direct primer) and *Hind*III (reverse primer) sites after the initiator and stop codons.

The amplified fragment, digested with *Nde*I and *Hind*III and ligated using T4 ligase into the same sites of plasmid pET-22b (Novagen), was used to transform *E. coli* DH5α cells (from Clontech), allowing the isolation of plasmid pNAGK-PA25, which carries in its insert the NAGK gene (as shown by DNA sequencing). To overexpress the gene, *E. coli* BL21 (DE3) cells (from Novagen) were transformed with pNAGK-PA25 and were grown at 310 K with aeration to a cell density of A_{600} = 0.6, in 1.5 l LB broth containing 0.1 mg ml^{-1} ampicillin. Then, 1 mM isopropyl-β-D-thiogalactoside (IPTG) was added, the incubation was continued for 3 h and the cells were collected by centrifugation. Subsequent steps were carried out at 277 K. The pellet, suspended in 15 ml per g cells of 0.1 M Na-phosphate pH 7.0, 0.2 mM dithioerythritol, was sonicated, and after centrifugation (30 min, 35,000 g) the supernatant was precipitated sequentially with $(NH_4)_2SO_4$ at 30 and 65 % saturation at 273 K. The 65 %-saturation precipitate was dissolved in 20 ml of buffer A (20 mM Na phosphate pH 8.0, 1 mM dithioerythritol) and dialysed against the same buffer before applying it to a 1×18 cm column of Q-Sepharose Fast Flow (Pharmacia Biotech) that had been equilibrated with buffer A. After a 100 ml wash with buffer A, a 400-ml linear gradient of 0-0.5 M NaCl in this buffer was applied. Enzyme-rich fractions (monitored by SDS-PAGE) were eluted at approximately 0.12 M NaCl. They were pooled, dialysed against 20 mM Na phosphate buffer pH 7.0, containing 20 mM $MgCl_2$ and 1 mM dithioerythritol, and were applied to a 2 x 20 cm column of Affigel

Blue (Bio-Rad) that had been equilibrated with the same solution. After washing with 180 ml of the solution, a 500-ml linear gradient of 0-1.0 M NaCl in this solution, devoid of $MgCl_2$, was applied to the column, and NAGK was eluted at approximately 0.75 M NaCl. The enzyme was concentrated by ultrafiltration to 5 mg. ml^{-1}, supplemented with 10 % glycerol, and stored at 253 K.

The sparse-matrix sampling vapor diffussion method (Jancarik & Kim, 1991) was used for crystallization tests carried out at 277 and 294 K in drops hanging in multiwell plates,

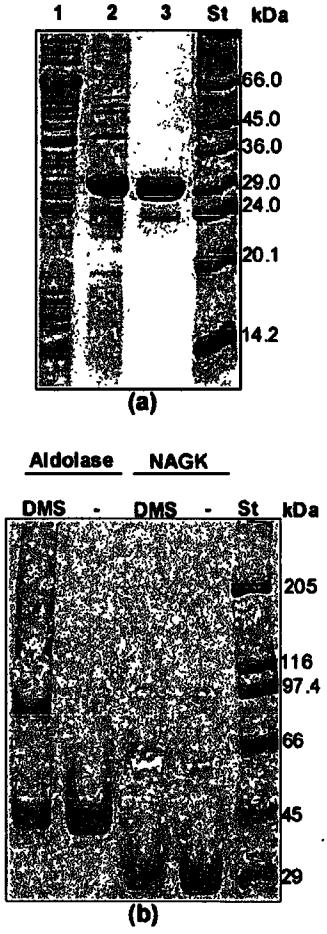


Figure 1. Expression, purification and cross-linking of NAGK. St, protein standards of the indicated masses. (a) SDS-PAGE of the sonicates of *E. coli* BL21 (DE3) cells transformed with either plasmid pET-22b (lane 1) or pNAGK-PA25 (lane2), or of the purified enzyme (lane 3). (b) Cross-linking with dimethylsuberimidate (DMS; from Pierce), using the method of Davies &Stark (1970), of 5 mg ml^{-1} NAGK or aldolase (used as a cross-linking control; from rabbit muscle; monomer mass 40 kDa). The minus sign indicates omission of DMS.

utilizing commercial kits (Crystal Screen I and II, from Hampton Research). The drops contained 1.5 µl of reservoir fluid and 1.5 µl of a 10 mg. ml^{-1} NAGK solution prepared by repeated centrifugal ultrafiltration (Microsep 10K, Pall Filtron) of the enzyme in 20 mM HEPES pH 7.5, 1 mM dithioerythritol, 10% glycerol and 0.02 % sodium azide. The best crystals were grown in the presence of 30 mM $MgCl_2$, 20 mM N-acetyl-L-glutamate and 10 mM ADP at 277 K in about two weeks using as reservoir fluid 0.1 M sodium cacodilate pH 6.5, 150-170 mM magnesium acetate and 13 % polyethylene-glycol 8000 (Hampton

Figure 2. Typical crystal of *P. aeruginosa* NAGK grown at 277 K in the presence 30 mM $MgCl_2$, 20 mM NAG and 10 mM ADP.

Research). Crystals were flash-cooled using an Oxford cryosystem and diffraction was carried out at 100 K using synchrotron radiation (ESRF, Grenoble; beamline ID14-4) and a Quantum ADSC Q4R CCD detector. The data set was processed and scaled with MOSFLM (Leslie, 1990) and SCALA (Evans, 1997). The self-rotation functions (Rossmann & Blow, 1962) were calculated with MOLREP (Vagin & Teplyakov, 1997).

Protein was determined by the method of Bradford (1976) using bovine serum albumin as standard. Densitometry of digitized gel images, carried out with the program Sigmagel (from Jandel Scientific) was used for quantification of proteins in Coomassie-stained electrophoretic gels. For all other experimental techniques, see Gil *et al.* (1999).

RESULTS AND DISCUSSION

SDS-PAGE (Fig. 1a) of *E. coli* BL21 (DE3) transformed with plasmid pNAGK-PA25 revealed a prominent band, representing 18 % (densitometric estimate) of the bacterial protein, which was not observed when the cells were transformed with pET 22b not carrying the insert. The electrophoretic mass estimate for this band (30 kDa) agrees, within experimental error, with a previous electrophoretic estimate of the mass of NAGK purified from *P. aeruginosa* (29 kDa, Haas & Leisinger 1975a) and with the sequence-deduced mass of the enzyme polypeptide (31.85 kDa). Furthermore, the extracts of the *E. coli* cells transformed with plasmid pNAGK-PA25 but not those of the cells transformed with pET22b exhibited important NAGK activity (4.01 μmol. min^{-1}. mg^{-1}) that was essentially abolished by the addition to the assay of 1 mM arginine, again as expected for *P. aeruginosa* NAGK (Haas & Leisinger, 1975b). In contrast, pure *E. coli* NAGK (Gil *et al.*, 1999) was not inhibited significantly even by 10 mM arginine.

A three-step purification protocol (see Experimental section) lasting 2-3 days yielded the 95 % pure enzyme (purity monitored by SDS-PAGE using densitometry; Fig. 1a, lane 3). The specific activity of the purified enzyme (88 ± 7 μmol.min^{-1}.mg^{-1}) was comparable, although somewhat higher, than that reported for the purest preparation from *P. aeruginosa* (68.8 μ mol.min^{-1}.mg^{-1}, Haas & Leisinger, 1975a). The enzyme activity was inhibited 13, 54 and 96 % by 0.1, 0.5 and 1 mM arginine, respectively. Automated Edman sequencing (performed by the sequencing service of this Institute) yielded the expected N-terminal sequence TLSRDDAAQVAKVLSEA, except for the lack of the initial methionine, which thus, is removed post-translationally. MALDI-TOF mass spectrometry (kindly performed by Dr. J.J. Calvete, from this Instituto, with the Voyager System 6221; from Applied Biosystems) of the pure protein yielded a mass of 31,711 Da, to be compared with the mass of 31,718 Da deduced from the gene sequence after removal of Met1. In sum, all the properties indicate that the recombinant enzyme corresponds to the genuine *P. aeruginosa* NAGK. In addition, similarly to *E. coli* NAGK (Gil *et al.*, 1999; Ramón-Maiques *et al.*, 2002), the enzyme from *P. aeruginosa* seems to be dimeric, since a band corresponding to the dimer was observed on SDS-PAGE (Fig. 1b), in addition to the band of the monomer, upon cross-linking of the enzyme with dimethylsuberimidate.

Prismatic crystals of approximately 0.5 mm maximal dimension were obtained in the presence of 30 mM $MgCl_2$, 20 mM NAG and 10 mM ADP at 277 K (Fig 2). Upon diffraction with synchrotron radiation, data were collected to 2.75 Å resolution (91.4 % completeness, R_{merge} 5.7). The space group was triclinic (P1), with unit cell parameters a= 71.86 Å, b= 98.78 Å, c= 162.9 Å, α= 91.49°, β= 92.03°, γ= 107.56°. Packing density considerations (Matthews, 1968) for a monomer mass of 31,714 Da conform with 6-18 monomers (V_m=5.79-1.93 Å^3.Da^{-1} for a solvent content range of 79-36 %). To cross-check the space group the self-rotation function (Rossmann & Blow, 1962) was calculated with MOLREP (Vagin & Teplyakov, 1997), yielding peaks that were always weaker than 25% of the origin peak, supporting the absence of crystallographic axes with rotational symmetry, in agreement with a triclinic space group P1. For X=180°, the self-rotation gives a number of well defined peaks (Fig. 3). The sharpest 3 to 7 peaks, all situated with orientations almost perpendicular to a*, could correspond to molecular dyad axes representing 3 to 7 dimers (6 to 14 subunits) in the crystal unit cell. The broad peak at an angle of about 15° from the a* axis could correspond to a relationship close to 180° between the orientation of different dimers (Fig. 3). Attempts at

finding an initial solution by molecular replacement (Navaza, 1994) using a polyalanine model of residues 1-258 of NAGK from *E. coli* (Ramón-Maiques *et al.*, 2002) are under way, but have given thus far no conclusive solutions.

Figure 3. Representation of the X= 180° section of the self-rotation function from the *P. aeruginosa* NAGK crystals. Start level at 1 σ and step size of 0.75 σ. The orthogonalization code follows the PDB convention. See text for interpretation.

Acknowledments

Supported by grants Rayos X of the Fundación Ramón Areces and GV01-259 of the Generalitat Valenciana. We thank the EU, ESRF and EMBL Grenoble for financial support for data collection. MLFM, SRM and FGO are fellows of Ministerio de Educación y Cultura de España, I3P of CSIC and of Fundación Ramón Areces, respectively. We thank J.J. Calvete, for N-terminal sequencing and mass spectrometry, D. Haas for providing the *P. aeruginosa* PAO1 cells and HG Britton for critical reading of the manuscript.

REFERENCES

Bradford, M.M. (1976) *Anal. Biochem.* **72,** 248-254.

Cunin, R., Glansdorff, N., Piérard, A., & Stalon, V. (1986). *Microbiol. Rev.* **50,** 314-352.

Davies, G.E. & Stark, G.R. (1970) *Proc. Natl. Acad. Sci. USA*, **66,** 651-656.

Evans, P. R. (1997). *Jnt. CCP4/ESF-EACBM Newslett. Protein Crystallogr.* **33,** 22-24.

Gil, F., Ramón-Maiques, S., Marina, A., Fita, I & Rubio, V. (1999) *Acta Cryst.* **D55,** 1350-1352.

Haas, D. & Leisinger, T. (1975a) *Eur. J. Biochem.* **52,** 365-375.

Haas, D. & Leisinger, T. (1975b) *Eur. J. Biochem.* **52,** 377-383.

Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24,** 409-411.

Leslie, A. G. W. (1990). *Crystallographic Computing.* Oxford University Press.

Leisinger, T & Haas, D. (1975) *J. Biol. Chem.* **250,** 1690-1693.

Matthews, B.W. (1968) *J. Mol. Biol.* **33,** 491-497.

Navaza, J. (1994) *Acta Cryst.* **A50,** 157-163.

Ramón Maiques, S., Marina, A., Gil , F., Fita, I. & Rubio, V. (2002). *Structure*, in press.

Rossmann, M.G & Blow, D.M. (1962) Acta Cryst. **15,** 24-31.

Shargool, P.D., Jain, J.C. & McKay, G. (1988) *Phytochemistry.* **6,** 1571-1574.

Vagin, A. & Teplyakov, A. (1997) *J. Appl. Cryst.* **30,** 1022-1025.

RESUMEN
DE LOS RESULTADOS

RESUMEN DE LOS RESULTADOS

Los resultados obtenidos se han estructurado en cinco artículos, que constituyen los cinco capítulos de la presente tesis. A continuación se presenta un breve resumen de en cada uno de ellos.

Capítulo 1:

N-Acetyl-L-glutamate kinase from *Escherichia coli*: cloning of the gene, purification and crystallization of the recombinant enzyme and preliminary X-ray analysis of the free and ligand-bound forms.

Publicado en *Acta Crystallographica Section D* (1999) **D55**, 1350-1352.

En este capítulo se describe la clonación mediante PCR en un vector plasmídico adecuado para expresión (pET-15b; de Novagen), del gen que codifica la NAGK de *Escherichia coli*. También se describe la hiperexpresión en *E. coli* de dicho gen, y la purificación del enzima recombinante, así como su caracterización básica, cristalización, y estudios cristalográficos preliminares mediante difracción de rayos X.

La secuencia del gen, determinada originalmente por Parsot et al. (1988), se utilizó para diseñar cebadores adecuados para su amplificación mediante PCR a partir de DNA genómico de *E. coli* de la cepa Bl21(DE3) (de la casa Novagen) usando una polimerasa termoestable de alta fidelidad (Deep Vent DNA polimerasa, de New England Biolabs). Los cebadores se diseñaron para que los extremos del amplicón incluyeran nuevos sitios para los enzimas de restricción *Bsp*HI y *Blp*I, con la finalidad de que tras la digestión con tales enzimas se generaran lugares adecuados para la ligación direccional dentro del vector pET-15b previamente digerido con *Nco*I y *Blp*I. El análisis electroforético (tras linearización con *Blp*I) de los plásmidos aislados de colonias individuales obtenidas en medio con ampicilina (a la que confiere resistencia el plásmido) de células DH5α (de Clontech) transformadas permitió aislar el plásmido con el inserto adecuado, como demostró la secuenciación del mismo, denominando a dicho plásmido pNAGK24. La transformación con el mismo de células de *E. coli* BL21(DE3), seguido de la siembra en medio líquido de colonias individuales obtenidas tras esa transformación y del crecimiento hasta el final de fase exponencial y la inducción durante 3 horas con 1 mM IPTG, hizo posible el aislamiento de colonias con un alto nivel de expresión (revelada mediante SDS-PAGE y tinción de Coomassie). A partir de una de ellas se ha obtenido el enzima tras preparar cultivos de unos 3 L en medio LB

con ampicilina, llevados al final de la fase exponencial con agitación e inducidos del mismo modo con IPTG.

La purificación del enzima se ha realizado en unos pocos pasos: rotura de las células mediante sonicación, centrifugación para clarificar el extracto, dos cortes de sulfato amónico (40 y 60 % de saturación), diálisis y cromatografía de intercambio iónico (Q-Sepharose fast flow, de Amersham Pharmacia) seguida de cromatografía de afinidad en azul cibacron-agarosa (Affi-Gel Blue, de Bio-Rad), con elución del enzima con alta pureza (>90 %) mediante la aplicación de ATP y NAG a la columna. En promedio se obtuvieron 10 mg de enzima por litro inicial de cultivo.

La masa del polipéptido determinada mediante SDS-PAGE fue de 27 kDa, en excelente acuerdo con lo esperado a partir de la secuencia (27.15 kDa). La actividad enzimática (60 U/mg, actividad determinada a 37°C) fue similar a la observada previamente para el enzima purificado de *Pseudomonas aeruginosa* (la NAGK bacteriana mejor caracterizada hasta ahora, Haas & Leisinger, 1975a). La secuenciación N-terminal arrojó la secuencia esperada MNPLIIK excepto por la perdida de la Met1 que, como es usual, había sido eliminada postranscripcionalmente. El entrecruzamiento con dimetilsuberimidato probó el carácter dimérico del enzima.

La cristalización se realizó íntegramente por el método de difusión de vapor en gota colgante. Se exploraron las condiciones iniciales mediante el uso de matrices de cristalización (Crystal Screen, de Hampton). Los mejores cristales sin sustratos se obtuvieron en aproximadamente una semana a 20°C, usando como precipitante polietilenglicol 4K en presencia de 0.1 M citrato sódico pH 5,6 y 0,1-0,3 M sulfato amónico. En presencia de sustratos (12 mM NAG, 15 mM $MgCl_2$, 3 mM AMPPNP) los mejores cristales se obtuvieron usando el éter monometílico del polietilenglicol 2 K en un rango de 27-32 %, en presencia de 0,1-0,3 M sulfato amónico y de 0,1 M acetato sódico pH 4,6. Se obtuvieron cristales de hasta 0,6 mm en la dimensión máxima, y se difractaron a -170°C utilizando una fuente convencional de rayos X hasta resoluciones de 2.3 y 2.95 Å, en presencia y ausencia de sustratos, respectivamente. Los cristales preparados sin sustratos pertenecieron al grupo espacial $P6_122$ con parámetros de celdilla a=b=78.6, c=278.0 Å, estimándose a partir del volumen de celdilla y de la masa molecular de la proteína, la presencia de dos monómeros por unidad asimétrica. Por otro lado, los cristales del complejo con MgAMPPNP y NAG eran ortorrómbicos, pertenecientes al grupo espacial $C222_1$, con parámetros de celdilla a=60.0 Å, b=71.9 Å y c=107.4 Å, estimándose que hay un monómero de proteína por unidad asimétrica.

Capítulo 2:

Structure of acetylglutamate kinase, a key enzyme for arginine biosynthesis and a prototype for the amino acid kinase enzyme family, during catalysis.

Publicado en *Structure* (2002) **10**, 329-342.

En este capítulo se describe la estructura tridimensional a nivel atómico de la N-acetil-L-glutamato quinasa, el miembro de la familia aminoácido quinasa que cataliza el paso segundo y frecuentemente punto de control de la biosíntesis de arginina. La estructura cristalina de la NAGK de *E. coli* a 1.5 Å de resolución revela un homodímero de una subunidad de 258 residuos nucleado por una hoja β molecular central abierta que abarca todo el dímero y que consta de 16 elementos, emparedada entre α hélices. En cada subunidad se une una molécula de AMPPNP, formando un complejo de sus fosfatos αβγ con Mg^{2+}, dispuesta a lo largo del borde C-terminal de la hoja β, mientras que el acetilglutamato (NAG) se une cerca del eje diádico, con su γ-COO^- alineado con el γ-fosforilo, a corta distancia del mismo, indicando la existencia de transferencia asociativa del γ-fosforilo asistida por: (1) la formación del complejo con Mg^{2+}; (2) las cargas positivas en la Lys8, Lys217 y en dos dipolos en los extremos N-terminales de dos α-hélices; y (3) mediante formación de puentes de hidrógeno entre la cadena polipeptídica y el γ-fosfato. La similitud estructural con la carbamato quinasa y el alineamiento de las secuencias sugiere que la NAGK es un prototipo estructural y funcional de la familia aminoácido quinasa, que difiere de forma importante de otras maquinarias de fabricación de acilfosfatos representadas por la fosfoglicerato quinasa, la acetato quinasa y la biotina carboxilasa.

En este trabajo de cinco autores no soy primer autor, sino co-segundo autor (Alberto Marina, quien figura el segundo en la lista de autores, y yo hemos contribuido igualmente al trabajo, como se señala en nota al pie de la línea de autores) habiendo estado implicado en la preparación del enzima con sustitución de la metionina por selenometionina (SeMet), paso que ha sido esencial para la determinación de las fases cristalográficas, en el refinado de la estructura hasta resoluciones atómicas, habiendo sido el responsable principal del trazado del nucleótido unido al enzima, y habiendo participado también en el análisis de la estructura y de las interacciones de los sustratos, y en los demás aspectos de la preparación del manuscrito, incluyendo la preparación de algunas de las figuras.

Con el fin de obtener las fases cristalográficas, se preparó el enzima con sustitución de la metionina por selenometionina, mediante transformación con pNAGK24 de la cepa de *E. coli* B834 (DE3) pLysS, que es auxótrofa para el aminoácido metionina, haciéndose crecer en un

medio en el que se sustituyó la metionina por SeMet, induciéndose la sobreexpresión del modo usual con IPTG. Aplicando técnicas de criocristalografía, cristales de tamaño adecuado, que habían sido cocristalizados con ligandos, fueron difractados con radiación de sincrotrón a tres longitudes de onda (MAD), calculándose la fases con el programa SOLVE. De este modo se obtuvieron datos de difracción con los que se construyó el modelo de la proteína sustituida con selenometioninas a una resolución de 1.85 Å. Separadamente se difractó también la proteína nativa (sin sustitución de la Met por SeMet), cristalizada con los mismos ligandos, consiguiendo a partir de estos nuevos datos de difracción, mediante reemplazo molecular a partir del modelo para la forma con SeMet, un nuevo modelo que pudo afinarse hasta 1.5 Å de resolución. La buena calidad de los datos de difracción y de los mapas resultantes de densidad electrónica han permitido construir un modelo preciso del enzima unido a una molécula de MgAMPPNP y a una molécula de NAG, que también incluye un buen número de moléculas de agua de posición bien definida.

El enzima es un homodímero que exhibe un plegamiento α-β, cuyo esqueleto principal consiste en una gran hoja β abierta de 16 elementos que se extiende de un extremo a otro de la proteína, cruzando la superficie de dimerización, mediante interacciones antiparalelas no canónicas, y estando rodeada por hélices α en ambas caras. Cada subunidad está compuesta por un lóbulo N-terminal y otro C-terminal, entre los que se define una cavidad en la que se localiza el centro activo del enzima. El lóbulo N incluye desde el residuo 1 al 173, y ocupa la parte central del dímero, aportando todos los residuos implicados en la superficie de interacción entre los dos monómeros, y formando el lugar de unión para el NAG. El lóbulo C incluye desde el residuo 174 al 258 y forma el sitio de unión para la porción ADP del ATP. La superficie de interacción entre los dos monómeros está formada íntegramente por residuos del lóbulo N predominantemente no polares, involucrando las hebra β5 y β9 y las hélices B, C y D.

El NAG se une en el lóbulo N, sobre la gran hoja β central, en el interior de una cavidad formada por tres lazos de éste lóbulo, por αC y, lateralmente, por αB, separada por αC de la superficie de dimerización. La apertura de la cavidad, donde se asienta el grupo γ-carboxilato del NAG, se orienta hacia el lóbulo C y está limitada por dos segmentos muy conservados que constituyen verdaderas firmas de NAGK: el segmento β2-αB, que contiene la secuencia ^{39}VIVHGGGxxV, y el β10-αE con la secuencia ^{158}NVNAD. La entrada a la cavidad está limitada por la horquilla β3-β4 que se sitúa sobre el NAG a modo de tapadera. La molécula de proteína establece numerosas interacciones con el NAG a través de los residuos Arg66 y Asn158, ambos

altamente conservados y pertenecientes a secuencias firma de la NAGK. Estas interacciones, el tamaño limitado de la cavidad (que la inhabilita para albergar sustituyentes mayores que el grupo N-acetilo), la conformación extendida del NAG (que hace que el N-acetil-L-aspartato sea demasiado corto para unirse productivamente), y la carga positiva del grupo guanidinio de la Arg66 (que repelería al grupo α-amino del glutamato) contribuyen a la gran especificidad de la NAGK por el NAG.

El AMPPNP, en conformación extendida, se une en un surco situado en la parte C-terminal de la hoja β principal, con la cadena de polifosfatos apuntando hacia la superficie de dimerización. La parte correspondiente al ADP se sitúa íntegramente en el lóbulo C, mientras que el γ-fosfato se interna en el lóbulo N enfrentado al γ-carboxilato del NAG. El AMPPNP establece numerosas interacciones con la proteína y está acomplejado a través de sus tres fosfatos a un ión Mg^{2+}, que a la vez está coordinado con dos moléculas de agua, una de las cuales interacciona con la proteína y la otra con el NAG. La adenina, en conformación *syn,* esta rodeada por dos bucles del lóbulo C. La purina reposa apilada sobre cadenas laterales de residuos hidrofóbicos, en especial de la Met214. La otra cara de la adenina está expuesta. El γ-fosfato mantiene contactos con dos secuencias firma de la NAGK ricas en glicinas situadas en β1-αA (^{6}I(I/f)KxGG) y β2-αA (^{39}VIVHGGGxxV). El nucleótido también establece contactos con los grupos ξ-amino de Lys8 y Lys217, de las que la primera interacciona con el γ-fosfato y la segunda con el β-fosfato. Ambas participan en la neutralización de las cargas negativas del nucleótido, junto con los extremos positivos de los dipolos de las hélices B y E. Los grupos amino de Lys8 y Lys217 interaccionan con el carboxilato del Asp162, un residuo absolutamente conservado que puede tener una función organizadora crucial.

Los átomos de $O_{\beta\gamma}$ y el P_γ del ATP así como el O atacante del NAG están casi en línea (ángulo OPO, 155º). Llama la atención la presencia de una intensa densidad electrónica entre el γ-carboxilato del NAG (grupo atacante) y el γ-fosfato del AMPPNP (grupo que se transfiere), así como la corta distancia entre el átomo de P_γ y el O atacante, de sólo 2.8 , mucho menos que la suma de los radios de van der Waals para estos dos átomos (3.3). Todo ello es indicativo de la formación parcial de enlace entre el fosforilo que se va a transferir y el grupo atacante, por lo que esta estructura representa una instantánea del enzima en el proceso de catalizar la reacción. La corta distancia P_γ-O_{NAG} (2.8 Å) indica que el mecanismo de transferencia de fosforilo es al menos 10.8 % asociativo, y por tanto que debe producirse un intermediario o estado de transición de fósforo pentavalente y carácter bipiramidal. La disposición de los residuos catalíticos del enzima

también está de acuerdo con una transferencia con características asociativas. En especial cabe destacar la presencia de las lisinas 8 y 217 altamente conservadas, que orientan sus cadenas laterales cargadas positivamente hacia el centro de la cavidad en que sucede la transferencia y que podrían desempeñar un papel clave en la neutralización de cargas negativas durante el proceso. El resto de grupos implicados en la catálisis serían el átomo de Mg^{2+}, el extremo N-terminal y por tanto positivo de αB y de KE, y enlaces donadores de hidrógenos con las glicinas 11 y 44, perteneciente a las dos secuencias firma de la NAGK. En este trabajo queda incierto el papel de la Lys61, un residuo situado demasiado lejos del centro activo aunque apunta hacia el lugar de la transferencia; en CK una lisina localizada en una posición similar interacciona con un sulfato y se ha propuesto participe en la catálisis. Sin embargo, los resultados del capítulo 4 excluyen la implicación catalítica de la Lys61 de NAGK.

Aunque no se resume aquí, este capítulo superpone los centros activos de NAGK y de la carbamato quinasa de *P. furiosus*, para concluir que la porción ADP del AMPPNP en la NAGK y el ADP unido a la CK adoptan esencialmente la misma conformación y se unen de modo muy parecido a residuos similares, siendo un punto de diferencia el hecho de que en *P. furiosus* la adenina esté emparedada entre la metionina que forma el suelo de su lugar de unión, y una tirosina que la recubre, lo que seguramente justifica en parte la mucha mayor afinidad por el nucleótido del enzima de *P. furiosus* que de la NAGK.

También se comparan las estructuras de NAGK y CKs para concluir que esencialmente son similares, diferenciándose en la presencia en la CK de un subdominio pequeño y lejano que en la NAGK se sustituye por una horquilla β, y en la constitución de la segunda hoja β de ambos enzimas, que exhibe ciertas diferencias entre uno y otro. En todo caso predominan fuertemente las similitudes sobre las diferencias. De hecho, en este capítulo se extiende la comparación a los otros enzimas de la familia aminoácido quinasa, cuyas secuencias se superponen sobre el alineamiento basado tanto en similitud de secuencia como en estructura de CK y NAGK. El resultado de esta superposición es la conclusión de que todas las enzimas de la familia aminoácido quinasa exhiben la arquitectura básica de la NAGK, y que probablemente comparten también la localización de los sitios de unión de los sustratos, y el modo de unión de estos.

Se concluye también que otros enzimas que transfieren un grupo fosforilo a un grupo carboxilato, de los incluidos en el grupo EC 2.7.2, como la acetato quinasa o la fosfoglicerato quinasa, exhiben arquitecturas distintas a la de la NAGK, lo que indica que este grupo consta de al menos tres familias estructurales distintas. Además se señala que la biotina carboxilasa y carbamil fosfato sintetasa, de las que originalmente se pensó compartirían el mismo tipo de

estructura con la CK, presentan en realidad distintos patrones organizativos a los de la familia de la NAGK.

Capítulo 3:

A crystallographic glimpse of a nucleotide triphosphate (AMPPNP) bound to a protein surface. External and internal AMPPNP molecules in crystalline N-acetyl-L-glutamate kinase.

Publicado en *Acta Crystallographica Section D* (2002) **D58**, 1892-1895.

En este capítulo se describe la aparición en el cristal de la acetilglutamato quinasa de *E. coli* de un gran volumen de densidad electrónica inexplicada que hemos interpretado como una molécula de AMPPNP externa, muy extendida y libre de metal, que ocupa dos posiciones alternativas y que establece contactos con la proteína exclusivamente a través de su grupo γ-imidofosfato. Además, este nucleótido externo se compara con el nucleótido unido al centro activo y se analizan las razones de su conformación extendida, ausencia de metal unido y unión periférica a la proteína. Por otro lado, se discute la posibilidad de que este AMPPNP externo se encuentre esperando para ocupar el centro activo del enzima.

Durante el trazado de la NAGK de *E. coli* ligada a MgAMPPNP y NAG, sorprendía la aparición de un gran volumen de densidad electrónica con apariencia, en proyección, de rana, localizada entre las diferentes moléculas del enzima en el eje binario cristalográfico coincidente con el eje diádico del homodímero de NAGK, que contactaba con el enzima a través de su parte inferior (las patas de la rana), específicamente con la conexión entre αB y β3, y que no podía ser interpretado como parte de la cadena polipeptídica de la proteína.

Teniendo en cuenta que el número de componentes moleculares presentes en la solución de cristalización era limitado, concluimos que se trataba de una molécula de AMPPNP externo ocupando dos posiciones alternativas, con un 50 % de ocupación cada una, alrededor del eje binario del cristal. Unos pocos ciclos de refinado de la molécula fueron suficientes para la construcción del modelo, cuyos ángulos y longitudes de enlace están dentro de los esperados para el AMPPNP. La adenina y la ribosa se superponen con la posición alternativa relacionada por simetría, y encajan en la cabeza y cuerpo de la rana, y las dos posiciones alternativas de la cadena de polifosfatos del AMPPNP encajan con las patas traseras de la rana. La presencia de varios átomos en la cercanía de la superficie de la densidad electrónica es debido a que son átomos

electrónicamente débiles que no se superponen con la molécula relacionada por simetría. Las únicas densidades positivas no explicadas por el modelo son las que generan la apariencia de las patas delanteras de la rana, que posiblemente correspondan a solvente o a posiciones alternativas con baja ocupación de la base adenina.

La molécula de AMPPNP externo presenta una buena estereoquímica: el anillo de adenina se dispone en conformación *anti* y la ribosa en conformación C3´-*endo*, con ángulo de torsión γ (enlace C4´-C5´) en el límite para –sinclinal, permitiendo el que los sustituyentes en los C4' y C5' no se eclipsen y la formación de un puente de hidrógeno canónico intramolecular entre los átomos de O 3' y 5' de la ribosa, puente que puede contribuir a la estabilidad de la forma extendida del nucleótido.

Pese a la alta resolución del complejo no se observa la presencia de magnesio coordinado con el nucleótido externo, lo que no es de extrañar ya que al pH de la cristalización y a las concentraciones de AMPPNP y Mg^{2+} usadas sólo la tercera parte del AMPPNP debe formar un complejo con el AMPPNP. Sin embargo, la molécula de AMPPNP unida al centro activo sí está unida como complejo de sus tres fosfatos con un ión Mg^{2+}. Se compara la estructura de este nucleótido externo con el unido en el centro activo, pero dicha comparación no se resume aquí.

Las únicas interacciones que establece el nucleótido externo con la proteína están mediadas por el grupo γ-imidofosfato terminal. Este grupo interacciona con tres residuos: Gly54, Asn56 y Lys53, que pertenecen a αB y a β3. Para que estos escasos contactos se establezcan, adquiere gran importancia la presencia del átomo de N puente β-γ del AMPPNP, así como el pH de cristalización.

Se discute la significación del nucleótido externo, señalando la posibilidad de que se encuentre en espera para incorporarse al centro activo. Aunque atractiva, esta posibilidad parece improbable, pero en todo caso la estructura descrita representa una rara observación del choque de un nucleótido prácticamente en solución, con una molécula de proteína.

Capítulo 4:

The course of phosphorus in the reaction of N-acetyl-L-glutamate kinase, determined from the structures of crystalline complexes, including a complex with an AlF_4^- transition state mimic.

Aceptado en *J. Mol. Biol.*

En este capítulo se determinan tres nuevas estructuras cristalinas de la NAGK: formando complejo con MgADP, NAG y el análogo del estado de transición tetrafluoruro de aluminio (AlF_4^-); con MgADP y NAG; y con ADP y SO_4^{2-}. La comparación de estas estructuras con el complejo con MgAMPPNP-NAG permite delinear tres etapas sucesivas durante la transferencia del fosforilo: al comienzo, cuando los átomos de oxígeno atacante y saliente y el átomo de P se encuentran alineados de forma imperfecta y la distancia entre el átomo de oxígeno atacante y el átomo de P es de 2.8 Å; a medio camino, cuando se forma el intermediario bipiramidal, con un alineamiento casi perfecto y una distancia de 2.3 Å; y cuando la transferencia se ha completado. La transferencia se da en línea y es fuertemente asociativa, con la Lys8 y Lys217 estabilizando el estado de transición y el grupo saliente, respectivamente, mientras que la Lys61 no participa, en contra de lo que se propuso anteriormente. Además, tres moléculas de agua encontradas en los tres complejos desempeñan un papel estructural crucial junto con el Asp162 y el Mg^{2+}. También es muy importante el papel de dos lazos ricos en glicinas (β1-αA y β2-αB) moviéndose en los diferentes complejos en concierto con los ligandos y estableciendo puentes de hidrógeno con ellos, situándolos en la posición adecuada para reaccionar o estabilizando el estado de transición. Se observa que el centro activo es demasiado estrecho para acomodar los sustratos sin comprimir los grupos reaccionantes, produciendo una deformación compresiva de los mismos en la dirección del estado de transición, que parece ser un componente crucial del mecanismo catalítico de la NAGK, y quizá también de otros enzimas de la familia aminoácido quinasa, tales como carbamato quinasa. Debido a esto, la unión inicial de los dos sustratos requeriría una conformación del enzima diferente a la observada, en la que su sitio activo sería más ancho. La energía de unión de los sustratos sería usada para cambiar la conformación del centro activo, causando la deformación de los sustratos hacia el estado de transición.

Se utilizó el método de difusión de vapor en gota colgante para preparar tres nuevos complejos de la NAGK con sustratos: con MgADP y NAG presentes con (cocristalizando con 80 mM NaF y 10 mM $AlCl_3$) y sin el análogo del estado de transición AlF_4^- interpuesto y con ADP y

SO_4^{2-}. Los cristales obtenidos eran de 0.3 mm aproximadamente y crecían en 1 o 2 semanas, permitiendo la recogida de datos de difracción a una resolución de 1.9 Å. Los modelos dan estructuras homodiméricas con plegamiento general idéntico al del complejo con MgAMPPNP-NAG. Las diferencias más importantes aparecen en el giro 21-KA (^{10}GGVLLD^{15}) y en la conexión entre 22 y KB (^{43}GGG^{45}) que parecen estar fuertemente implicadas en la transferencia del grupo fosforilo. Mientras que en los complejos MgAMPPNP-NAG y MgADP-AlF_4^--NAG el giro 21-KA permanece en una posición intermedia, en los complejos MgADP-NAG y ADP-SO_4^{2-} se mueve en direcciones opuestas, interaccionando con el 2-fosfato del ADP. A la vez, en el complejo ADP-SO_4^{2-} la conexión 22-KB adquiere la forma de un meandro que acomoda al sulfato.

La disposición del NAG en todos los complejos es casi idéntica a la encontrada en el complejo con MgAMPPNP-NAG, excepto en el complejo con ADP-SO_4^{2-} que carece de NAG. En su lugar, el hueco dejado por el NAG es ocupado por una molécula de acetato en posición similar al grupo acetamido del NAG, y por cinco moléculas de agua, dos de las cuales se localizan aproximadamente en la posición de los oxígenos del K-carboxilato y otra localizada en una posición similar a la del átomo de oxígeno no atacante del γ-carboxilato, mientras que un átomo de oxígeno del ión sulfato se sitúa en la posición del oxígeno atacante del NAG. El sitio de unión del sustrato a fosforilar aparece un poco deformado debido al ligero movimiento en los elementos de estructura secundaria que lo rodean. Además, la Lys61 se desplaza hacia el Asp212 formando un puente salino con él y la Arg66 exhibe doble conformación.

La orientación de la molécula de ADP es en general la misma en todos los complejos excepto en el complejo ADP-SO_4^{2-}, el cual carece del ión magnesio, resultando en una diferente orientación del 2-fosfato. En este complejo los residuos Ser180, Asp181 y Asp212 se reorientan, acomodándose a la nueva posición del 2-fosfato. En el resto de complejos las diferencias más importantes corresponden a la doble conformación de los residuos Ser180 y Asp212 y al desplazamiento de aproximadamente 1 Å en direcciones opuestas en los complejos MgADP-NAG y ADP-SO_4^{2-} de la porción AMP del nucleótido, así como de los lazos que la rodean y de la Met214 que forma la base sobre la que se asienta la adenina.

Un hallazgo de gran importancia es la identificación de una densidad plana con aspecto cuadrado localizada entre el 2-fosfato del ADP y el γ-carboxilato del NAG, que se ajusta a la esperada para el análogo del estado de transición AlF_4^- (forma esperada a al pH de la cristalización). En este complejo se observa una menor distancia del átomo de aluminio al átomo de oxígeno atacante del NAG, en comparación con el P_γ del AMPPNP (con distancias de 2.33 y

2.8 Å, respectivamente), y con una disposición casi perfectamente alineada entre el O saliente, el Al y el O atacante (ángulo OAlO, 176°). Estas características acercan este complejo al estado de transición para la transferencia en línea del grupo fosforilo en la reacción catalizada por la NAGK, incluso más que el complejo descrito previamente con AMPPNP y NAG. En el complejo MgADP-AlF_4^--NAG el átomo de aluminio está coordinado de forma octaédrica con los cuatro átomos de flúor ecuatoriales y con los átomos de oxígeno atacante y saliente del NAG y del ADP, y únicamente un átomo de flúor está expuesto al solvente. El resto de átomos de flúor establecen las mismas interacciones que los átomos de oxígeno del γ-fosfato del AMPPNP, con pequeñas diferencias que indican interacciones similares o incluso más fuertes. El resto de interacciones con la proteína de éste complejo son idénticas a las del complejo MgAMPPNP-NAG. Así, el átomo de magnesio, el Asp162 y tres aguas fijas (W1, W2 y W3) actúan de elementos organizadores clave del centro activo, orientando adecuadamente, por ejemplo, las lisinas 8 y 217.

En el complejo MgADP-NAG la estructura general del centro activo es esencialmente la misma que la observada en los complejos MgAMPPNP-NAG y MgADP-AlF_4^--NAG. Sin embargo, aparecen dos moléculas más de agua, además de las descritas en el párrafo anterior. W4 ocupa una posición equivalente al átomo de oxígeno del γ-fosfato del AMPPNP que forma parte de la esfera de coordinación del ión magnesio; y W5 ocupa la posición del átomo de oxígeno del γ -fosfato del AMPPNP que contacta con la Lys8, interaccionando con el 2-fosfato del ADP y con el γ-carboxilato del NAG, de modo que ambas aguas ocupan parte del espacio dejado por el γ-fosfato, imitando sus conexiones. Así, el 2-fosfato en este complejo esta fijado por una extensa red de interacciones, incluyendo una nueva interacción con la Lys217 y otra a través de una molécula de agua con el Asp181. Estas interacciones pueden imitar a aquellas establecidas por el ADP en el complejo de la NAGK con los productos. Otra característica de este complejo es la perfecta disposición octaédrica de la esfera de coordinación del magnesio, con la sexta posición de coordinación ocupada, en contraste con el resto de los complejos, que carecen de la sexta agua y en los que la geometría de la esfera de coordinación está algo distorsionada, con distancias de enlace en general mayores que en los otros complejos, sugiriendo que la coordinación del magnesio en esos otros complejos está sometida a tensión.

En el complejo ADP-SO_4^{2-} se localizó un ión sulfato en una posición similar al encontrado en la CK de *Enterococcus faecalis*. Este ión sulfato podría corresponder al fosfato del producto NAG-fosfato (NAGP), ya que: a) está más próximo al NAG que el γ-fosfato del AMPPNP o el tetrafluoruro de aluminio; b) uno de sus átomos de oxígeno ocupa aproximadamente la posición

del átomo de oxígeno atacante del NAG; y c) la disposición de los oxígenos es la que adoptaría el fosfato del NAG-fosfato tras la transferencia del fosforilo con inversión de configuración. En este complejo la conexión β2-αB se dispone en una conformación adecuada para acomodar el fosfato transferido, y el giro β1-αA se mueve para interaccionar con el β-fosfato del ADP. Las aguas fijas W1, W2 y W3 se disponen en igual posición y establecen las mismas interacciones que en el resto de los complejos, excepto con el Mg^{2+}, que está ausente en este complejo y es sustituido por una molécula de agua (W6) que imita las conexiones del metal. En este complejo, otra molécula de agua (W7) sustituye al átomo de oxígeno no atacante del NAG, imitando sus conexiones, e interaccionando además con el ión sulfato. Ésta última conexión probablemente esté ausente en el complejo genuino con los productos, ya que ninguno de los átomos implicados dispone de un átomo de hidrógeno para compartir. W7 también interacciona con otra molécula de agua (W8), que ayuda a fijar el sulfato a través de la Gly11, siendo otra interacción dudosa en el complejo con los productos.

Estos nuevos complejos cristalinos de la NAGK, isomorfos con el complejo con MgAMPPNP-NAG, han revelado que, al contrario que ocurre en este último, no presentan densidad electrónica alguna en el eje diádico cristalográfico correspondiente a una molécula de nucleótido externo que pueda representar el equivalente al AMPPNP descrito en el capítulo 3, por lo que la molécula de AMPPNP externa encontrada en el complejo MgAMPPNP-NAG es sin duda un hallazgo restringido al AMPPNP, posiblemente sin impacto funcional, excepto por el hecho de representar un nucleótido colisionando con una molécula de proteína.

Capítulo 5:

Towards structural understanding of feed-back control of arginine biosynthesis: cloning and expression of the gene for the arginine-inhibited N-acetyl-L-glutamate kinase from *Pseudomonas aeruginosa*, purification and crystallization of the recombinant enzyme and preliminary X-ray studies.

Publicado en *Acta Crystallographica Section D* (2002) **D58**, 1045-1047.

En este capítulo se describe la clonación mediante PCR en un vector plasmídico adecuado para expresión (pET-22b; de Novagen), del gen que codifica la NAGK de *Pseudomonas aeruginosa*. También se describe la hiperexpresión en *E. coli* de dicho gen, y la purificación del

enzima recombinante, así como su caracterización básica, cristalización, estudios cristalográficos preliminares mediante difracción de rayos X y análisis de los datos obtenidos del cristal.

Soy el tercer autor de este trabajo de cinco autores, habiendo estado implicado en la identificación del gen para la NAGK en el genoma de *P. aeruginosa*, entonces incompleto, no anotado, y depositado en una oscura base de datos (Proyecto Genoma de *Pseudomonas aeruginosa*, PA5323; http://www.pseudomonas.com). También participé en la clonación del gen mediante el diseño y verificación de los cebadores empleados, así como en la cristalización del enzima.

La secuencia del gen se utilizó para diseñar cebadores adecuados para su amplificación mediante PCR a partir de DNA genómico de *P. aeruginosa* de la cepa PAO1, (donada por el Dr. D. Haas, Universidad de Lausanne) usando una polimerasa termoestable de alta fidelidad (Deep Vent DNA polimerasa). El procedimiento usado para la clonación en pET22b y para la hiperexpresión fue en todo como en el capítulo 1, excepto por el uso de los enzimas de restricción *NdeI* y *HindI* para la digestión de amplicón y plásmido parental. Se denominó pNAGK-PA25 al plásmido con el inserto correspondiente al gen.

La purificación del enzima ha sido también muy parecida a la que se utilizó para la NAGK de *E. coli*, excepto por la elución del enzima de la columna final de Affi-Gel Blue con un gradiente de NaCl.

La masa del polipéptido determinada mediante SDS-PAGE (30 kDa) y espectrometría de masas (31.711 Da) estaba en excelente acuerdo con lo esperado a partir de la secuencia (31.718 Da). La actividad enzimática (88 U/mg, actividad determinada a 37°C) fue comparable, incluso superior, a la observada previamente para el enzima más puro obtenido de *Pseudomonas aeruginosa* (Haas & Leisinger, 1975a), inhibiéndose, como se esperaba, con 1 mM de arginina. La secuenciación N-terminal arrojó la secuencia esperada TLSRDDAAQVAKVLSEA excepto por la perdida de la Met1 que había sido eliminada postranscripcionalmente. El entrecruzamiento con dimetilsuberimidato probó el carácter dimérico del enzima, pero demostró también finas bandas correspondientes a trímero y quizá a tetrámero, confirmando que el enzima puede agregar a formas superiores al dímero.

La estrategia de exploración del espacio de cristalización fue la misma que para el enzima de *E. coli*. Los mejores cristales con sustratos (10 mM NAG, 15 mM $MgCl_2$, 5 mM AMPPNP) se obtuvieron en aproximadamente dos semanas a 20° C, usando como precipitante polietilenglicol 8K en presencia de 0.1 M cacodilato sódico pH 6.5 y 0.15-.17 M acetato magnésico. Se obtuvieron cristales de hasta 0,5 mm en la dimensión máxima, y se difractaron a -170°C

utilizando radiación de sincrotrón hasta una resolución de 2.75 Å. Los cristales preparados pertenecieron al grupo espacial P1 (triclínico), con parámetros de celdilla a= 71.86 Å, b= 98.78 Å, c= 162.9 Å, α= 91.49°, β= 92.03°, γ= 107.56°, estimándose a partir del volumen de celdilla y de la masa molecular de la proteína que puede haber de 6-18 monómeros por unidad asimétrica. El cálculo de la función de rotación propia del cristal indicó la ausencia de ejes cristalográficos con simetría rotacional, confirmando el grupo espacial triclínico. Además, la presencia en X= 180° de 3-7 picos bien definidos, que podrían corresponder a ejes diádicos moleculares, representarían de 3 a 7 dímeros (6-14 monómeros) en la unidad asimétrica del cristal.

Discusión General

DISCUSIÓN GENERAL

Entre otras contribuciones, este trabajo de tesis esclarece la estructura de la acetilglutamato quinasa, describe el modo de unión de los sustratos y propone un mecanismo de catálisis para este enzima, delineando el curso del fosforilo a lo largo del camino reaccional, y proveyendo una visión generalizable a todos los enzimas de la familia estructural aminoácido quinasa. Discutiremos aquí algunos de estos aspectos.

Como ya se ha indicado (Capítulo 2), el plegamiento de la NAGK es casi idéntico al de la carbamato quinasa (CK) (Marina et al., 1999; Ramón-Maiques et al., 2000). Ambos enzimas se estructuran en dos lóbulos (Fig. 1), el N-terminal que proporciona la superficie de dimerización y en el que se une el sustrato a fosforilar (demostrado en la NAGK e inferido para la CK), y el C-terminal, que une la porción ADP del nucleótido. Como ilustra la Fig. 1, la topología de los elementos de estructura secundaria de ambos enzimas difiere sólo en la presencia en la CK del llamado "subdominio lejano" (Marina et al., 1999), que no existe en la NAGK, habiendo sido reemplazado aparentemente en este último enzima por la horquilla β3-β4, a su vez ausente de la CK, y que, ocupando un lugar aproximadamente similar en el espacio, está constituida por una porción no equivalente de la secuencia del enzima. La función de estos dos elementos estructurales disimilares de ambos enzimas no ha sido aclarada, aunque ambos parecen cerrar el bolsillo para el sustrato a fosforilar, y podrían experimentar cambios conformacionales importantes durante el ciclo catalítico. El hecho de que ambos elementos emerjan del lóbulo N-terminal en la dirección del C-terminal, sobre la cavidad en la que sucede la catálisis, hace

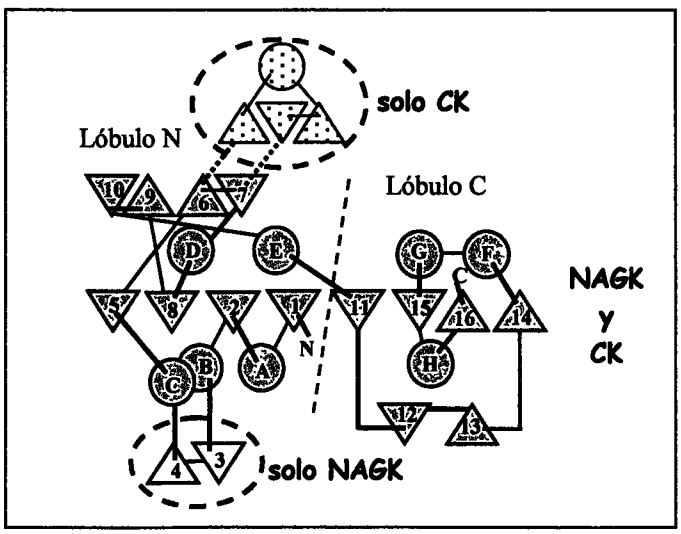


Figura 1. Topología de la NAGK de *E. coli* donde se resalta las similitudes con las carbamato quinasas. La denominación de los elementos de estructura secundaria es la de la NAGK de *E. coli*. La recta segmentada muestra la separación entre los lóbulos N y C.

concebible que sean elementos importantes de interrelación entre los lazos de ambos lóbulos que son verdaderamente los elementos funcionales, pudiendo participar en la propagación de movimientos coordinados que afecten a todo el centro activo. La observación previa de que mutaciones en residuos del subdominio lejano de la CK de *E. faecalis* (Marina et al., 1999) no tienen efecto importante alguno sobre la actividad del enzima, no excluye la posibilidad de que este dominio, y por extensión la horquilla β3-β4 de NAGK, tenga un papel importante, ya que las mutaciones que se introdujeron afectaron a residuos muy expuestos que miraban al exterior, y que por tanto no era previsible participaran en la catálisis o en la unión de los sustratos. No existen estudios comparando las estructuras de formas con y sin sustratos para la misma CK, ni la descripción estructural del complejo ternario con este enzima, pero existen diferencias estructurales importantes, caracterizadas como movimientos de cuerpo rígido, en el subdominio lejano de la CK cuando se compara la estructura del enzima de *E. faecalis*, que carece de sustratos, con la de *P. furiosus*, que contiene ADP unido (Ramón-Maiques et al., 2000). En el caso de la NAGK hemos postulado (capítulo 4) la necesidad de cambios conformacionales en el centro activo que hagan posible el acceso inicial de ambos sustratos sin compresión mutua, pero todavía no hemos determinado la estructura de la forma del enzima sin sustratos, y por tanto, aún no hemos podido demostrar si existen cambios importantes asociados a la unión de los sustratos y si estos cambios afectan a la horquilla β3-β4. Sin embargo, se han observado ya algunos pequeños cambios en dicha horquilla al comparar los diferentes complejos estudiados en el presente trabajo (Capítulo 4). Quizá en la forma de NAGK sin sustratos la horquilla β3-β4 se desplace hacia afuera para abrir el bolsillo para el NAG, facilitando el acceso de este sustrato.

La comparación de las estructuras de CK y NAGK permite establecer los rasgos básicos de la arquitectura que son comunes a ambos enzimas, y quizá a cualquiera de los de la familia aminoácido quinasa (Figs. 1 y 2). La hoja β central y las dos capas de hélices α que la rodean proveen de la cimentación o andamiaje básico del que emergen lazos en el lado del borde C-terminal de la hoja β con los que se construyen los lugares de unión de los sustratos. Esto es análogo a lo que sucede con los barriles αβ (del tipo barril TIM) (Nagano et al., 2002), en los que los sitios de unión de los sustratos se ubican sobre el polo C-terminal de la hoja β cerrada que constituye el barril. Como la hoja β molecular de la NAGK no se cierra sobre sí misma, es importante proveer de elementos de cierre del centro activo, de modo que pueda restringirse el acceso del agua durante el proceso catalítico, ya que un ataque nucleofílico por el agua sobre el fósforo haría del enzima una ATPasa. Desempeñan papeles importantes de cierre del centro activo tanto en la CK como en la NAGK la hélice B y la porción de hélice C que emergen del

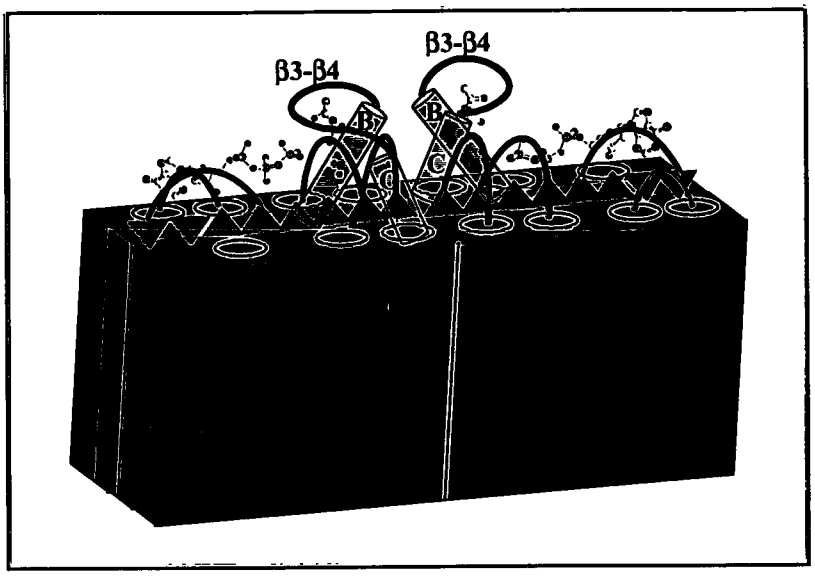


Figura 2. Arquitectura básica del dímero de NAGK, como una hoja β central (marrón, las puntas de flecha rojas indican el sentido de las hebras β) emparedada entre dos capas de hélices α (en azul; las hélices individuales se señalan como círculos verdes), de la que emergen, en el lóbulo C-terminal, dos bucles (en azul), y en el lóbulo N-terminal, las hélices B y C (señaladas con letras), el bucle β3-β4 y otros dos bucles que definen la segunda hoja β de cuatro elementos. Estos bucles se representan en magenta. Se esquematiza también en cada subunidad la posición de los dos sustratos.

borde C-terminal de la hoja β principal (Fig. 2), así como la horquilla β3-β4 de la NAGK o el subdominio lejano de la CK, y dos lazos adicionales (Fig. 2, en magenta) que definen una segunda hoja β (Fig. 1). El carácter dimérico de estos enzimas parece estar relacionado con la necesidad de conferir estabilidad estructural a esta maquinaria de cierre del centro activo, sin que existan evidencias de cooperatividad (Marshall & Cohen, 1966; Marina et al., 1998; y resultados nuestros no publicados sobre la NAGK de *E. coli*) que sugieran la existencia de un flujo importante de información entre los centros activos de las dos subunidades. Hay que destacar que una de las diferencias notables entre la CK de *P. furiosus*, que resiste temperaturas de 100° C (Durbecq et al. 1997), y la de *E. faecalis*, que se inactiva a los 55° C (Marina et al., 1998) es la mayor superficie de interacción entre subunidades en la primera, y un componente hidrofóbico más importante de esa superficie (Ramón-Maiques et al., 2000). Por tanto, esta observación es altamente sugerente de que el carácter dimérico sea también una aspecto importante para conferir estabilidad a estas proteínas enzimáticas.

El lugar de unión del nucleótido presenta unas características que son comunes también a CK (Ramón-Maiques et al, 2000) y NAGK, y que seguramente caracterizan a toda la familia aminoácido quinasa. El sitio para la porción adenosina viene delimitado entre el lazo que conecta las hélices antiparalelas αF y αG y el lazo de conexión β11-β12 (nomenclatura de NAGK; en

azul en la Fig. 2), apoyándose el anillo de la adenina sobre un plano hidrofóbico provisto por residuos que emergen de las primeras vueltas de αG. El fosfato γ del nucleótido interacciona principalmente con las conexiones β1-αA y β2-αB. De la unión al enzima del grupo atacante del otro sustrato puede decirse que los diferentes enzimas de la familia deben tener en común la interacción con la conexión β2-αB, así como con el comienzo de la hélice αE. Es difícil definir más las interacciones con los otros elementos del sitio para el sustrato a fosforilar, que es el que difiere entre los distintos enzimas de esta familia, aunque no cabe duda que deben participar en la constitución de dicho sitio la hélice αB y el elemento β10 de la segunda hoja β y su conexión con αE, pudiendo eventualmente participar también, como en el caso de NAGK, la horquilla β3-β4 o, en otros enzimas, el elemento estructural equivalente (subdominio lejano en la CK).

Con respecto a los elementos que participan en la catálisis, la comparación de NAGK y CK provee de unos pocos elementos comunes que parecen cruciales (Fig. 3). Estos elementos están constituidos por las dos cargas positivas de las cadenas laterales de dos lisinas, que emergen en la superficie de interacción con los fosfatos β y γ del ATP, los extremos positivos de los dipolos correspondientes a las hélices αB y αE, y el establecimiento de puentes de hidrógeno entre las conexiones β1-αA y β2-αB y el fosforilo terminal del ATP. A estos elementos hay que añadir la cadena lateral de un aspartato (D162 en la NAGK) que aunque no participa directamente en las interacciones con los sustratos, organiza las dos cargas positivas de las

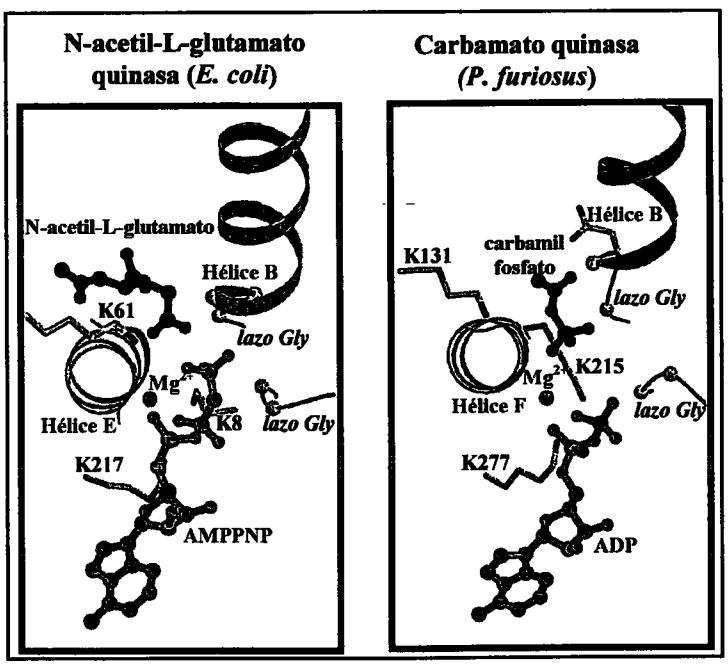


Figura 3. Comparación de los elementos de los centros activos de la NAGK de *E. coli* y de la carbamato quinasa de *P. furiosus* que parecen implicados en la catálisis. En la CK se ha modelado el carbamil fosfato.

lisinas implicadas, estructurando el centro activo. No siempre es fácil identificar todos estos elementos en las secuencias de los demás enzimas de la familia estructural. Así, una de las dos lisinas catalíticas (Lys8 de NAGK) procede de un lugar muy distante de la secuencia de la CK, aunque el grupo ξ-amino se ubica en el mismo lugar a pesar de las diferentes procedencias en NAGK y CK (Marina et al., 1999; Ramón-Maiques et al., 2000). En todo caso, el alineamiento de las secuencias (capítulo 2) permite identificar con alto grado de seguridad en todos los enzimas de la familia aminoácido quinasa las secuencias correspondientes a las conexiones β1-αA y β2-αB, que están altamente conservadas en todos los casos, los extremos N-terminales de las hélices αB y αE, el aspartato equivalente al 162 de NAG, y una de las dos lisinas catalíticas, que en todos los enzimas de la familia, excepto en las CKs, se alinea estrictamente con la Lys8 de la NAGK, quedando por identificar la segunda lisina catalítica, correspondiente a la 217 de la NAGK, en aspartoquinasas y UMP quinasas.

Sobresale entre las observaciones derivadas del estudio de los complejos de la NAGK la de que el enzima exhibe una notable complementariedad conformacional con el estado de transición, lo que nos ha llevado a concluir que el enzima estabiliza en alto grado dicho estado de transición desplazando el equilibrio entre el mismo y los sustratos en la dirección del estado de transición. Esta conclusión está plenamente de acuerdo con la vieja propuesta de Pauling (1946) de que éste es el mecanismo principal por el que los enzimas aceleran las reacciones que catalizan. Nuestra conclusión, contenida en el trabajo que constituye el capítulo 4 de esta Tesis, se ha visto corroborada muy recientemente en una publicación aparecida en el número de 28 de marzo de la revista *Science* (Lahiri et al., 2003), fecha posterior a nuestro envío del trabajo del Capítulo 4, y plenamente dentro del periodo final de la escritura de esta tesis. En esta publicación se realizan observaciones aún más directas para otra reacción de transferencia de fosforilo, la de la fosfoglucomutasa. Los autores han tenido la fortuna de atrapar en un cristal del complejo con

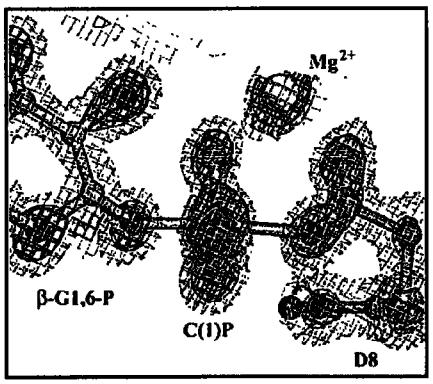


Figura 4. Imagen de la densidad electrónica $2F_o$-F_c y del modelo en representación de bolas y bastones del intermediario fosforano pentavalente en la reacción de la fosfoglucomutasa. Tomado de Lahiri et al, 2003.

Mg^{2+} y glucosa 1,6-(bis)fosfato de la β-fosfoglucomutasa de *Lactococcus lactis*, el intermediario de fósforo pentavalente bipiramidal (Fig. 4) postulado para la transferencia de fosforilo en línea, lo que proporciona una instantánea nueva e inédita del fosforilo a mitad de transferencia entre la posición 1 de la glucosa y el grupo carboxilo de la cadena lateral del aspartato en posición 8, demostrando que el proceso sucede en línea y tiene un fuerte carácter asociativo. Esta visión del estado de transición del enzima es muy parecida a la obtenida con la acetilglutamato quinasa para los complejos con MgAMPPNP-NAG y MgADP-AlF_4^--NAG, estando más próxima al complejo en presencia del análogo del estado de transición, puesto que muestra (Lahiri et al. 2003) unas distancias de enlace entre los dos átomos de oxígeno apicales y el átomo de fósforo, de 2.0 y 2.1 Å, con un ángulo de 174° ± 3°, valores cercanos a los 1.88 y 2.33 Å con un ángulo de 176°, obtenidos por nosotros. Como titula Jeremy Knowles (2003) en su comentario editorial a dicho artículo, "Seeing is believing", sin duda el artículo de *Science* documenta visualmente con el intermediario genuino lo inferido por nosotros para la NAGK sobre la base de sus complejos con análogos o de carácter abortivo ("dead-end").

En otro orden de cosas, la presente tesis abre las puertas al estudio de los mecanismos por los cuales se produce la inhibición por arginina en las NAGKs reguladas mediante inhibición de retroalimentación (Capítulo 5). Con este objetivo hemos dado los pasos preliminares que nos han conducido a la obtención de cristales de la NAGK de *P. aeruginosa*, así como de datos de difracción a una resolución de 2.75 Å. Cuando se purificó el enzima por primera vez a partir de su fuente natural (Haas & Leisinger, 1975a), la actividad se eluyó de una columna de Sephadex G-150 en una posición correspondiente a una masa molecular de 230 kDa, lo que representaría un agregado de siete monómeros (la masa del monómero es 31,85 kDa), pero otros experimentos de gel filtración con muestras purificadas o crudas del enzima revelaron al menos dos picos, indicando un sistema en el que el dímero parecía ser la unidad básica, pero en el que dependiendo de las condiciones se daban formas variadas de agregación. Los experimentos de entrecruzamiento con dimetilsuberimidato (capítulo 5) confirman este punto de vista: además del dímero evidencian formas de agregación superiores (particularmente experimentos adicionales no presentados aquí y realizados por la Dra. Fernández-Murga). La coexistencia de formas variables de agregación no es una condición favorable para el desarrollo de cristales (McPherson, 1999). Por tanto, cabía esperar un complicado proceso de cristalización del enzima de *P. aeruginosa*, de modo que el hecho de obtener cristales de esta proteína ha sido un logro muy importante que esperamos nos aporte información de gran relevancia que permita esclarecer el mecanismo de inhibición por retroalimentación de la NAGK, y, por extensión de

aspartoquinasa y glutamato 5-quinasa, otros dos enzimas de la misma familia que son también fuertemente inhibidos por el producto final.

CONCLUSIONES

CONCLUSIONES

1. La sobreexpresión del gen para la N-acetil-L-glutamato quinasa de *Escherichia coli*, clonado en un vector plasmídico, ha permitido aislar el enzima en forma pura, y cristalizarlo en una forma apropiada para el análisis de difracción de rayos X. Los cristales obtenidos en presencia de nucleótido son ortorrómbicos, del grupo $C222_1$, con dimensiones de celda unidad a, b y c, 59,4, 71,7 y 107,4 Å, respectivamente, conteniendo un monómero en la unidad asimétrica. Sin ligandos se han obtenido cristales hexagonales del grupo $P6_122$, y dimensiones de celda a=b=78.7 Å, c=277.7 Å, con dos monómeros en la unidad asimétrica.

2. Los estudios de difracción de rayos X han permitido determinar las estructuras del complejo del enzima con MgAMPPNP y N-acetilglutamato, a 1,5 Å de resolución, y las de los complejos con MgADP y NAG con o sin tetrafluoruro de aluminio interpuesto, o de ADP y sulfato, todos ellos a una resolución de 1,9 Å.

3. Todas las estructuras concuerdan en un mismo plegamiento básico, constituido por un homodímero nucleado por una hoja β molecular central de 16 elementos, rodeado de dos capas de hélices α, y con bucles y dos hélices emergiendo del borde C-terminal de la hoja β central de cada subunidad, formando dichos bucles los sitios de unión de los sustratos. La estructura concuerda en su mayor parte con la descrita previamente por el laboratorio para la carbamato quinasa, pudiendo constituir la acetilglutamato quinasa el paradigma para los demás enzimas de la familia aminoácido quinasa. La presencia en la acetilglutamato quinasa de acetilglutamato unido permite la caracterización por primera vez del sitio de unión del sustrato a fosforilar en estos enzimas.

4. Los diferentes complejos han permitido dilucidar el modo de unión del nucleótido y establecen las bases de la especificidad para el mismo, de la catálisis, así como el curso del grupo fosforilo desde los sustratos a los productos. La comparación con la carbamato quinasa revela un mismo modo de unión de nucleótido a ambos enzimas. Los complejos con MgAMPPNP-acetilglutamato y con MgADP-tetrafluoruro de aluminio-acetilglutamato revelan que la transferencia de fosforilo sucede en un sólo paso, en línea, con carácter fuertemente asociativo y con formación de un intermediario pentavalente bipiramidal. Las cargas positivas en dos lisinas conservadas en, los extremos N-terminales de dos hélices α, y

la formación de una red de puentes de hidrógeno del fosfato que se transfiere con la proteína, son elementos catalíticos clave. Un aspartato conservado, tres moléculas fijas de agua y el catión metálico divalente parecen elementos adicionales clave en la organización del centro activo.

5. El centro activo comprime los sustratos en la dirección del intermediario o estado de transición, proponiéndose que la complementariedad del enzima con dicho intermediario estabiliza este último y es un elemento catalítico clave con este enzima. Se propone también que el centro activo sufre fuertes cambios conformacionales con la unión de los sustratos, y que parte de la energía de dicha unión de los sustratos se utiliza para la generación de la conformación catalíticamente productiva.

6. La presencia de una molécula de AMPPNP unida periféricamente al enzima, muy extendida, no acomplejada con Mg^{2+}, nos informa acerca de la conformación que adopta el nucleótido en solución o en sus colisiones con moléculas de proteína. La hipótesis de que este nucleótido tenga importancia funcional para la reacción de la acetilglutamato quinasa no parece apoyada por la ausencia del nucleótido periférico en los demás complejos.

7. Se han iniciado los estudios para la caracterización de las bases moleculares de la inhibición "feed-back" de la acetilglutamato quinasa por arginina, mediante la cristalización de la acetilglutamato quinasa de *Pseudomonas aeruginosa*, que a diferencia del enzima de *Escherichia coli*, está sometida a este tipo de inhibición. El procedimiento utilizado ha sido como el usado para el enzima de *Escherichia coli*, excepto por el uso como molde del genoma de *Pseudomonas aeruginosa*. Se han obtenido cristales con MgADP y acetilglutamato, del grupo P1 (triclínico) con celda unidad de dimensiones a=71.86 Å, b=98.78 Å, c=162.9 Å, y ángulos α=91.49°, β=92.03°, γ=107.56°. El estudio de la función de rotación propia del cristal indica la ausencia de ejes cristalográficos con simetría rotacional y la presencia de 6 a 18 monómeros en la celda unidad del cristal.

BIBLIOGRAFÍA

correspondiente a la introducción general,

resumen de los resultados y discusión general

BIBLIOGRAFÍA CORRESPONDIENTE A LA INTRODUCCIÓN GENERAL, RESUMEN DE LOS RESULTADOS Y DISCUSIÓN GENERAL

Abadjieva, A., Pauwels, K., Hilven, P. y Crabeel, M. (2001) A new yeast metabolon involving at least the two first enzymes of arginine biosynthesis: acetylglutamate synthase activity requires complex formation with acetylglutamate kinase. *J. Biol. Chem.* **266**, 42869-42880.

Allen, C.M. y Jones, M.E. (1964) Decomposition of carbamylphosphate in aqueous solutions. *Biochemistry* **3**, 1238-1247.

Alonso, E. y Rubio, V. (1989) Participation of ornithine aminotransferase in the synthesis and catabolism of ornithine in mice. Studies using gabaculine and arginine deprivation. *Biochem. J.* **259**, 131-138.

Alonso, E., Cervera, J., García-España, A., Bendala, E. y Rubio, V. (1992) Oxidative inactivation of carbamoyl phosphate synthetase (ammonia). Mechanism and sites of oxidation, degradation of the oxidized enzyme, and inactivation by glycerol, EDTA and thiol protecting agents. *J. Biol. Chem.* **267**, 4524-4532.

Alonso, E. y Rubio, V. (1995) Affinity cleavage of carbamoyl phosphate synthetase I localizes regions of the enzyme interacting with the molecule of ATP that phosphorylates carbamate. *Eur. J. Biochem.* **229**, 337-384.

Auerbach, G., Huber, R., Grattinger, M., Zaiss, K., Schurig, H., Jaenicke y R., Jacob, U. (1997) Closed structure of phosphoglycerate kinase from *Thermotoga maritima* reveals the catalytic mechanism and determinants of thermal stability. *Structure* **5**, 1475-83.

Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S.R., Griffiths-Jones, S., Howe, K.L., Marshall, M. y Sonnhammer, E.L. (2002) The Pfam Protein Families Database. *Nucleic Acids Research* **30**, 276-280.

Caldovic, L. y Tuchman, M. (2003) N-acetylglutamate and its changing role through evolution. *Biochem. J.*

Caplow, M. (1968) Kinetics of carbamate formation and breakdown. *J. Am. Chem. Soc.* **90**, 6795-6803.

Chabre, M. (1990) Aluminofluoride and beryllofluoride complexes: a new phosphate analogs in enzymology. *Trends Biochem. Sci.* **15**, 6-10.

Chu, K., Vojtchovsky, J., McMahon, B.H., Sweet, R.M., Berendzen, J. y Schlichting, I. (2000) Structure of a ligand-binding intermediate in wild-type carbonmonoxy myoglobin. *Nature* **403**, 921-923.

Cleland, W.W. y Hengge, A.C. (1995) Mechanisms of phosphoryl and acyl transfer. *FASEB J.* **9**, 1585-94.

Cunin, R., Glansdorff, N., Piérard, A. y Stalon, V. (1986) Biosynthesis and metabolism of arginine in bacteria. *Microbiol. Rev.* **50**, 314-352.

Dénes, G. (1973) N-Acetyl-5-Phosphotransferase. *The Enzymes, 3thed.* Academic Press, New York y London.

Durbecq, V., Legrain, C., Roovers, M., Pierard, A. y Glansdorff, N. (1997) The carbamate kinase-like carbamoyl phosphate synthetase of the hyperthermophilic archaeon *Pyrococcus furiosus*, a missing link in the evolution of carbamoyl phosphate biosynthesis. *Proc. Natl. Acad. Sci. U S A.* **94**, 12803-12808.

Fersht, A. (1999) Structure and mechanism in protein science. A guide to enzyme catalysis and protein folding. W. H. Freeman & Co., New York.

Fry, P.A. (1992) Nucleotidyltransferases and phosphotransferases: stereochemistry and covalent intermediates. *The enzymes* 20, 141-186. Academic Press, NY.

Haas, D. and Leisinger, T. (1975a) N-acetylglutamate 5-phosphotransferase of *Pseudomonas aeruginosa*. Purification and ligand-directed association-dissociation. *Eur. J. Biochem.* **52**, 365-375.

Haas, D. and Leisinger, T. (1975b) N-acetylglutamate 5-phosphotransferase of *Pseudomonas aeruginosa*. Catalytic and regulatory properties. *Eur. J. Biochem.* **52**, 377-383.

Jencks, W.P. (1975). Binding energy, especificity, and enzimic catalysis: The circe effect. *Adv. Enzymol. Relat. Areas Mol. Biol.* **43**, 219-410.

Jin, L., Stec, B., Lipscomb, W.N. y Kantrowitz, E.R. (1999) Insights into the mechanisms of catalysis and heterotropic regulation of *Escherichia coli* aspartate transcarbamoylase based upon a structure of the enzyme complexed with the bisubstrate analogue N-phosphonacetyl-L-aspartate at 2.1 Å. *Proteins* **37**, 729-742.

Kack, H., Gibson, K.J., Lindqvist, Y. y Schneider, G. (1998) Snapshot of a phosphorylated substrate intermediate by kinetic crystallography. *Proc. Natl. Acad. Sci.* **95**, 5495-5500.

Knowles, J.R. (1980) Enzyme-catalyzed phosphoryl transfer reactions. *Annu. Rev. Biochem.* **49**, 877-919.

Knowles, J.R. (1982) Phospho transfer enzymes: lessons from stereochemistry. *Fed. Proc.* **41**, 2424-2431.

Lahiri, S.D., Zhang, G., Dunaway-Mariano, D., Allen, K.N. (2003) The pentacovalent phosphorus intermediate of a phosphoryl transfer reaction. *Science* **299**, 2067-2071.

Li, Y.F., Hata, Y., Fujii, T., Hisano, T., Nishihara, M., Kurihara, T. y Esaki, N. (1998) Crystal structures of reaction intermediates of L-2-haloacid dehalogenase and implications for the reaction mechanism. *J. Biol. Chem.* **273**, 15035-15044.

Marina, A, Bravo, J., Fita, I. y Rubio, V. (1994) Crystallization, characterization and preliminary crystallographic studies of carbamate kinase of *Streptococcus faecium*. *J. Mol. Biol.* **235**, 1345-1347.

Marina, A., Uriarte, M., Barcelona, B., Fresquet, V., Cervera, J. y Rubio, V. (1998) Carbamate kinase from *Enterococcus faecalis* and *Enterococcus faecium*: cloning of the genes, studies on the enzyme expressed in *Escherichia coli*, and sequence similarity with N-acetyl-L-glutamate kinase. *Eur. J. Biochem.* **253**, 280-91.

Marina, A., Alzari, P.M., Bravo, J., Uriarte, M., Barcelona, B., Fita, I. y Rubio, V. (1999) Carbamate kinase: New structural machinery for making carbamoyl phosphate, the common precursor of pyrimidines and arginine. *Protein. Sci.* **8**, 934-40.

Marshall, M y Cohen, P.P. (1966) A kinetic study of the mechanism of crystalline carbamate kinase. *J. Biol. Chem.* **241**, 4197-4208.

Matte, A., Tari, L.W. y Delbaere, L.T. (1998) How do kinases transfer phosphoryl groups? *Structure* **6**, 413-9.

McPherson, A. (1999) Crystallization of Biological Macromolecules. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Mildvan, A.S. y Fry, D.C. (1987) NMR studies of the mechanism of enzyme action. *Adv. Enzymol. Relat. Areas Mol. Biol.* **59**, 241-313

Mildvan, A.S. (1997) Mechanisms of signaling and related enzymes. *Proteins Struc. Func. Gen.* **29**, 401-416.

Nagano, N., Orengo, C.A. y Thornton, J.M. (2002) One fold with many functions: the evolutionary relationships between TIM barrel families based on their sequences, structures and functions. *J. Mol. Biol.* **321**, 741-65.

Parsot, C., Boyen, A., Cohen, G.N. y Glansdorff, N. (1988) Nucleotide sequence of *Escherichia coli* argB and argC genes: comparison of N-acetylglutamate kinase and N-acetylglutamate-gamma-semialdehyde dehydrogenase with homologous and analogous enzymes. *Gene* 68, 275-283.

Pauling, L. (1946) Molecular architecture and biological reactions. *Chem. Eng. News.* **24**, 1375-1377.

Ramon-Maiques, S., Marina, A., Uriarte, M., Fita, I. y Rubio, V. (2000) The 1.5 resolution crystal structure of the carbamate kinase-like carbamoyl phosphate synthetase from the

hyperthermophilic archaeon *Pyrococcus furiosus*, bound to ADP, confirms that this thermostable enzyme is a carbamate kinase, and provides insight into substrate binding and stability in carbamate kinases. *J. Mol. Biol.* **299**, 463-76.

Rubio, V. y Grisolía, S. (1977) Mechanism of mitochondrial carbamoyl-phosphate synthetase. Synthesis and properties of active CO_2, precursor of carbamoyl phosphate. *Biochemistry* **16**, 321-329.

Rubio, V., Britton, H.G., Grisolía, S., Sproat, B.S. y Lowe, G. (1981) Mechanism of activation of bicarbonate ion by mitochondrial carbamoyl phosphate synthetase: formation of enzyme-bound adenosine diphosphate from the adenosine triphosphate that yields inorganic phosphate. *Biochemistry* **20**, 1969-1974.

Rubio, V., Britton, H.G., Rodríguez-Aparicio, L. y Climent, I. (1990) Carbamate synthases and kinases. *Enzymatic and model carboxylation and reduction reactions for carbon dioxide utilization* (Aresta, M. y Schloss, J.V., eds.), pp. 221-238. Kluwer Academic Publishers, Boston.

Rubio, V., Cervera, J., Lusty, C. J., Bendala, E. y Britton, H. G. (1991) Domain structure of the large subunit of *E. coli* carbamoyl phosphate synthetase. Location of the binding site for the allosteric inhibitor UMP in the COOH-terminal domain. *Biochemistry* **30**, 1068-1075.

Rubio, V. (1994) Structure-activity correlations on carbamoyl phosphate synthetases. En *Carbon dioxide fixation and reduction in biological and model systems* (Branden, C.I. and Schneider, G., eds.), pp. 249-264. Oxford University Press, London.

Rubio, V. y Cervera, J. (1995) The carbamyl phosphate synthetase family and carbamate kinase: structure-function studies. *Biochem. Soc. Transact.* **23**, 879-883.

Rubio, V., Llorente, P. y Britton, H.G. (1998) Mechanism of carbamoyl phosphate synthetase from *Escherichia coli*. Binding of the ATP molecules used in the reaction and sequestration by the enzyme of the ATP molecule that yields carbamoyl phosphate. *Eur. J. Biochem.* **255**, 262-270.

Schlichting, I. y Reinstein, J. (1999) pH influences fluoride coordination number of the AlF_x phosphoryl transfer transition state analog. *Nature Struc. Biol.* **6**, 721-723.

Thoden, J.B., Holden, H.M., Wesenberg, G., Raushel, F.M. y Rayment, I. (1997) Structure of carbamoyl phosphate synthetase: a journey of 96 A from substrate to product. *Biochemistry* **36**, 6305-6316.

Uriarte, M., Marina, A., Ramon-Maiques, S., Fita, I. y Rubio, V. (1999) The carbamoyl-phosphate synthetase of *Pyrococcus furiosus* is enzymologically and structurally a carbamate kinase. *J. Biol. Chem.* **274**, 16295-16303.

Vogel, H.J. y McLellan, W.L. (1970) N-acetyl-γ-glutamokinase (*Escherichia coli*). *Methods Enzymol.* **17A**, 251-255.

Waldrop, G.L., Rayment, I. y Holden, H.M. (1994) Three-dimensional structure of the biotin carboxylase subunit of acetyl-CoA carboxylase. *Biochemistry* **33**, 10249-56.

Wimmer, M.J., y Rose, I.A. (1978) Mechanisms of enzyme-catalyzed group transfer reactions. *Ann. Rev. Biochem.* **47**, 1031-1078.

Wittinghofer, A. (1997) Signaling mechanistics: aluminum fluoride for molecule of the year. *Curr. Biol.* **7**, R682-5.

Xu, Y.W., Morera, S., Janin, J. y Cherfils, J. (1997) AlF_3 mimics the transition state of protein phosphorylation in the crystal structure of nucleoside diphosphate kinase and MgADP. *Proc. Natl. Acad. Sci USA* **94**, 3579-3583.

Yount, R.G., Babcock, D., Ballantyne, W. y Ojala, D. (1971) Adenylyl imidodiphosphate, an adenosine triphosphate analog containing a P-N-P linkage. *Biochemistry* **10**, 2484-2489.