UNIVERSITEIT GENT

FACULTEIT WETENSCHAPPEN
Vakgroep Biochemie, Fysiologie en Microbiologie
Laboratorium voor Eiwitbiochemie en Eiwitengineering

Structural insights into extremozymes:
a study of *Sulfolobus acidocaldarius* and *Moritella profunda* aspartate carbamoyltransferase, and of *Pseudoalteromonas haloplanktis* xylanase pXyl

Dirk De Vos

Proefschrift voorgelegd tot het bekomen van de graad van
Doctor in de Wetenschappen – richting Biochemie

Promotor: Prof. Dr. J. Van Beeumen
Academiejaar 2005-2006

Word of thanks

I would like to express my sincere gratitude to everyone who was involved directly or indirectly in the realization of this thesis.

In the first place my promoter Prof. Van Beeumen for offering me the opportunity to perform this work at his lab, and for the support whenever needed.

I would also like to thank the FWO-Vlaanderen for the doctoral grant I have benefited from.

I am greatly indebted to Filip Van Petegem for introducing me into the "art" (according to some) of crystallography and for his help and encouragement, even after leaving the lab.

I wish to thank all those at the Laboratory for Protein Biochemistry and Protein Engineering for their support and kindness. Thanks, Kris and Frank for the mass spectrometric analyses, Isabel for the N terminal sequencing, and Wesley for the computer-related help. Savvas, I have appreciated your company, as much as your efforts in making life easier at the crystallography unit. I am indebted to Paco for his eternal willingness to help out on all sorts of occasions. Jan, thanks for the pleasant company and support, for instance during those first stressful expeditions to synchrotron facilities. Bjorn, thank you for your aid, availability and experience. Your criticism has been of great help. I especially think of Bjorn, Jan, Kris and Paco for their friendship. The numerous talks, discussions and laughs were worth enduring hundreds of "croques madames/spéciales" for.

I am most thankful to the people from other groups with whom I had the pleasure of collaborating:
-Thanks to Prof. Glansdorff for the supply of the ATCases.
-Thanks to Prof. Gerday, Prof. Feller and Tony Collins for the collaborations on the xylanase. Thanks Tony, for your enthusiasm and dedication made it a real pleasure to work on this subject.
-Thanks to Prof. Claeyssens and Wim Nerinckx for helping out with the docking studies. Thanks Wim for the lively discussions on glycosidase mechanisms.
-Thanks to Tony Aerts for collaboration on the ultracentrifugation studies.

Thanks Gert for the support and friendship ever since our first year in Ghent.

I am deeply grateful to my family for supporting and encouraging me. I owe a special word of thanks to my sweetheart Ilse for putting up with me, esp. during the last months and for her help with the realization of this thesis. A big kiss also for Kobe for making me smile no matter when.

Ghent, February 2006

Table of contents

Preface

In this thesis a compilation is presented of a number of studies aimed to unravel the structure-function relationships of three enzymes from extremophilic micro-organisms: the aspartate carbamoyltransferases (ATCases) from the hyperthermophilic archaeon *Sulfolobus acidocaldarius* and from the psychrophilic deep-sea bacterium *Moritella profunda*, and a glycoside hydrolase family 8 xylanase (termed pXyl) from the psychrophilic bacterium *Pseudoalteromonas haloplanktis*. What connects these enzymes is that they face the common challenge of optimally performing their function at so-called 'extremes' of life. Therefore, in the first part of this work an attempt is made to describe the influence of the environmental extremes considered relevant to these enzymes and their adaptational strategies towards these conditions. Given their obviously distinct structural, functional and physiological character, the two ATCases are dealt with in a section (Part II) separate from that of the xylanase (Part III). The first chapter of each part consists of a brief introduction that allows the reader to catch up with the following chapters which give a description of the experiments, results and conclusions of the studies, mainly in the form of manuscripts that have been published, are submitted, or are in preparation for submission.

Part I gives a general introduction into the diverse types of extremophiles and defines which environmental extremes are relevant to their proteins. Stability, probably the most intensively studied property in relation to protein adaptation, is discussed in sufficient depth to illustrate our current understanding of molecular adaptation to high and low temperatures. The influence of temperature on enzyme activity and the relationship between activity, flexibility and stability is shortly debated. Finally, the potential role for pressure in enzyme adaptation is discussed.

Part II deals with the ATCase from *S. acidocaldarius* and from *M. profunda*. The first chapter of this section is a short introduction into the physiological background of the enzyme, into the *Escherichia coli* ATCase as an allosteric model system, and gives an overview of extremophilic ATCases charaterized to date. Chapter 4 consists of the paper "Crystal structure of T state aspartate carbamoyltransferase of the hyperthermophilic archaeon *Sulfolobus acidocaldarius*", presents the results of the structural description and molecular analysis of this enzyme, and compares it to the previously determined crystal structures of the *E. coli* enzyme in its main allosteric states and of the *Pyrococcus abyssi* R state catalytic trimer. Chapter 5 is a mainly technical paper entitled: "Expression, purification, crystallization and preliminary X-ray crystallographic studies of a cold-adapted aspartate carbamoyltransferase from *Moritella profunda*". Chapter 6 is a manuscript in preparation entitled "Structural basis for cold regulation: biochemical properties and crystal structure of aspartate carbamoyltransferase from the psychrophilic bacterium *Moritella profunda*", and presents the results of the functional characterization, structure determination and molecular analysis of the recombinant *M. profunda* ATCase. Chapter 7 provides a description of supplementary experiments on the allosteric regulation of the *S. acidocaldarius* and *M. profunda* ATCases. Preliminary results on the crystal structures of effector complexes and ultracentrifugation (sedimentation velocity) experiments probing the quaternary structural changes of both ATCases are presented within.

Part III deals with a glycoside hydrolase family 8 xylanase, pXyl, from *P. haloplanktis*. Chapter 8 provides a short introduction in xylanases and their substrate xylan, and describes the functional and structural characteristics of this cold-adapted enzyme. Chapter 9 consists of the recently published research paper: "Study of the active site residues of a glycoside hydrolase family 8 xylanase". Whereas the crystal structure of the native enzyme was previously determined at our lab, a detailed structural analysis of several active site mutants is now presented in combination with their kinetic and biophysical properties. The principal aim of this study was to determine the identity of the catalytic proton donor and proton acceptor. Finally, chapter 10 consists of a manuscript, submitted for publication, entitled: "Oligosaccharide binding in family 8 glycosidases: crystal structures of active site mutants of the β-1,4-xylanase pXyl from *Pseudoalteromonas haloplanktis* in complex with substrate and product". This paper further explores the catalytic mechanism and substrate specificity of pXyl.

At the end of this thesis, a general conclusion and proposals for future work on these topics will be formulated.

Voorwoord

In dit proefschrift worden een aantal studies voorgesteld met als doel het ontrafelen van de structuur-functie verwantschap van drie enzymen van extremofiele micro-organismen: de aspartate-carbamoyltransferases (ATCasen) van de hyperthermofiele archaebacterie *Sulfolobus acidocaldarius* en de psychrofiele diepzeebacterie *Moritella profunda*, en een xylanase van de glycosidehydrolase-familie 8 (genaamd pXyl) van de psychrofiele bacterie *Pseudoalteromonas haloplanktis*. Het verband tussen deze enzymen is de gemeenschappelijke uitdaging optimaal te functioneren in zogenaamde 'extreme' leefomstandigheden. Daarom wordt in het eerste deel van dit werk gepoogd de invloed van extreme leefomstandigheden die relevant zijn voor deze enzymen te beschrijven, alsook hun strategieën om zich aan te passen aan deze omstandigheden. Gezien hun duidelijke verschillen in structuur, functie en fysiologische rol, worden de twee ATCasen in een apart onderdeel (Part II) behandeld, naast het xylanase (Part III). Het eerste hoofdstuk van beide delen is opgevat als een korte inleiding tot de daarop volgende hoofdstukken, die een beschrijving geven van de experimenten, resultaten en besluiten van het onderzoek. Dit gebeurt voornamelijk onder de vorm van manuscripten die reeds gepubliceerd zijn, of ingediend voor publicatie, of in voorbereiding.

Deel I geeft een algemene inleiding in de verschillende types extremofielen en definieert de types van extreme leefomstandigheden die relevant zijn voor hun eiwitten. Stabiliteit, waarschijnlijk de meest intensief bestudeerde eigenschap in verband met eiwitadaptatie, wordt grondiger besproken om ons huidig begrip van moleculaire aanpassing aan hoge en lage temperaturen te duiden. De invloed van temperatuur op enzymactiviteit en het verband tussen activiteit, flexibiliteit en stabiliteit wordt kort besproken. Tenslotte wordt de potentiële rol van druk in enzymadaptatie toegelicht.

Deel II bespreekt de ATCasen van *S. acidocaldarius* en *M. profunda.* Het eerste hoofdstuk van dit onderdeel (Hoofdstuk 3) is een korte inleiding in de fysiologische achtergrond van het enzym, in het *Escherichia coli* ATCase als allosterisch modelsysteem, en geeft een overzicht van de extremofiele ATCasen die tot op heden gekarakteriseerd zijn. Hoofdstuk 4 bestaat uit de publicatie "Crystal structure of T state aspartate carbamoyltransferase of the hyperthermophilic archaeon *Sulfolobus acidocaldarius*", en stelt de resultaten van de structurele beschrijving en moleculaire analyse van dit enzym voor, naast een vergelijking met de eerder bepaalde kristalstructuur van het *E. coli* enzyme in zijn belangrijkste allosterische toestanden en met de R toestand van de katalytische trimeer van *Pyrococcus abyssi*. Hoofdstuk 5 bestaat uit een publicatie van technische aard met als titel: "Expression, purification, crystallization and preliminary X-ray crystallographic studies of a cold-adapted aspartate carbamoyltransferase from *Moritella profunda*". Hoofdstuk 6 is een manuscript in voorbereiding, getiteld "Structural basis for cold regulation: biochemical properties and crystal structure of aspartate carbamoyltransferase from the psychrophilic bacterium *Moritella profunda*", en geeft de resultaten van de functionele karakterisering, structuurbepaling en moleculaire analyse van het recombinant *M. profunda* ATCase. Hoofdstuk 7 bevat een beschrijving van aanvullende experimenten met betrekking tot de allosterische regulatie van de *S. acidocaldarius* en *M. profunda* ATCasen. Hierin worden ook voorlopige onderzoeksresultaten voorgesteld i.v.m.

kristalstructuren van effectorcomplexen en ultracentrifugatie- (sedimentatiesnelheids-) experimenten die de quaternaire structuur-veranderingen van beide ATCasen nagaan.

Deel III handelt over een glycosidehydrolase-familie 8 xylanase, pXyl, van *P. haloplanktis*. Hoofdstuk 8 geeft een korte inleiding over xylanases en hun substraat, xylaan, en beschrijft de functionele en structurele karakteristieken van dit enzym aangepast aan koude omstandigheden. Hoofdstuk 9 omvat een recent gepubliceerd artikel: "Study of the active site residues of a glycoside hydrolase family 8 xylanase". Hoewel de kristalstructuur van het natief enzym reeds bepaald werd op ons labo, wordt nu een gedetailleerde structurele analyse voorgesteld van verschillende mutanten van het actief centrum, gecombineerd met hun kinetische en biofysische eigenschappen. Het voornaamste doel van dit onderzoek was het bepalen van de identiteit van de katalytische proton donor en proton acceptor. Hoofdstuk 10 tenslotte omvat een manuscript, ingediend voor publicatie en getiteld: "Oligosaccharide binding in family 8 glycosidases: crystal structures of active site mutants of the β-1,4-xylanase pXyl from *Pseudoalteromonas haloplanktis* in complex with substrate and product". Dit artikel gaat dieper in op het katalytisch mechanisme en de substraatspecificiteit van pXyl.

Op het einde van dit proefschrift formuleren we een algemeen besluit en doen we voorstellen voor toekomstig onderzoek.

Part I
Introduction

1 Extremophiles

An extremophile (terminology from Macelroy, 1974) is an organism that thrives in an extreme environment. 'Extremes' typically include physical extremes (for example, temperature, pressure or radiation) and geochemical extremes (for example: salinity, pH, dessication, oxygen species or redox potential). Yet, it could be argued that extremophiles should include organisms thriving in biological extremes (for example, nutritional extremes or extremes of population density). What one classifies as 'extreme' inevitably requires a definition of 'normal' which in this case perhaps tends to be anthropocentric. While intuitively acceptable, it presents a potential pitfall, i.e. the misconception that these organisms have had to develop adaptive strategies to colonize their extreme habitats, implying a derived character. Indeed, from an evolutionary perspective it may well be that some of these environmental conditions are more similar to those found on primitive Earth. E.g. throughout most of the early history of life the Earth has been anaerobic. Until relatively recently a high-temperature origin of life was widely favoured because hyperthermophiles (maximum growth above 80 °C) are claimed to be the oldest organisms on the Earth (Levy and Miller, 1998; Xu and Glansdorff, 2002). Considerations like these have to be kept in mind, particularly when trying to recognize molecular mechanisms of adaptation.

The focus of this work is the study of enzymes originating from three extremophilic micro-organisms. A glycoside hydrolase family 8 (GH-8) xylanase from *Pseudoaltermonas haloplanktis*, and the aspartate carbamoyltransferases (ATCases) from *Moritella profunda* and *Sulfolobus acidocaldarius*. Firstly, *P. haloplanktis*, isolated from the Antarctic, and *M. profunda*, isolated from deep Atlantic sediments, are both Gram-negative psychrophilic ('cold-loving'), moderately halophilic ('salt loving') bacteria. *M. profunda*, furthermore, is a piezophile, meaning it thrives at high pressures. *S. acidocaldarius* is a hyperthermophilic, acidophilic (low pH-loving) archaeon, originally isolated from a sulphur-rich hot spring in Yellowstone National Park, USA. All three organisms thrive in multiple 'extreme' conditions and can therefore be called 'polyextremophiles'.

Temperature creates a series of challenges, from the structural damage brought about by ice crystals at one extreme, to the denaturation of biomolecules at the other. The solubility of gasses is correlated with temperature, creating problems at high temperature for aquatic organisms requiring O_2 or CO_2. Temperatures approaching 100 °C normally denature proteins and nucleic acids, and increase the fluidity of membranes to lethal levels. Chlorophyl degrades above 75 °C, excluding photosynthesis. Yet, in nature, thermal preferences range from hyperthermophilic (maximum growth >80 °C) to psychrophilic (maximum growth <15 °C). The most hyperthermophilic organisms are archaea, with *Pyrolobus fumarii* (Crenarchaeota) being capable of growing at the highest temperatures of up to 113 °C (Blochl *et al.*, 1997). Representatives of all major taxa inhabit temperatures just below 0 °C. Many microbes and cell lines can be preserved successfully at –196 °C (liquid nitrogen), but the lowest recorded temperature for active microbial communities is substantially higher, at –18 °C (Clarke, 2003). At low temperatures water freezes. The resulting ice crystals can rip cell membranes, and solution chemistry stops in the absence of liquid water.

Pressure challenges life because it forces volume changes. Pressure compresses the packing of lipids, resulting in decreased membrane fluidity (Bartlett and Bidle, 1999). If a chemical reaction results in an increase in volume, as often is the case, it will be inhibited by an increase in pressure (Van Dover, 2000). High pressure can damage DNA and proteins in particular, so survival necessitates avoidance of damage or high repair rates. The Mariana trench is the world's deepest sea floor at 10,898 m, yet it harbours organisms that can grow at standard temperature and pressure. It has also yielded obligatory piezophilic species that can grow at 70 to 80 MPa (~700–800 times atmospheric pressure), but not below 50 MPa (Kato *et al.*, 1998).

Organisms live within a range of salinities, from essentially distilled water to saturated salt conditions. Osmophily refers to the osmotic aspects of life at high salt concentrations, especially turgor pressure, cellular dehydration and dessication. Halophily refers to the ionic requirements for life at high salt concentrations. Although these phenomena are physiologically distinct, they are environmentally linked. Thus a halophile must cope with osmotic stress.

Acidophiles thrive at low pH. The archaeal iron-oxidizing *Ferroplasma acidarmanus* has been described growing at pH 0 in acid mine drainage, thriving in a brew of sulphuric acid and high levels of copper, arsenic, cadmium and zinc, with only a cell membrane and no cell wall (Edwards *et al.,* 2000). Indeed, probably the most critical factor for acidophily is the cytoplasmic membrane, as most of these organisms maintain their cytoplasm near neutrality (Rothschild and Mancinelli, 2001) in order to prevent destruction of acid-labile macromolecules of the cell. Consequently, when studying the aspartate carbamoyltransferase (ATCase) from *S. acidocaldarius* adaptation to low pH will, in principle, not have to be considered as it is an intracellular enzyme.

2 Proteins at extremes

While realizing that other cell components (membrane lipids, DNA, …) need to be equally well adapated to life in these conditions, we will focus our attention on a single, yet essential, cellular constituent of the types of extremophiles that are relevant to this study, i.e. their proteins and more specifically their enzymes. Like proteins in general, enzymes have a three-dimensional fold that is only marginally stable and is, among other factors, susceptible to changes in temperature, pressure and salinity. Rather than being a single static rigid structure, all the atoms in the folded state are subject to small temperature-dependent fluctuations. The molecule as a whole undergoes breathing and every atom is constantly in motion. In addition to these small breathing movements there can also be larger conformational changes which are usually essential for function (e.g. in enzymes). Stability, hence, is only one aspect of protein adaptation. Enzymes are presented with an extra challenge. Firstly, they have to achieve appropriate values of catalytic activity. Due to the complexity of any catalytic activity, the particular influence of each rate constant on the reaction velocity depends in turn on the values of the rest of the constants, through an equally complicated relationship. Therefore the achievement of appropriate values of enzyme activity is a question of multivariable optimization (Melendez-Hevia *et al.*, 1994). Most frequently, the role of enzyme activity is the maintenance of a steady flux of products (Cornish-Bowden, 1976). In other cases (e.g. ATCase) the enzyme plays a more central role in the control or regulation of a metabolic

pathway, and therefore its activity is intricately regulated by the presence or balance of one or more effector molecules which trigger the switching between the different states that make up an allosteric enzyme system. One thereby has to bear in mind that adaptive changes of an enzyme are always context-dependent in that any enzyme is just a tool of one or more metabolic pathways (Dykhuizen *et al.*, 1987; Fell, 1997).

2.1 Temperature dependence of protein stability

Stability is probably the most well-studied aspect of enzyme adaptation to extreme conditions, more specifically in relation to extremes of temperature. Protein stability curves describe the temperature-dependent variation of protein stability, i.e. the Gibbs free energy of unfolding (ΔG = G_u(unfolded or denatured state) - G_f(native folded state)) in function of the absolute temperature (T). For a protein or a protein domain which (i) folds in a reversible two-state manner, (ii) is stable over a temperature range, and (iii) has a constant (greater than zero) heat capacity change (ΔC_p) in this range (which was established as a valid assumption by Privalov and Khechinashvii (1974)), the following equation can be used to plot its stability

$$\Delta G = \Delta H - T\,\Delta S = \Delta H^{\,0} - T\,\Delta S^{\,0} + \Delta C_p\,[T - T^{\,0} - T\ln(T / T^{\,0})]\ ,$$

where $T^{\,0}$ is any reference temperature and $\Delta S^{\,0}$ and $\Delta H^{\,0}$ are the entropy and enthalpy changes at that temperature, respectively. The three parameters ΔC_p, $\Delta S^{\,0}$ and $\Delta H^{\,0}$ in this equation are sufficient to establish the course of the free energy over de temperature range for which the third postulate is valid. It follows that, for all of the proteins that follow a two-state transition, there are two transition temperatures where $\Delta G(T) = 0$. These are T_G and T'_G, termed the heat- and cold-denaturation transition temperatures, respectively.

To microscopically understand the macroscopic parameters of protein stability Makhatadze and Privalov (1995), in a highly authoritative paper, dissect the free energy of unfolding into separate enthalpic and entropic contributions from nonpolar (including aliphatic and aromatic amino acid residues, denoted with subscript ‘npl’) and polar groups (polar residues, denoted with ‘pol’). A further distinction is made between (i) net hydration effects (denoted with ‘hyd’) resulting from the transfer of a molecule from a fixed position in an ideal gas phase into a fixed position in (liquid) water, (ii) internal noncovalent interactions, and (iii) a conformational entropy term (denoted with ‘cnf’). The internal noncovalent interactions, contributing to the enthalpy of unfolding in vacuum, are all electrostatic in nature and present a combination of Coulombic, dipole, quadrupole, etc. interactions (Sharp and Honig, 1990). Nevertheless, in considering the interactions in proteins, it is convenient to classify them as salt links, hydrogen bonds, weak polar, and van der Waals interactions. Salt links (superscript ‘SL’) includes the Coulombic interactions between closely located opposite-charged groups (Barlow and Thornton, 1983). Hydrogen bonds (superscript ‘HB’) include the interactions between electronegative atoms involving hydrogen (Baker and Hubbard, 1984). The interactions between the slightly polar aromatic groups are classified as weak polar interactions (Burley and Petsko, 1988; superscript WP). Van der Waals interactions (superscript ‘vdW’) are the London dispersion forces between induced dipoles. Because dispersion forces are proportional to the polarizability of groups, van der Waals interactions must be especially large between aliphatic groups (Makhatadze and Privalov, 1995).

The overall enthalpy of protein unfolding in water can then be represented as follows:

$$\Delta H = (\Delta H^{HB} + \Delta H^{SL} + \Delta H^{hyd})_{pol} + (\Delta H^{vdW} + \Delta H^{WP} + \Delta H^{hyd})_{npl}.$$

The enthalpy of hydration proceeds from direct noncovalent interactions with water, and also from the rearrangement of the hydrogen-bonding network of water. As temperature increases, the van der Waals (and weak polar) interactions between nonpolar groups do not change much, but the enthalpy of hydration of these groups increases (becomes less favourable). Therefore, at high temperatures, the enthalpic effect of interactions between nonpolar groups in the interior of a native protein overbalances the hydration enthalpy of these groups on protein unfolding. The positive enthalpy of interaction between polar groups in a protein does not depend significantly on temperature, while the negative hydration enthalpy of these groups increases in magnitude linearly with increasing temperature. Therefore, above a given temperature, the hydration enthalpy of polar groups overcompensates the enthalpy of hydrogen bonding between polar groups. The stronger interaction of polar groups with water than with each other is in accord with the observation that the partial volumes of polar side chains are larger in proteins than in water (Harpaz *et al.*, 1994).

The overall entropy of protein unfolding in water is comprised of three terms: the conformational (also called configurational) entropy, ΔS^{cnf}, the entropy of hydration for polar groups, $(\Delta S^{hyd})_{pol}$, and the entropy of hydration for nonpolar groups, $(\Delta S^{hyd})_{npl}$.

$$\Delta S = \Delta S^{cnf} + (\Delta S^{hyd})_{pol} + (\Delta S^{hyd})_{npl}.$$

Calculation of these terms (per mole of amino acid residues) for several model proteins shows that the negative hydration entropies of nonpolar groups converge to zero at about 125 °C (Makhatadze and Privalov, 1995). The negative hydration entropies of polar groups slightly diverge and increase in magnitude with increasing temperature. Despite many doubts (Dill, 1990; Williams, 1991), it appears that the contribution of hydrogen bonds to the stabilization of protein structure is significant, due (to a large extent) to this negative entropy of hydration of polar groups (Privalov and Makhatadze, 1993). According to Shirley *et al.* (1992) who studied ribonuclease T1 mutants, and Pace *et al.* (1996) who analyzed hydrogen bonding mutants in a number of proteins, the contribution of the hydrogen bond to stabilization of protein structure is between 3 and 7 kJ/mol (i.e. comparable to hydrophobic interactions). This contrasts with the traditional view of Kauzmann (1959) who regards hydrophobic interactions as the primary source of stability and hydrogen bonding as the source of specificity. The positive (favourable for unfolding) conformational entropies do not change much with an increase in temperature (Makhatadze and Privalov, 1995).
In considering the hydrophobic effect in proteins, the entropy of association of nonpolar groups on protein folding is usually assigned to the conformational entropy of the polypeptide chain. Therefore, the effect which we can regard as a hydrophobic interaction (denoted with 'HPH') in proteins should include only the enthalpy of the van der Waals interactions between nonpolar groups and the Gibbs energy of hydration of these groups:

$$\Delta G^{HPH} = \Delta G^{hyd} + \Delta H^{vdW}.$$

At room temperature the enthalpy of hydration of nonpolar groups becomes equal in magnitude and opposite in sign to the enthalpy of van der Waals interactions between these groups. These two enthalpic effects compensate each other, and it is only the entropy of hydration of nonpolar groups which contributes to the Gibbs energy of their transfer into water, i.e., to the magnitude of the hydrophobic effect at this temperature. At high temperature (around 75 °C), the entropy of hydration of nonpolar groups decreases to zero and Gibbs energy of the hydrophobic effect becomes completely enthalpic (Baldwin, 1986; Privalov and Gill, 1988). The dual character of the hydrophobic effect is nicely illustrated in a study by Eriksson *et al.* (1992) on six "cavity-creating" mutants in T4 lysozyme. This study demonstrates that the energy, associated with the replacement of a buried nonpolar residue by a less bulky nonpolar residue consists of two parts. The first part is constant and presumed to depend only on the identities of the two amino acids being compared. Physically, this can be considered as the difference in hydration energy. The second part of the change in protein stability depends on the context within the three-dimensional structure and the way in which the protein adjusts in response to the substitution, which can be considered as resulting from the removal of favourable van der Waals contacts.

From the less traditional (and more convenient) viewpoint of separating hydration effects and intramolecular interactions, Makhatadze and Privalov (1995) conclude that in spite of the positive contribution of hydration of nonpolar groups (aromatic not included according to Makhatadze and Privalov, 1995), hydration effects are destabilizing the compact native state, and this effect increases with decreasing temperature. The destabilizing action also comes from a thermal dissipative force, which is proportional to the gain of conformational entropy on protein unfolding and the absolute temperature – $T\Delta S^{cnf}$. This negative force thus decreases with temperature. The destabilizing effects are then counterbalanced by enthalpic interactions between the groups tightly packed in the protein interior, which do not depend significantly on temperature.

Kumar *et al.* (2001) propose a two-state molecular model of the water structure which is useful to understand the temperature dependence of protein stability. In such a model, water consists of different dynamically transforming intermolecular hydrogen-bonding types (Vedamuthu *et al.*, 1994; Robinson and Cho, 1999). The first one is the enthalpically favoured, "normal" icelike, tetrahedrally connected hexagonal hydrogen bonding type, with optimal hydrogen-bonding networks. The second is the entropically favourable, enthalpically unfavourable, highly fluctuating liquid form. At high temperatures, the denser liquid types dominate, with a fluctuating gradient of interconverting hydrogen-bonding types from low to high temperatures. The water structure is dynamic, with the hydrogen-bonds continuously broken and created on a very short time scale (Ruelle and Kesselring, 1998).

Below the cold-denaturation temperature ($T < T'_G$), the fraction of normal hexagonal ice in solvent water increases and the protein denatures because of the loss of the hydrophobic effect. In the denatured state of the protein, the exposure of the nonpolar surface promotes optimally-ordered hydrogen-bond formation at the first shell of the water molecules, propagating to the outer shells. The entropy contribution to the Gibbs free energy change is negative. However, this effect is overcome by the optimized hydrogen-bond networks resulting in a reduced enthalpy.

At temperatures above the protein heat-denaturation temperature ($T > T_G$), the protein denatures because of an increase in the entropy of the system. Liquid water dominates, with favourable entropy and unfavourable enthalpy terms. Order cannot propagate in highly fluctuating water molecules. The increase in the entropy of the system overcomes the enthalpically unfavourable exposure of the nonpolar surface area of the protein, and the denatured state of the protein is energetically favourable. While giving an accurate description of the hydration effects of nonpolar groups, it ignores the role of the growth in hydration of polar groups with decreasing temperature (Makhatadze and Privalov, 1995). The main difficulty appears to be the difference in temperature dependencies of the hydration enthalpy and entropy of polar groups compared to nonpolar groups (first noted by Edsall, 1935), or in other words: why do the heat capacity increments of hydration of these groups have opposite signs. This finding suggests that the presence of aliphatic groups in water intensifies its ice-like structure, while the polar groups reorganize the water structure. Apparently this is achieved more readily at higher temperatures for polar groups.

2.2 Adaptation strategies of (hyper)thermophilic proteins

After the discovery of the first hyperthermophilic bacterium (Brock *et al.*, 1972), studies of proteins extracted from these microorganisms showed that they retain their thermal resistance after purification, suggesting that this property is intrinsic to protein structure, independent of the presence of some particular compounds produced in the microorganism (although some stabilizing factors as polyamines and ectoines, so called intrinsic factors, have been found occasionally in some extremophiles, Lippert and Galinski, 1992). A further confirmation of the intrinsic thermostability of the proteins from extremophiles was obtained by the production of a thermoactive and thermostable recombinant protein in a mesophilic host. Thermotolerance was thus encoded in the genetic blueprint and (as the same 20 canonical are found in these organisms) could be ascribed to the presence of a different and/or increased number of weak interactions that are essential for the stability of the native three-dimensional structure (Tanaka *et al.*, 1981; Jaenicke, 1991).

The high level of similarity encountered in the core of mesophilic and hyperthermophilic protein homologues suggests that even mesophilic proteins are packed almost as efficiently as possible, and that there is not much room left for stabilization inside the protein core. Stabilizing interactions are often found in the less conserved areas of the protein. Enough experimental evidence (e.g. sequence, mutagenesis, structure, and thermodynamics) has been accumulated on hyperthermophilic proteins to conclude that no single mechanism is responsible for their remarkable stability. Increased stability must be found, instead, in a small number of highly specific mutations that often do not obey any obvious traffic rules (Vieille and Zeikus, 2001).

Protein amino acid composition has long been thought to be correlated to its thermostability. The first statistical analyses comparing amino acid compositions in mesophilic and thermophilic proteins indicated a trend towards substitutions such as Gly → Ala and Lys → Arg. A higher alanine content in thermophilic proteins was supposed to reflect the fact that Ala is the best helix-forming residue (Argos *et al.*, 1979). Nevertheless, the comparison of residue contents in hyperthermophilic and mesophilic proteins based on genome sequences of mesophilic and hyperthermophilic organisms shows some general

trends. More charged residues are found in hyperthermophilic proteins than in mesophilic proteins, mostly at the expense of uncharged polar residues (Cambillau and Claverie, 2000). Hyperthermophilic proteins also contain slightly more hydrophobic and aromatic residues than mesophilic proteins do (Vieille and Zeikus, 2001). These data obtained from genome sequencing cannot be generalized, since large variations exist among hyperthermophilic genomes themselves. Thus, a bias in a hyperthermophilic protein amino acid composition might often be evolutionary relevant, rather than being an indication of its adaptation to high temperatures. Indeed, a recent analysis of structures of several hyperthermostable proteins from various sources reveals two major physical mechanisms of thermostabilization. The first mechanism is "structure-based", whereby some hyperthermostable proteins (namely hyperthermophilic archaea) are more compact than their mesophilic homologues, while no particular interaction type appears to cause stabilization. Organisms that evolved as mesophiles but later recolonized a hot environment (*in casu Thermotoga maritima*) are suggested to employ an alternative, "sequence-based" mechanism (Berezovsky and Shakhnovich, 2005).

Probably more relevant to thermostability are the distribution of the residues and their interactions in the protein (Vieille and Zeikus, 2001). In relation to the idea that protein stability is determined by the stability and tight packing of its core, the propensity of the individual residues to participate in helical or strand structures was studied as a potential stability mechanism. The only trend that was detected by Facchiano *et al.* (1998) was a decreasing content in (destabilizing) β-branched residues (Val, Ile and Thr) in the helices of thermophilic proteins. On the other hand, according to Dill (1990), Ile might be favoured for its ability to better fill various voids that occur during protein core packing (it has more favourable rotamers than its isomer leucine). Several properties of Arg residues suggest that they would be better adapted to high temperatures than Lys residues: the Arg δ-guanidino moiety has a reduced chemical reactivity due to its high pK_a an its resonance stabilization. Arg, furthermore, has more surface area for charged interactions than Lys. The average Arg/Lys ratios in the protein pools of the mesophiles and hyperthermophiles are associated with large standard deviations (Vieille and Zeikus, 2001). If an increased Arg is indeed stabilizing, this mechanism is not universally used among hyperthermophiles. A decreased number of thermolabile residues, e.g. Asn and Gln which are susceptible to deamidation, or Cys and Met which are sensitive to oxidation, has also been reported as a possible adaptation strategy (Russell *et al.*, 1997). Matthews *et al.* (1987) proposed that proteins of known three-dimensional structure can be stabilized by decreasing their entropy of unfolding. In the unfolded state, glycine is the residue with the highest conformational entropy. Proline, which can adopt only a few conformations and restricts the possible conformations of the preceding residue (Sriprapundh *et al.*, 2000), has the lowest conformational entropy. This technique has been used to engineer enzymes that are more thermodynamically stable (Eijsink *et al.*, 1993), and a number of thermophilic and hyperthermophilic proteins also use this stabilization mechanism (Nakai *et al.*, 1999). For the same reason disulfide bridges are proposed to be thermostabilizing (Matsumura *et al.*, 1989). Interestingly, substitution of residues (in a left-handed helical conformation) with Gly has been demonstrated to increase thermodynamic stabilities of two proteins by conformational strain release (Kimura *et al.*, 1992; Kawamura *et al.*, 1996).

Although it was concluded by several studies (Karshikoff and Ladenstein, 1998; Kumar *et al.*, 2000) that packing and fractional polar and nonpolar surface areas are not common thermostabilization mechanisms, a few examples exist of hyperthermophilic

proteins that gain part of their stability from better packing (Chan *et al.*, 1995; Russell *et al.*, 1997). Thermodynamic stabilization was demonstrated experimentally for *Methanobacterium formicicum* histone by filling of a solvent-accessible cavity by bulkier hydrophobic side chains (Li *et al.*, 1998).Several examples exist of a reduced hydrophobic accessible surface areas (Auerbach *et al.*, 1998; Grabarse *et al.*, 1999) but, more frequently, extra hydrophobic interactions at subunit interfaces of oligomers are reported in structural studies. Kirono *et al.* (1994) furthermore demonstrated with *Thermus thermophilus* 3-isopropylmalate dehydrogenase mutants that extra (compared to the *E. coli* homologue) hydrophobic interactions make the dimer more resistant to dissociation.

Two types of noncovalent interactions involving aromatic residues can be distinguished: aromatic-aromatic interactions (aromatic pairs) defined by a distance of less than 7 Å, and cation-π interactions between positive charges (e.g. metal cations or cationic side chains of Arg and Lys) and an aromatic ring center. Both types are electrostatic in nature but are potentially stabilizing because of the lower desolvation penalty for aromatic residues (Burley and Petsko, 1985; Dougherty, 1996). Some examples of the first type are suggested to play a role in stabilizing thermophilic and hyperthermophilic proteins (Ishikawa *et al.*, 1993; Teplyakov *et al.*, 1990). The second type has not been intensively studied in relation to thermostability (Massant).

As mentioned earlier, the role of hydrogen bonds in the stability of proteins has always been controversial (Dill, 1990; Pace *et al.*, 1996; Privalov). Because the identification of H bonds is highly dependent on the distance cutoff and because a number of hyperthermophilic protein structures have not been refined to sufficiently high resolutions, studying the role of H bonds in thermostability by comparative structure analysis has not provided clear-cut answers. An interesting discrepancy is found between the results of the large-scale statistical studies of Vogt *et al.* (Vogt and Argos, 1997; Vogt *et al.*, 1997) and those of Szilagyi and Zavodszky (2000). The former found hydrogen bonds to be the most important stabilizing factor, whereas the latter did not found a significant difference in hydrogen bonds between mesophilic and thermophilic proteins, possibly because of the better balance in quality of the mesophilic and thermophilic datasets (Szilagyi and Zavodszky, 2000). Finally, one study (Tanner *et al.*, 1996) propagates the potentially stronger stabilizing role of charged-neutral hydrogen bonds due to (i) the lower desolvation penalty in comparison with ion pairs, and (ii) the greater enthalpic reward compared to neutral-neutral hydrogen bonds.

Because ion pairs are usually present in small numbers in proteins and because they are not highly conserved, they seem less important in protein folding (Dill, 1990; Makhatadze and Privalov (1995) do not distinguish salt links from other polar groups in their study of protein stability). However, the most consistent trend that comes out of individual and systematic studies is the increase in the number of ion pairs (Vogt *et al.*, 1997; Kumar *et al.*, 2000, Szilagyi and Zavodszky, 2000). On the other hand, some theoretical and a number of experimental studies have indicated that salt bridges usually destabilize, or at most only slightly stabilize the native state of proteins (Sali *et al.*, 1991; Hendsch and Tidor, 1994; Waldburger *et al.*, 1995). Most of these studies, however, were made at room temperature. The major reason for the low stability of ion pairs at room temperature is the large desolvation penalty upon association, which is not fully compensated for by the electrostatic energy of the ion pair. In a theoretical study, Elcock (1998) shows that, as the temperature increases, the unfavourable change in the electrostatic solvation of the separated molecules results in a lowering of the desolvation

penalty for bringing the two together. This effect is only partially compensated by the hydration free energy of the salt bridge itself, being adversely affected by the increase in temperature. This effect can be partly ascribed to a decrease of the water dielectric constant (from 80.20 at 0°C to 55.12 at 100 °C) in his solvation model. Apart from demonstrating the potential role of electrostatic interactions in the thermodynamic stability of hyperthermophilic proteins, the model of Elcock suggests that a sizeable energetic barrier exists for breaking a salt bridge, and the level of this barrier increases with temperature. A similar barrier is not seen with hydrophobic isosteres. The presence of this barrier may go some way toward explaining the apparent role of salt bridges in increasing the kinetic barrier to unfolding (Pappenberger *et al.*, 1997). An interesting observation is that the calculated stability curves for the hyperthermophilic model proteins studied by Elcock differ from the curves of their mesophilic counterparts in that they are mainly shifted along the temperature axis towards higher temperatures. This corresponds to one of the three major ways of promoting thermal stability as proposed by Nojima *et al.* (1977), which are illustrated in figure 2.1 (in this case curve c).

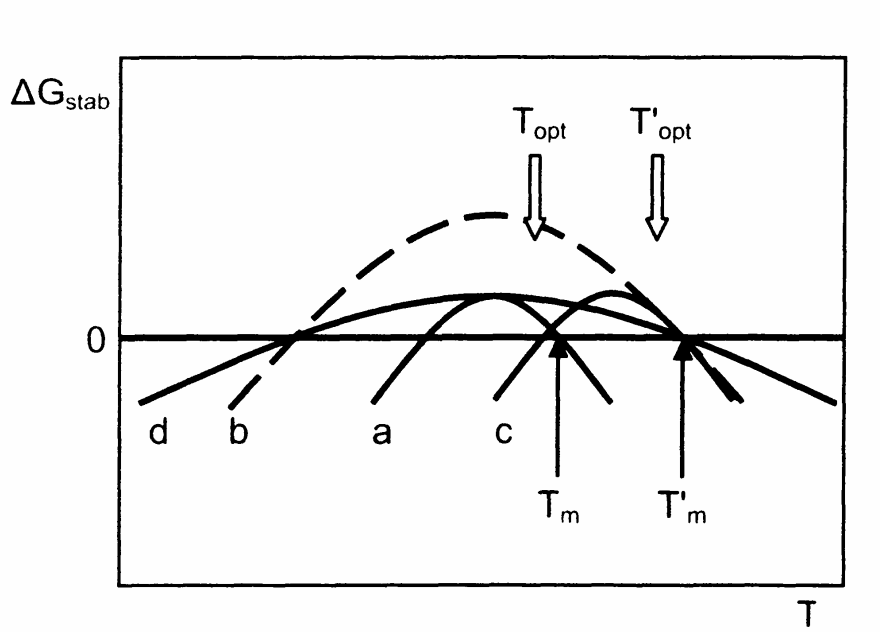


Figure 2.1 Hypothetical temperature profile of the free energy of stabilization (ΔG_{stab}) for (a) mesophilic and (b-d) thermophilic proteins. ΔG_{stab} is defined as the difference in the free energies between the native and denatured proteins. T_m and T'_m are the melting temperatures (corresponds to the heat-denaturation transition temperature T_G) of the mesophilic and thermophilic variants, respectively. The maximum stability for a given protein is observed at a temperature that is much below the optimal growth temperature (T_{opt} and T'_{opt}) of the respective mesophilic or thermophilic organism. Adapted from Jaenicke and Böhm (1998) and Jaenicke (2000).

Although the assumptions involved in this approach are too extensive to illustrate a general principle, it clearly illustrates that if ion pairs are the dominant factor in the stability of hyperthermophilic proteins, one might expect destabilization of the native state at mesophilic temperatures. In a comparative study of thermophilic/mesophilic (T/M) proteins within different protein families, Kumar *et al.* (2001) found that greater protein stability and resistance to higher temperatures are generally obtained by an upshift and

broadening of their stability curves (corresponding with curves b and d of Figure 2.1). Indeed, regardless of the heat transition temperatures (T_G), the temperature of maximal stability (T_S) falls frequently around room temperature and, in most cases, the estimated cold-denaturation temperatures are also lower for the thermophilic proteins. An upshift of the curve would be the outcome of a larger ΔH_G (responsible for the slope of the curve at T_G) resulting from specific additional noncovalent interactions. If the curvature (specified largely by the heat capacity of unfolding ΔC_p) is smaller, the curve would be broader, leading to the third way of increasing thermostability. While it has been recognized that a reduced ΔC_p may represent a common mechanism by which thermophilic proteins achieve thermostability, the origin of this phenomenon is still controversial. It is well established that hydration of the hydrophobic core of a protein upon unfolding contributes to ΔC_p (Privalov and Makhatadze, 1990; Makhatadze and Privalov, 1990). Various equations have been developed to give an estimate of ΔC_p based on the change of accessible-surface area upon unfolding (ΔASA) (Murphy and Freire, 1992; Myers *et al.*, 1995). However, these equations fail to account for the big differences in ΔC_p values observed for T/M pairs of proteins, which are very similar in structure, and are expected to have similar ΔASA values (Hollien and Marqusee, 1999; Motono *et al.*, 2001). A large positive value is taken to indicate the dominance of hydrophobic interactions in driving protein folding, because of the well-kown fact that exposure of nonpolar compounds to water also gives rise to a large positive ΔC_p (Baldwin, 1986; Livingstone *et al.*, 1991; Makhatadze and Privalov, 1995). Based on heat capacity data for transferring model compounds to water, it was also contended that the exposure of polar groups to water gives rise to a negative ΔC_p (Spolar *et al.*, 1992; Murphy and Freire, 1992; Makhatadze and Privalov, 1995). Theoretical calculations based on a simple electrostatic model predict that favourable electrostatic interactions should reduce the value of ΔC_p (Zhou, 2002). As more favourable electrostatic interactions and reduced ΔC_p are both common in thermophilic proteins, it is likely that the reduced ΔC_p is contributed by electrostatic interactions. In a recent study Lee *et al*. (2005) have tested this hypothesis by creating charge-to-neutral mutants of *Thermococcus celer* ribosomal protein L30e. Their results demonstrate that charge-to-neutral mutants (having disrupted electrostatic interactions) that destabilize the protein have increased ΔC_p values. However, the technical difficulty in the accurate measurement of ΔC_p and the relative scarcity of data available on protein thermodynamics have to be noted.

A particularly important role appears to be reserved for ion pairs that participate in networks often prominent at the surface or partially buried at intersubunit interfaces (Lebbink *et al.*, 1999; Rahman *et al.*, 1998; Van Boxstael *et al.*, 2003). Compared to isolated ion pairs, the desolvation cost for each new pair is cut in half: only one additional residue must be desolvated and immobilized. An extremely large ion pair network (composed of 24 residues connected by 18 ion pairs) has been shown to be important for thermostability of *Pyrococcus furiosus* glutamate dehydrogenase (Yip *et al.*, 1995; 1998). However, one has to keep in mind that local environment and geometrical arrangements of interaction partners define the potential for stabilization. Optimization of these parameters, independent from simply increasing numbers of interactions, may even be sufficient to significantly increase protein stability (Xiao and Honig, 1999; Kumar and Nussinov, 1999).

Frequently, intersubunit interactions are mentioned as a potential major stabilization mechanism (Grabarse *et al.*, 1999; Chi *et al.*, 1999; Dams *et al.*, 2000). Interestingly, there is no single type of intersubunit interaction responsible for this stabilization (electrostatic interactions, hydrophobic interactions or disulfide bridges). In addition to a strengthened or increased surface area buried upon oligomerization, various

examples exist of hyperthermophilic proteins with a higher oligomerization state than their mesophilic homologues (Voorhorst *et al.*, 1995; Hess *et al.*, 1995; Villeret *et al.*, 1998; Chi *et al.*, 1999). For *Thermotoga maritima* phosphoribosylanthranilate isomerase, experimental evidence is available demonstrating that dimerization is a stabilization factor (Thoma *et al.*, 2000).

Finally, a multitude of other strategies for protein stabilization, which we will not consider in detail here, have been reported: deletion or shortening of loops, increased secondary structural content, helix dipole stabilization, docking of termini, anchoring of flexible loops, metal binding, posttranslational modifications and extrinsic parameters (Kumar *et al.*, 2000;Vieille and Zeikus, 2001).

2.3 Cold enzymes: stability - flexibility - activity

From homology-based models and a (small) number of X-ray structures, it appears that all structural factors currently known to stabilize the protein molecule can be attenuated in strength and number in cold-active enzymes (Russell, 2000; Sheridan *et al.*, 2000; Feller and Gerday, 2003). This involves the clustering of glycine residues (providing local mobility), the disappearance of proline residues in loops (providing enhanced chain flexibility between secondary structures), a reduction in arginine residues capable of forming multiple salt bridges and hydrogen bonds, as well as a lower number of ion pairs, aromatic interactions or hydrogen bonds compared to mesophilic enzymes. Frequently, stabilizing cofactors bind weakly, and loose or relaxed protein extremities seem to favour unzipping. The protein surface is generally characterized by the disappearance of several solvent-exposed ion pairs, the exposure of a higher proportion of nonpolar groups to the surrounding medium and an excess of negative charges that favour interactions with the solvent. These factors are thought to improve interactions with the solvent, which may be of prime importance in the acquisition of flexibility near zero degrees (Feller *et al.*, 1999). In multimeric enzymes, the cohesion between monomers is also reduced by decreasing the number and strength of interactions involved (Bell *et al.*, 2002). However, as observed for thermophilic proteins (Zavodzsky *et al.,* 1998), it appears that each protein family adopts its own strategy to decrease stability by using one or a combination of these structural alterations. By using this strategy the α-amylase from the Antarctic bacterium *Pseudoalteromonas haloplanktis* appears to have reached the lowest possible stability of its native state (Feller *et al.*, 1999) as indicated by a global collapse of its stability curve compared to its mesophilic homologue (corresponding to the reverse of strategy b in Figure 2.1).

Low temperatures slow down and strongly inhibit chemical reaction rates as is basically described by the Arrhenius equation:

$$k = A\mathrm{e}^{-Ea/RT},$$

where k is the rate constant, A is the pre-exponential factor, E_a is the so-called activation energy, R is the gas constant (8.31 kJ/mol), and T is the absolute temperature. Accordingly, any decrease in temperature will induce an exponential decrease of the reaction rate, the extent of which depends on the value of E_a. Therefore, the main challenge faced by psychrophilic enzymes is to maintain high catalytic activities (or rather catalytic efficiencies, as many enzymes evolved by optimizing k_{cat}/K_m; Wolfenden, 1983)

at low temperatures. Although it has been demonstrated by directed evolution experiments that low temperature activity and thermal stability can be improved simultaneously (Wintrode and Arnold, 2000), it remains a fact that such enzymes are not generally found in nature. A possible explanation may be that the low stability of psychrophilic enzymes is the simplest adaptive strategy to provide a gain in activity in the absence of selection for stable proteins in cold environments but in the presence of a strong selective pressure for highly active enzymes (Giver *et al.*, 1998; Cherry *et al.*, 1999; Roovers *et al.*, 2001). Similarly, the existence of highly thermostable enzymes that are more active than their mesophilic counterparts, even at 37 °C, suggests that thermostability is not incompatible with high activity (Sterner *et al.*, 1996; Merz *et al.*, 1996; Ichikawa and Clarke, 1998). Hyperthermophiles probably only need enzymes with activities at their living temperatures comparable to those of their homologues (Vieille and Zeikus, 2001). Interestingly, Kumar *et al.* (2001) found the free energy of unfolding at the living temperature of the source organism (i) to be uncorrelated with the living temperature, and (ii) to be relatively constant within T/M families, which indicates that the stability of individual proteins in the source organism may relate to their function. In other words: too high of a stability may hinder protein function. This observation is a manifestation of the concept of corresponding functional states, i.e. a high degree of similarity found in the physical and biochemical properties of homologous proteins from mesophiles and extremophiles under their respective optimum growth conditions (Somero 1983; Somero 1995).

A growing body of experimental data (including frequency domain fluorometry and anisotropy decay, hydrogen-deuterium exchange and tryptophan phosphorescence experiments) support the hypothesis that highly thermostable enzymes are more rigid (less flexible) than their mesophilic homologues and that rigidity is a prerequisite for high protein stability (Bönisch *et al.*, 1996; Jaenicke and Böhm, 1998; Zavodszky *et al.*, 1998; Manco *et al.*, 2000; Gershenson *et al.*, 2000). Lazaridis *et al.* (1997) argue that there is no fundamental reason for stability and rigidity to be correlated. Moreover, this "rigidity hypothesis" suffers from the lack of a single measure of flexibility. The spectrum of relaxation times characterizing conformational transition dynamics spreads over many orders of magnitudes from 10^{-11} s (local side chain rotations or hydrogen bond rearrangments on the protein surface) to hours or even years (the mean waiting time for protein spontaneous unfolding in physiological conditions) (Kurzynski *et al.*,1998). In other words, a protein can be rigid on a nanosecond scale but flexible on a millisecond scale. Furthermore, the contribution of dynamical effects on protein stability is poorly characterized (Daniel *et al.*, 2003).

It has also been proposed that excessive rigidity explains why hyperthermophilic enzymes are often inactive at low temperatures (Jeanicke and Böhm, 1998). One set of evidence that tends to support this hypothesis is that denaturants, detergents and solvents often activate hyperthermophilic enzymes at suboptimal temperatures. This activation tends to disappear as the temperature gets closer to the enzyme's temperature of maximal activity (Beaucamp *et al.*, 1997). At that temperature, the enzyme is flexible enough in the absence of a denaturant to show full activity. Concerning psychrophilic enzymes, the current accepted hypothesis is that, in order to perform catalysis at low temperatures, they have to increase their flexibility (Gerday *et al.*, 1997; Zavodszky *et al.*, 1998). The high flexibility of cold-active enzymes is strongly supported by experiments using dynamic quenching of fluorescence that show an increased permeability to a small quencher molecule (Feller and Gerday, 2003). An increase in flexibility can, however, be limited to only a small but crucial part of the protein and hence not lead to a uniformly unstable

protein, which in part may explain why stabilities and activities are not always inversely correlated (Jeanicke, 2000). One hypothesis is that global flexibility may be required when concerted motions of the whole molecule are involved in activity (possibly when large substrates are processed), whereas localized flexibility might only concern the essential components of the catalytic activity (possibly when small substrates are processed) (Feller and Gerday, 2003). In a study by Fields and Somero (1998), comparison of lactate dehydrogenase A_4 (A_4-LDH) kinetic properties and amino acid sequences of notothenioid teleosts having different adaptation temperatures, suggests that these enzymes have adapted to low temperatures by increases in flexibility in small areas of the molecule that affect the mobility of adjacent active site structures. The authors furthermore propose a model that explains a linked temperature-adaptive variation in K_m and k_{cat}, starting from the assumption that flexibility (necessary for function) causes each enzyme to occupy an ensemble of conformational states (Hilser and Freire, 1996; Zavodszky *et al.*, 1998). According to this model temperature-adaptive increases in k_{cat} will occur concomittantly with increases in K_m. The higher conformational mobility required for high activity (high k_{cat}) in the cold-adapted enzyme will lead to a higher number of conformational states available to the molecule and, as a result, to a larger proportion of enzyme conformations that bind substrates poorly or not at all. This will yield a higher K_m in the cold-adapted enzyme (at least at the same temperature). Indeed, most psychrophilic enzymes have higher K_m values than their mesophilic counterparts. For *Moritella abyssi* ornithine carbamoyltransferase (OTCase) and for *Moritella profunda* dihydrofolate reductase (DHFR) the higher K_m has apparently even resulted in a much lower catalytic efficiency (k_{cat}/K_m), suggesting that optimization of key metabolic enzymes at low temperatures may be constrained by natural limits (Xu *et al.*, 2003a,b).

When explained in terms of the activated complex theory[1], the increase in local flexibility may account for the lower activation enthalpies ($\Delta H^{\#}$) (Low *et al.*, 1973; Lonhienne *et al.*, 2000). Conversely, activation entropies ($\Delta S^{\#}$) for cold-adapted enzymes are higher, revealing the greater degree of ordering (or, in case of a positive $\Delta S^{\#}$, a smaller degree of disordering) these enzymes must undergo to form the activated complex (Figure 2.2).

[1] equation $k_{cat} = A(T)\, e^{-\Delta G\#/RT} = A(T)\, e^{-(\Delta H\#/RT - \Delta S\#/R)}$, with $A(T)$ the pre-exponential factor

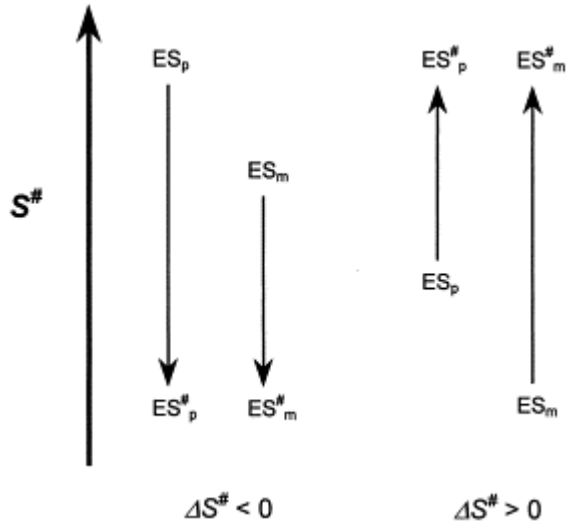


Figure 2.2 Graphical representation for the possible origin of $\Delta(\Delta S^{\#})_{p\text{-}m} < 0$. The transition state intermediate $ES^{\#}$ can be reached through a decrease ($\Delta S^{\#} < 0$) or an increase ($\Delta S^{\#} > 0$) of the activation entropy $S^{\#}$. This model postulates that the activation entropy of $ES^{\#}$ for both psychrophilic (p) and mesophilic (m) enzymes is similar because the same reaction is catalyzed. As a result of more flexible catalytic structures in psychrophilic enzymes, the distribution of conformational states for ES is broader and is translated into a higher level of $S^{\#}$. It follows that $\Delta S^{\#}_{p} - \Delta S^{\#}_{m}$ is always negative. Adapted from Lonhienne *et al.* (2000).

This seemingly unavoidable enthalpic-entropic compensation effect then results in a minor difference in terms of free energy of activation ($\Delta G^{\#}$). Recent calculations on subtilisin indicate that the residual flexibility in the active site is indeed significant (Villa *et al.*, 2000). Even without this compensatory entropic effect, according to theoretical considerations and computer simulations (Albery and Knowles, 1976; Warshel, 1998), an enzyme optimized under the evolutionary pressure of increasing k_{cat}/K_m will not gain much catalytic power from changing its ground state energy without changing the transition state (TS) energy. It then follows that, since most psychrophilic enzymes seem to fit in the model by Fields and Somero, a decrease of the temperature dependence of the enzymatic reaction by means of a decrease of the activation enthalpy probably is their main adaptive parameter. Interestingly, a number of enzymes (typically operating at subsaturating concentrations of substrate) are known to react against this adaptive drift and optimize substrate binding (lowering their K_m), while even optimizing their k_{cat} values (Bentahir *et al.*, 2000; Hoyoux *et al.*, 2001; Lonhienne *et al.*, 2001). Apparently, in this case local flexibility seems to be compatible with a lower number of conformational states as required for optimum binding.

In the case of the psychrophilic phosphoglycerate kinase from the Antarctic *Pseudomonas* sp. TACI18, it has been hypothesized that a heat-stable domain may improve substrate binding (Bentahir *et al.*, 2000), while maintaining a high k_{cat} by virtue of a heat-labile domain. Clearly, significant progress in our understanding of the role of (static/dynamic) flexibility in catalytic activity is needed to explain these findings.

2.4 Proteins under pressure

Pressure is a key physical parameter which has influenced the evolution and distribution of both micro-organisms and macro-organisms (Yayanos, 1986; Somero, 1990). The oceans have an average depth of 3800 m and thus an average pressure of 38 MPa as well as a maximal depth of approximately 11,000 m (~110 MPa). Additional high-pressure environments include deep lakes and the deep subsurface regions. Communities of micro-organisms have been detected as deep as 3500 m below the surface (Szewzyk *et al.*, 2000), and it is predicted that the largest number of prokaryotes in the biosphere are likely to reside in this realm below our feet (Whitman *et al.*, 1998). Finally, while astrobiologists ponder the possibility of life on Jupiter's moon Europa (Chyba and Phillips, 2001), it is interesting to note that such life may exist at pressures approximately twice that found in the deepest ocean trench on earth.

Comparitive studies have shown that pressure sensitivities of enzymes, structural proteins, and membrane based systems differ markedly between shallow- and deep-living species, and have indicated that the terms deep and high pressure begin to apply at depths of only 500 m or even less (corresponding to ~5 MPa) (Somero, 1992). In the case of enzymes, adaptation is characterized by pressure insensitive K_m values and reduced k_{cat} values (Somero, 1990; Somero, 1992). A recent comparison of orthologous proteins from a non-barophile (*Pyrococcus furiosus*) and a barophile (*Pyrococcus abyssi*) identifies the amino acids arginine, serine, glycine, valine and aspartic acid as having the most barophilic character and tyrosine and glutamine as the least barophilic (Di Giulio, 2005).

Pressure denaturation of monomeric proteins commonly requires a pressure in the 100 MPa range, which indicates that protein denaturation cannot play a significant role as a stress phenomenon in the adaptation of micro-organisms towards deep-sea conditions. In contrast to this well-documented finding, pressure-induced dissociation of oligomeric and multimeric proteins is observed well within the biologically relevant range of pressures, and has therefore been discussed as a possible mechanism underlying the growth inhibition of microorganisms at high pressures (Jaenicke, 1981; Jaenicke, 1987). All pressure effects accompany system volume changes in physical or biochemical processes and the response of the system is governed by the principle of Le Chatelier. Two fundamental relationships describe the pressure effects:

$$(\partial \ln K / \partial p)_T = -\Delta V / RT \qquad (\partial \ln k / \partial p)_T = -\Delta V^{\#} / RT,$$

where K is the equilibrium constant, k the rate constant, p the pressure, T the absolute temperature, R the gas constant and ΔV the difference of the final and initial volumes in the entire system at equilibrium, including the solute and surrounding solvent. $\Delta V^{\#}$ is the activation volume according to the Eyring theory. When a reaction is accompanied by a volume decrease, it is inhibited by elevated pressure, and vice versa. Applied to the structure and stability of biological macromolecules Gross and Jaenicke (1994) consider three kinds of interactions: ionic, hydrophobic and hydrogen bonding. Ion pairs in aqueous solution are strongly destabilized by hydrostatic pressure. This effect is attributed to the electrostrictive effect of the separated charges: each of it arranges water molecules in its vicinity more densely than bulk water. Thus the overall volume change favours the dissociation of ionic interactions under pressure. Formation of hydrogen bonds in biomacromolecules is connected to a negligibly small reaction volume, which may be positive or negative depending on the model system. Concerning the hydrophobic effect

Kauzmann (1987) reported that the volume change upon protein unfolding (ΔV), largely accounted for in terms of exposure of apolar residues to the aqueous medium, is positive at low pressures but negative at pressures higher than 100 – 200 MPa, and that the transfer of hydrocarbons into water is governed by an opposite behaviour, i.e. it shows a negative ΔV at low pressures and a positive one at high pressures. This failure of the 'liquid-hydrocarbon model' for pressure effects on protein stability has been addressed by Hummer *et al.* (1998), who pointed out that pressure and heat denaturation represent processes underlying distinct mechanisms: the first corresponds to the incorporation of water into protein, the second to the transfer of nonpolar groups into water. Thus pressure-denatured proteins have reduced compactness, yet they retain a considerably more ordered structure with respect to heat-denatured proteins (Zhang *et al.*, 1995; Hummer *et al.*, 1998). Generally speaking, there seems to be a great deal of experimental evidence that the pressure denatured state may resemble a molten globule state (Vidugiris and Royer, 1998). As a general rule, excluding high temperatures, the volume change upon the unfolding (partial or complete) of proteins is negative. The magnitude of the volume change is so small that it may well represent the sum of large contributions of opposite sign. A substantial negative contribution to ΔV is the elimination of internal cavities and voids upon disruption of the folded structure (Frye *et al.*, 1996; Richards, 1979). The origin and sign of various other contributions remain subjected to debate (Royer, 2002).

Structural determinants of protein piezostability are poorly characterized. The majority of experimental findings shows that thermostable proteins are piezostable as well (Mombelli *et al.*, 2002). It has been found, in fact, that thermophilic proteins in many cases are also stabilized by pressure values which lead to inactivation of the mesophilic counterparts (Hei and Clark, 1994; Bec *et al.*, 1996; McLean *et al.*, 1998; Summit *et al.*, 1998). Studies on a 7 kDa DNA-binding from *S. solfataricus* (Sso7d) have shown that aromatic interactions are important for increased thermal and piezostabilility (Konisky *et al.*, 1995). It was shown for the piezostable cytochrome P450 from *Sulfolobus solfataricus* that mutations decreasing the volume of the active site increase pressure stability (Tschirret-Guth *et al.*, 2001). Amino acid substitutions that decrease chain flexibility in staphylococcal nuclease result in an increase of stability of the protein at high pressure (Royer *et al.*, 1993). Therefore, increased structural homogeneity may be favoured by the high-pressure environments of the deep-sea. The results on the single-stranded DNA-binding protein of *Schewanella* strains pointed in the same direction: a decreased number of helix breaking and helix destabilizing residues in the central region of the protein suggests diminished flexibility and compressibility (Chilikuri and Bartlett, 1997).

The increase in protein stability on increasing pressure may be accounted for on the basis of the pressure-temperature phase diagram (Figure 2.3). A few elaborate studies have revealed elliptic contours for p/T diagrams, which corresponds to the well-known fact that proteins can undergo heat and cold denaturation, but also indicates that, over a broad temperature range proteins are first stabilized and then destabilized by increasing pressure. The stabilizing effect should correspond, for each tested protein, to the lower pressure range of the diagram. This implies that, for thermophilic proteins, both stabilization and destabilization occurs in a higher range of pressure as compared to mesophilic counterparts (Mombelli *et al.*, 2002).

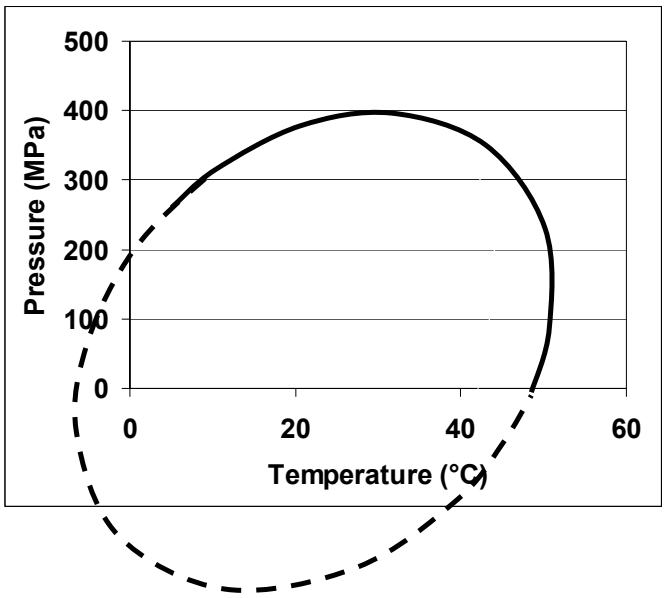


Figure 2.3 Stability phase diagram. The contour line connects points corresponding to zero Gibbs free energy of denaturation. Inside the ellipse, the protein is in the native state, outside in the denatured state. Following the full line from bottom right to top left, three different kinds of transition can be distinguished and explained in terms of the fundamental equation $\Delta G = \Delta V \mathrm{d}p - \Delta S \mathrm{d}T$. At high temperatures and low pressures, both ΔV and ΔS are positive, implying a positive slope for the contour lines. Increase of temperature or decrease of pressure may lead to denaturation under these conditions. At the pressure of maximum transition temperature, the sign of ΔV is reversed. From this point onward, denaturation at higher temperatures needs lower pressures: the protein can be denatured by increase in pressure or temperature. Finally, at the temperature of maximum transition pressure, ΔS also becomes negative, leading to the cold denaturation phenomenon. The denaturation pressure decreases with decreasing temperature, allowing cold denaturation at least at elevated pressure. Extrapolation of this curve to sub-zero temperatures led to the suggestion that cold denaturation should occur at atmospheric pressure in supra-cooled aqueous solutions. Adapted from Gross and Jaenicke, (1994) and Scharnagl *et al.*, (2005).

Pressure influences enzyme behaviour mainly by affecting the individual rate constants involved in the kinetic mechanism and that concur to the definition of k_{cat} and K_m. Although the prediction of these effects on K_m is not straightforward, it is expected that an increase in pressure will increase the k_{cat} in reactions for which $\Delta V^{\#}$ of the rate limiting step has a negative value (Gross and Jaenicke, 1994; Mozhaev *et al.*, 1996; Mozhaev *et al.*, 1994). On changing from sea level to the ocean floor, 20 °C difference in temperature may decelerate reaction rates by a factor of 4 – 10 (depending on the activation energy), whereas effects of the increase in pressure will hardly exceed 15 %. Thus it seems clear that, from the evolutionary point of view, adaptation to deep-sea conditions is dominated by low temperature rather than high pressure (Gross and Jaenicke, 1994). However, protein compressibility is directly related to the structural and conformational fluctuations of proteins at normal atmospheric pressure (Cooper, 1976; Gekko and Hasegawa, 1986). The reduced specific volume of a pressurized protein has been proposed to bring a reduced flexibility and hence an inhibition of activity in cases where flexibility is crucial for biological function (Tsou, 1986; Huber, 1988; Gross and Jaenicke, 1992).

2.5 Additional remarks: role of the cosolvent, allosteric regulation

The influence of different types of cosolvent on protein stability is a vast and diverse subject that will not be discussed in detail here. As a major type of cosolvent, ions have widely different effects according to their size and charge: e.g. all the major intracellular anions (phosphates, sulphates and carboxylates) are kosmotropes (interacting strongly with solvent molecules), whereas the major intracellular monovalent cations (K^+, ammonium, guanidinium and imidazolium groups) are chaotropes (interacting weakly with solvent molecules). Among physiologically relevant anions, Cl^- is the only chaotrope and tends to interact with basic groups of proteins with significant affinity (Collins, 2004).
The potential origins for salt interactions with proteins can be summarized as: (i) differential cation binding, (ii) differential anion binding, (iii) differential hydration, (iv) differential screening causing a variation of the macromolecular activity coefficient and (v) effects of electrolytes on the activity coefficient of a ligand (Record *et al.*, 1978).

Enzymes from extreme halophiles, while performing identical functions as their non-halophilic counterparts, have been shown to exhibit substantially different properties. Among them is the requirement for high salt concentrations, in the 1 – 4 M range, for activity and stability, and a high excess of acidic over basic amino acid residues. It is argued that the high negative surface charge of halophilic proteins renders them more flexible at high salt concentrations, conditions under which non-halophilic proteins tend to aggregate and become rigid. This high surface charge is neutralized mainly by tightly bound water dipoles (Mevarech *et al.*,). The requirement of high salt concentrations for the stabilization of halophilic proteins appears to result, according to Elcock and McCammon (1998), from repulsive electrostatic interactions between the abundant negative surface charges that remain destabilizing, even at high salt concentrations,. Consequently, the role of acidic residues may be more to prevent aggregation than to make a positive contribution to protein stability. Except for the need to maintain normal hydration, halophilic proteins do not seem to exhibit specific structural properties (Jaenicke, 2000).

The fundamental 'linkage relations' between stability and binding were developed by Wyman (1948). In general, the Gibbs free energy of any given protein state, i, is modified by ligand binding as follows:

$$\Delta G_i = \Delta G^0_i - RT \ln ((1 + K_{B,i}[L]) / (1 + K_{B,0}[L])),$$

where ΔG^0_i is the Gibbs free energy in the absence of ligand, $K_{B,0}$ is the binding constant to the native state, $K_{B,i}$ is the binding constant to state i, and [L] the free ligand concentration. This equation dictates that protein states are structurally stabilized proportionally to their binding affinity towards a ligand. The equation has usually been applied to the equilibrium between the native and the unfolded state in denaturation experiments but is equally valid for the equilibrium between two different allosteric states. Indeed, although we have not discussed the effect of physical variables (like pressure and temperature) on allosteric equilibria, the problem is fundamentally the same: as for protein stability, both (or multiple) states are differentially stabilized by means of the typical (non-covalent) interactions that themselves are a complex function of these physical variables.

If one considers proteins in their native state as an ensemble of fluctuating conformational states, binding of a ligand will result in a redistribution of the population of states, in which states that were only negligibly populated in the unliganded ensemble become significantly populated. This redistribution corresponds to a separate functional state (different biological activity) in an allosteric protein. Luque *et al.* (2002) have shown that the propagation of binding effects is more efficient when the binding site is located in a region of low stability (in other words a more extensive redistribution of states takes place). Stability and functional cooperativity hence seem to be linked, and attempts are even made to identify signal transmission pathways by mapping regions prone to local unfolding on protein structures (Freire, 2000).

As a final remark, we note the (potential and established) economic benefits of extremophilic enzymes (discussed or reviewed in: Vieille and Zeikus, 2001; Gerday *et al.*, 2000; Gross and Jaenicke, 1994; Mombelli *et al.*, 2002).

Part II

ATCase

3 Introduction to ATCase

3.1 ATCases: function and classification

The *de novo* synthesis of pyrimidines is universal. The pathway consists of six enzymatic steps leading to the formation of UMP, which is further converted into UTP, CTP, dCTP, and dTTP. Studies of de novo pyrimidine biosynthesis and salvage are actively pursued because of the relationship of these pathways to growth, development, and chemotherapy. Organisms from the Archaea, Bacteria and Eukarya all produce aspartate carbamoyltransferase (ATCase, E.C. 2.1.3.2), the enzyme that catalyzes the committed step of the pyrimidine biosynthetic pathway, i.e. the reaction of carbamoyl phosphate and L-aspartate to form *N*-carbamoyl-L-aspartate and inorganic phosphate (Jones *et al.*, 1955). There are four major classes or forms of quaternary structure known for aspartate carbamoyltransferases. In prokaryotes, ATCase is known to exist in three classes. Class A consists of 480 kDa dodecamers of two catalytic trimers and six subunits that are either dihydro-orotase (as in *Thermus aquaticus*; Van de Casteele *et al.*, 1997), the third enzyme in the pathway, or inactive dihydro-orotase homologues (e.g. as for *Pseudomonas aeruginosa*; (Schurr *et al.*, 1995; Vickrey *et al.*, 2002). Class B ATCases are 310 kDa dodecamers consisting of two catalytic trimers and three regulatory dimers that bind allosteric effectors (e.g. as in *Escherichia coli*; Wiley and Lipscomb, 1968). Class C ATCases are unregulated 100 kDa catalytic trimers (e.g. as in *Bacillus subtilis*; Brabson and Switzer, 1975). In eukaryotes, aspartate carbamoyltransferase is often part of a large multifunctional protein, such as those found in yeast and hamster (Simmer *et al.*, 1989). The core catalytic subunit is common to all known ATCases, because the active sites are composed of residues contributed by adjacent subunits within the trimer (Robey and Schachman, 1985).

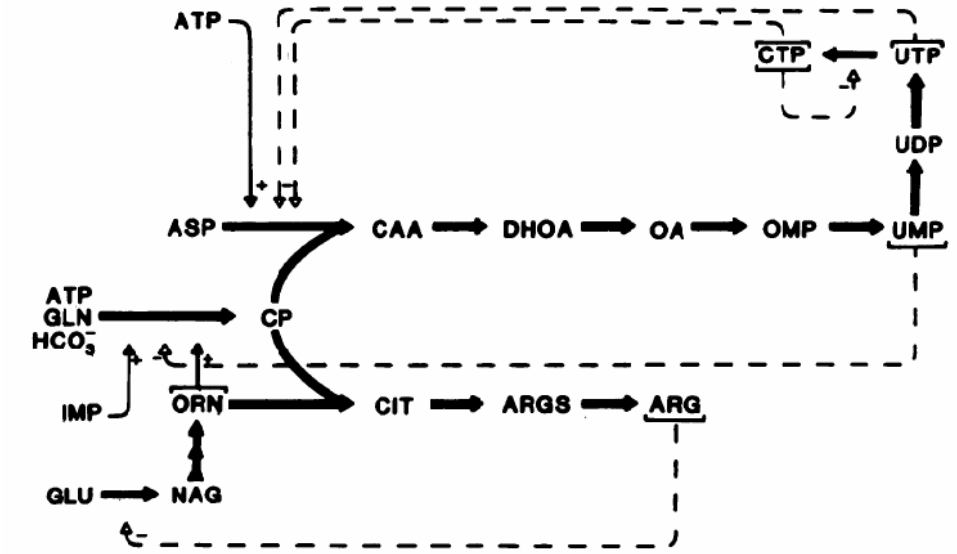


Figure 3.1 Allosteric control patterns for *de novo* pyrimidine and arginine biosynthesis in *E. coli*. CTP and UTP provide synergistic feedback inhibition of ATCase, whereas ATP provides competitive heterotropic activation. The feedback inhibition of carbamoylphosphate synthetase (CPSase) by UMP is subject to antagonism by ornithine (ORN) and the enzyme is activated by IMP. *N*-acetylglutamate synthase is subject to feedback inhibition by L-arginine (ARG). The biosynthetic intermediates of the *de novo* pyrimidine pathway are glutamine (GLN), carbamoylphosphate (CP), aspartate (ASP) carbamoyl aspartate (CAA), dihydroorotic acid (DHOA), orotic acid (OA), orotidine-5'-monophosphate (OMP) and the common uridylates (UMP, UDP and UTP). The identified intermediates of the arginine biosynthetic pathway are glutamate (GLU), *N*-acetylglutamate (NAG), ornithine (ORN), citrulline (CIT), arginosuccinate (ARGS) and arginine (ARG). Adapted from Wild *et al.* (1989).

Aspartate carbamoyltransferase (together with and independently from threonine deaminase; Umbarger, 1956) is the first enzyme for which feed-back inhibition was demonstrated (Yates and Pardee, 1956). Since then the enzyme from *Escherichia coli* has been extensively studied and has become the prototype of allosteric enzymes (Allewell, 1989; Hervé, 1989; Wild and Wales, 1990; Lipscomb, 1994). *E. coli* ATCase exhibits positive homotropic interactions between catalytic sites for binding of the substrate L-aspartate, which is explained by a concerted transition of the enzyme from a conformation of low affinity for aspartate (T or "taut" state) to a conformation of high affinity (R or "relaxed" state) for this substrate. The enzyme also exhibits heterotropic interactions between catalytic and regulatory sites. The enzyme is feedback inhibited by the end product CTP; UTP is a synergistic inhibitor in the presence of CTP, and ATP is an (antagonistic) activator. Wild *et al.* (1989) have pointed at the inherent metabolic logic when integrating these allosteric patterns with other findings on the regulation of *de novo* pyrimidine and arginine biosynthesis in *E. coli* (Figure 3.1). Elevated levels of CTP would exert two simultaneous regulatory functions, as a partial inhibitor for the entire de novo pathway through ATCase and for direct control over its own synthesis through competitive inhibition of CTP synthetase. Since the inhibition of ATCase is only partial, significant pyrimidine synthesis would continue. However, this would primarly serve to increase the UTP pool since the conversion of UTP to CTP would be restricted at CTP synthetase. Ultimately, CTP and UTP pools would be elevated, thus permitting the thorough inhibition of ATCase. Elevated ATP levels would oppose the effects of CTP and UTP on ATCase and thus provide for a balancing of pyrimidine nucleotide pools with their purine counterparts.

3.2 *E. coli* ATCase is a paradigm for allosteric regulation

The catalytic (c) and regulatory (r) chains of *E. coli* ATCase have molecular masses of 34 and 17 kDa, respectively (A). The three c chains of each catalytic trimer are related by a threefold axis, while three twofold axes relate pairs of c chains in the two trimers and pairs of r chains within the three regulatory dimers (Figure 3.2). Each chain consists of two domains (Monaco *et al.*, 1978). The N- and C-terminal domains of the c chains are termed the polar and equatorial domains, respectively; the corresponding domains of the r chains are designated the allosteric and zinc domains. Carbamoyl phosphate binds between the two c chains in such a way as to interact with the polar domains of both (Robey and Schachman, 1985; Krause *et al.*, 1985). Groups with which L-Asp interacts, except for Lys 84, are located in the equatorial domain of a single chain. The primary nucleotide binding

site is in the N-terminal domain of the r chain; the C-terminal domain contains a Zn ion coordinated by four sulfhydryl groups. There are five unique interchain interfaces in the unliganded protein: C1–C2, between c chains within catalytic trimers; C1–C4, between c chains in different catalytic trimers; C1–R1 and R1–C4, two types of c:r contacts; and R1–R6, the r:r contact within the regulatory dimers (Honzatko *et al.*, 1982) (Figure 3.2). The twofold molecular symmetry is distorted in the crystal in the space group P321, particularly when CTP is bound. Under these conditions, the C5–C6, C6–R6, C6–R3, and analogous interfaces in one half of the molecule are not equivalent to the C1–C2, C1–R1, C1–R4, and analogous interfaces, in the other half of the molecule (Kim *et al.*, 1987).

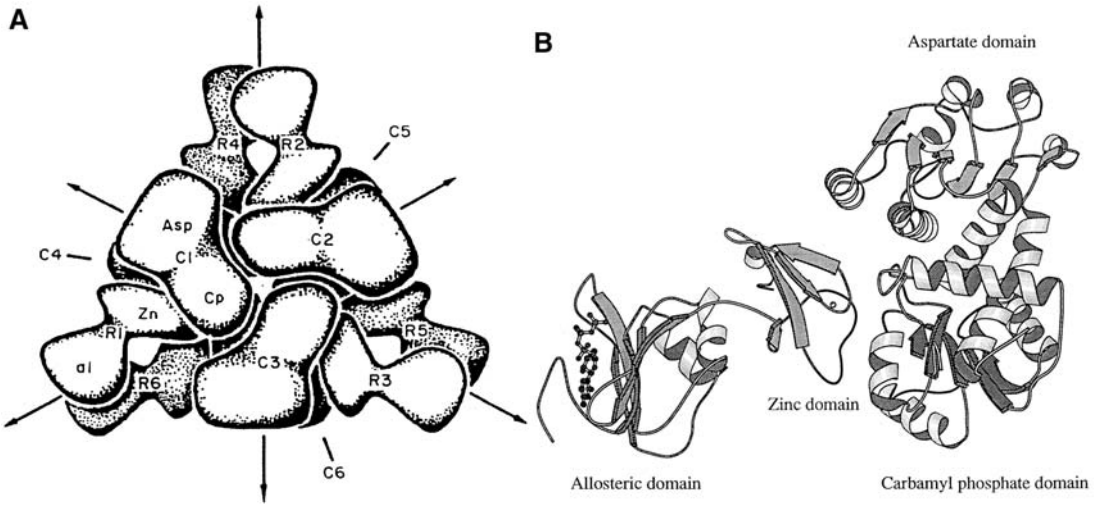


Figure 3.2 (A) Schematic diagram of the structure of ATCase. (B) Close-up of the structures of the catalytic and regulatory chains. Adapted from Roche and Field (1999).

High resolution crystal structures are available of *E. coli* ATCase in the unliganded form and in complex with substrates, products and/or isosteric inhibitors (Honzatko *et al.*, 1982; Ke *et al.*, 1984; Krause *et al.*, 1985; Krause *et al.*, 1987; Gouaux and Lipscomb, 1990; Huang and Lipscomb, 2004). These studies indicate that there are two distinct structural states, each characterized by a strong conservation of molecular symmetry among the monomers, in accordance with the symmetry model of allosteric regulation (Monod *et al.*, 1965). Binding of the bisubstrate analogue PALA (*N*-phosphonacetyl-L-aspartate) causes the enzyme to expand by 12 Å along the threefold axis while the catalytic trimers and regulatory dimers rotate 10° and 15°, respectively, about their symmetry axes (Ladner *et al.*, 1982) (Figure 3.3). A multitude of other studies seem to confirm that substrate binding induces a concerted shift in the equilibrium between two quaternary conformations of the oligomeric protein (Schachman, 1988). In contrast to the transitions from 'taut' to 'relaxed' structures postulated in early models of allostery, the T – R transition involves breaking of bonds at some interfaces and simultaneous formation of new bonds at others, as illustrated in Figure 3.4. Bonds are broken at the interfaces between zinc and allosteric domains in the r chain and at the C1–C4 and R1–C4 interfaces between subunits. New bonds form at the interface between polar and equatorial domains in the c chain and at the C1–R1 and C1–C2 interfaces between subunits. Changes at the R1–R6 interface are minor.

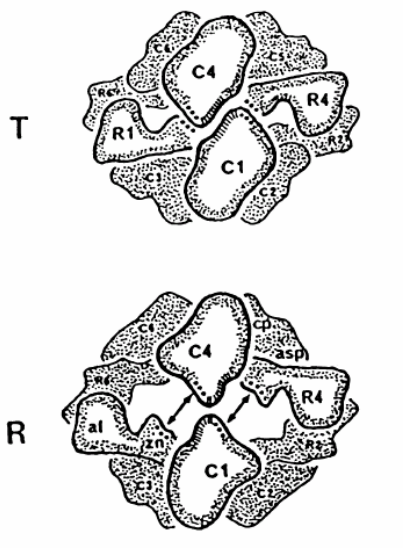


Figure 3.3 Dodecameric structure of T and R states of ATCase. The threefold axis is in the plane of the page, and the twofold axis is perpendicular to the plane of the page. The dots indicate the surfaces of the enzyme involved in the C1-R4 and symmetry-related R1-C4 interfaces. The slashes define the region involved in the C1-C4 interface. In the T to R transition, the separation of catalytic trimers increases by 10.8 Å along the threefold axes while the trimers rotate 12° relative to one another around the threefold axes. The regulatory dimers rotate 15° about each of the three noncrystallographic twofold axes, and the dimers separate by 1.4 Å. Adapted from Stevens and Lipscomb (1992).

The crystallographic results suggest the following sequence of events (Kantrowitz and Lipscomb, 1988; Ke *et al.*, 1988; Wang *et al.*, 2005). Binding of carbamoyl phosphate first triggers changes in tertiary structure in residues 50 – 55 and 73 – 85 of the polar domains, accompanied by a transformation of the active site electrostatics, thus creating the binding site for Asp. Binding of Asp would then cause the Asp domain, the 240s loop, and the 80s loop of the adjacent chain to close in on the two substrates. This motion has two consequences. Firstly, it forces the substrates together, thereby lowering the activation energy of the reaction, and secondly, it further weakens the intersubunit interactions that specifically stabilize the T quaternary state. The weakening of these interactions triggers the initiation of the global quaternary conformational change. In fact, the 240s loops cannot attain their final domain-closed conformation without an expansion along the threefold axis, which allows the 240s loops from the upper and lower catalytic subunits to slide past each other.

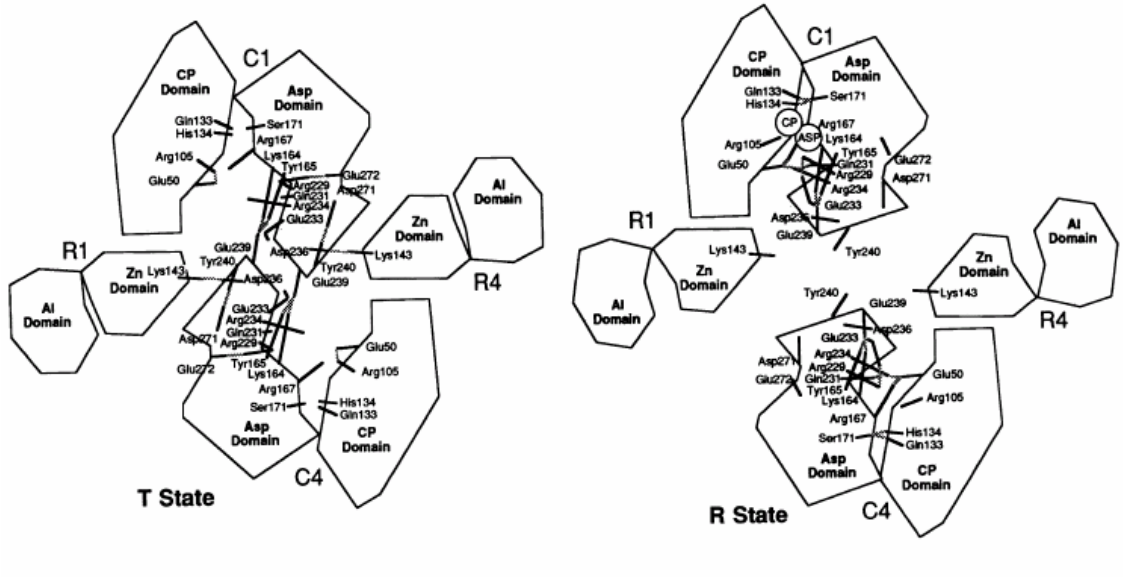


Figure 3.4 Schematic representation of interactions related to homotropic cooperativity in the T state (left) and R state (right) of the *E. coli* ATCase holoenzyme. Adapted from Lipscomb (1994).

Studies with hybrid enzymes (more specifically with c and r chains from a different source) have indicated that the heterotropic response of B-class ATCases is determined by the regulatory subunit, independent from the source of the catalytic polypeptides (Wild and Wales, 1990). Interestingly, the heterotropic regulatory behaviour of the *E. coli* enzyme may be more related to the sequential allosteric model (Koshland *et al.*, 1966), with "local" (pairwize) rather than global conformational changes as the primary mechanism. One of the main arguments opposing the simple two-state model is the possibility of (i) decoupling heterotropic and homotropic effects, and (ii) CTP and ATP effects, suggesting multiple mechanisms rather than effectors shifting the same equilibrium (Enns and Chan, 1978; Kantrowitz *et al.*, 1981; Baillon *et al.*, 1985; Foote *et al.*, 1985). Alternative structural models were subsequently proposed (Thiry and Hervé, 1978; Xi *et al.*, 1991; Stevens and Lipscomb, 1992).

Crystal structures are also available of effector (CTP and ATP) complexes in both allosteric states, and structural changes upon effector binding have been described in detail (Kim *et al.*, 1987; Stevens *et al.*, 1990; Gouaux *et al.*, 1990). These studies (but also solution X-ray scattering experiments; Hervé *et al.*, 1985; Fetler *et al.*, 1995) indicate that the structural changes caused by the binding of nucleotides are small. Binding of ATP and CTP causes 0.3 Å changes in the separation distance between the two chains in the regulatory dimers, and 0.4 – 0.5 Å changes in the separation distance between the catalytic trimers (Stevens and Lipscomb, 1992). Structural changes within the effector sites and the active sites were also observed. However, the structures reveal no information on the transmission of the information between the effector sites and the active sites. Site-directed mutagenesis has been more successful in identifying residues that appear to be involved in the communication pathway(s) for the regulatory signals (Aucoin *et al.*, 1996; De Staercke *et al.*, 1995; Xi *et al.*, 1994). Nevertheless, the mechanism of heterotropic regulation of *E. coli* ATCase remains debated.

3.3 Extremophilic ATCases

Several ATCases from extremophilic organisms have been characterized. Hyperthermophilic archaea such as *Pyrococcus abyssi*, *Sulfolobus acidocaldarius* and *Methanococcus jannaschii* produce B-class ATCases (as do *Enterobacteriaceae*) (Purcarea *et al.*, 1994; Durbecq *et al.*, 1999; Hack *et al.*, 2000; Labedan *et al.*, 2004). *P. abyssi* ATCase exhibits homotropic cooperative behaviour for aspartate, and heterotropic interactions for ATP (activator) and CTP and UTP (inhibitors). Although intrinsically thermostable, the catalytic trimers are further stabilized by association with the regulatory dimers. A remarkably high affinity for the highly thermostable substrate carbamoyl phosphate is a possible adaptation to high temperature. It is suggested to be correlated with the metabolic channelling of this substrate between the carbamoyl phosphate synthetase (CPSase) and the ATCase. The high resolution X-ray structure of the catalytic trimer of this enzyme has also been reported and has identified ion pairs, ion pair networks and shortening of loops and termini as potentially thermostabilizing features (Purcarea *et al.*, 1994; Van Boxstael *et al.*, 2003; Van Boxstael *et al.*, 2005). *S. acidocaldarius* ATCase (a subject of this work) also exhibits homotropic interactions for aspartate, yet the (holo-) enzyme is activated by the nucleotides ATP, CTP and UTP (these effectors exert an inhibitory effect on the catalytic subunits by competing with carbamoyl phosphate for the same binding site). Only monomers could be found in extracts of recombinant *E. coli* or *Saccharomyces cerevisiae* cells expressing the catalytic chain. In the presence of carbamoyl phosphate these monomers assemble into trimers. The *S. acidocaldarius*

catalytic trimer is not intrinsically stable and the stability of the enzyme is achieved by the association of the catalytic and regulatory subunits. A more comprehensive description of the biochemical and kinetic properties of the *S. acidocaldarius* ATCase has been given by Durbecq *et al*. (1999). *M. jannaschii* ATCase shows limited homotropic cooperativity for aspartate and has little if any regulatory properties (Hack *et al.*, 2000). The weak allosteric properties of the *Methanococcus* ATCase might be related to an inherently unstable association between catalytic and regulatory subunits.
Of the domain of the Bacteria, (hyper-)thermophilic ATCases are characterized from *Thermotoga maritima* and *Aquifex aeolicus*. *Thermotoga maritima* ATCase contains catalytic and regulatory regions within a single polypeptide (also called class B' ATCases). It shows homotropic cooperative interactions towards aspartate and heterotropic effects towards the nucleotides ATP (activator), CTP and UTP (inhibitors) (Chen *et al.*, 1998). The *Aquifex aeolicus* is a homotrimer of 34 kDa catalytic chains. The K_m values for both substrates are 30 – 40 fold lower than the corresponding values of the *E. coli* ATCase catalytic trimer. Furthermore, there is strong evidence for channelling and transient complex formation between *A. aeolicus* ATCase and CPSase (Purcarea *et al.*, 2003). The moderate thermophile *Thermus* ZO5 has an ATCase that is part of a multi-enzyme complex with a molecular mass of 480 kDa. The complex also contains a (ATCase stabilizing) DHOase and probably the *bbc*-encoded protein. The bbc-protein appears to play a role in (UTP) feedback inhibition of the ATCase (Van De Casteele *et al.*, 1997).

Psychrophilic ATCases have been described from the Antarctic bacterial strain *TAD1* and from the deep-sea *Vibrio* strain 2693 (later identified as *Moritella abyssi*; Xu *et al.*, 2003c). The first ATCase was studied in dialyzed cell-free extracts. It shows no cooperativity nor regulation by ATP, but (mutual antagonistic) inhibition by CTP and UTP (Sun *et al.*, 1998). The K_m for carbamoyl phosphate was 100 times higher than for *E. coli* ATCase, which was suggested to result from an increase in k_{cat} (instead of a lower affinity). The high activity at low temperatures was suggested to be caused by a discrete modification at the active site, having no influence on the global thermostability of the enzyme. The *Moritella abyssi* ATCase has a molecular weight of 320 kDa compatible with the [2(c3):3(r2)] class B architecture (Xu *et al.*, 1998). It exhibits homotropic interactions for aspartate and heterotropic inhibition by CTP, but no activation by ATP and no synergistic inhibition by CTP and UTP. The *M. abyssi* ATCase was studied in *M. abyssi* and recombinant *E. coli* extracts. In the present thesis we report on the functional characteristics and the crystal structure of the purified recombinant ATCase of the closely related organism *Moritella profunda*.

Finally, the ATCase has been studied of the extremely halophilic archaeon *Halobacterium cutirubrum*, and of the moderately halophilic eubacterium *Vibrio costicola*. The activity of the *H. cutirubrum* enzyme was dependent on high salt concentration, optimal activity requiring 3.5 – 4 M NaCl or KCl. With aspartate, sigmoidal curves indicative for homotropic interactions were found when extracts were assayed immediately after preparation. The enzyme was highly sensitive to retroinhibition by CTP; inhibition was also dependent on high salt concentration, but was independent on assay temperature. The estimated molecular weight is 160 kDa (Liebl *et al.*, 1969; Norberg *et al.*, 1973). The ATCase of *V. costicola* (estimated MW 130 – 150 kDa) also looses activity rapidly at lower salt concentrations, yet differs from the extremely halophilic archaebacterium in its salt-response patterns of activity and regulation. ATCase activity is optimal at about 1.5 M NaCl or 1 M KCl, although high activity was detected at 0.15 M NaCl. Significantly,

strong regulation by CTP takes place in low (0.15 – 0.3 M), but not high, NaCl concentrations (Ahonkhai *et al.*, 1989).

4 Crystal structure of T state aspartate carbamoyltransferase of the hyperthermophilic archaeon *Sulfolobus acidocaldarius*

doi:10.1016/j.jmb.2004.03.079 J. Mol. Biol. (2004) 339, 887–900

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Crystal Structure of T State Aspartate Carbamoyltransferase of the Hyperthermophilic Archaeon *Sulfolobus acidocaldarius*

Dirk De Vos[1], Filip Van Petegem[1], Han Remaut[1], Christianne Legrain[2] Nicolas Glansdorff[3] and Jozef J. Van Beeumen[1]*

[1]*Laboratorium voor Eiwitbiochemie en Eiwitengineering, Universiteit Gent, Ghent, Belgium*

[2]*Institut de Recherches Microbiologiques J-M Wiame (IRMW), Brussels, Belgium*

[3]*Erfelijkheidsleer en Microbiologie, Vrije Universiteit Brussel, Brussels Belgium*

**Corresponding author*

Aspartate carbamoyltransferase (ATCase) is a model enzyme for understanding allosteric effects. The dodecameric complex exists in two main states (T and R) that differ substantially in their quaternary structure and their affinity for various ligands. Many hypotheses have resulted from the structure of the *Escherichia coli* ATCase, but so far other crystal structures to test these have been lacking. Here, we present the tertiary and quaternary structure of the T state ATCase of the hyperthermophilic archaeon *Sulfolobus acidocaldarius* ($SaATC_T$), determined by X-ray crystallography to 2.6 Å resolution. The quaternary structure differs from the *E. coli* ATCase, by having altered interfaces between the catalytic (C) and regulatory (R) subunits, and the presence of a novel C1–R2 type interface. Conformational differences in the 240s loop region of the C chain and the C-terminal region of the R chain affect intersubunit and interdomain interfaces implicated previously in the allosteric behavior of *E. coli* ATCase. The allosteric-zinc binding domain interface is strengthened at the expense of a weakened R1–C4 type interface. The increased hydrophobicity of the C1–R1 type interface may stabilize the quaternary structure. Catalytic trimers of the *S. acidocaldarius* ATCase are unstable due to a drastic weakening of the C1–C2 interface. The hyperthermophilic ATCase presents an interesting example of how an allosteric enzyme can adapt to higher temperatures. The structural rearrangement of this thermophilic ATCase may well promote its thermal stability at the expense of changes in the allosteric behavior.

Keywords: Sulfolobus acidocaldarius; aspartate carbamoyltransferase; crystal structure; thermostability; allosterism

Introduction

Aspartate carbamoyltransferase (ATCase, EC 2.1.3.2) catalyzes the formation of phosphate and *N*-carbamoyl-L-aspartate from carbamoyl phosphate and L-aspartate, initiating the first step in the synthesis of pyrimidines.[1,2] ATCase is extensively studied as a model for structure–function relationships in cooperativity and allosteric regulation mechanisms. The structure and properties of the *Escherichia coli* enzyme have been studied in detail.[3–8] This dodecameric enzyme is composed of two types of polypeptide chains: the 34 kDa catalytic chains (c) encoded by the *pyrB* gene, and the 17 kDa regulatory chains (r) encoded by the *pyrI* gene.[9–11] The quaternary structure of this enzyme results from the association of two catalytic chain trimers (c_3) held together through their interactions with three regulatory chain dimers (r_2). Both chain

Present addresses: F. Van Petegem, Cardiovascular Research Institute, University of California San Francisco, San Francisco, USA; H. Remaut, School of Crystallography, Birkbeck College, University of London, London, UK.

Abbreviations used: ASA, accessible surface area; ATCase, aspartate carbamoyltransferase; c, catalytic subunit; PDB, RCSB Protein Data Bank; r, regulatory subunit; rmsd, root-mean-square-deviation; $EcATC_T$, *Escherichia coli* T state ATCase; $EcATC_R$, *E. coli* R state ATCase; $SaATC_T$, *Sulfolobus acidocaldarius* T state ATCase.

E-mail address of the corresponding author: jozef.vanbeeumen@ugent.be

types consist of two domains, carbamoyl phosphate and aspartate binding domains in the case of the catalytic chains and allosteric and zinc (Zn) domains in the case of the regulatory chains. The catalytic sites are located at the interface between two catalytic chains belonging to the same trimer and involve residues belonging to both chains.[12,13] The allosteric domain contains the binding site for nucleotide effectors and the zinc binding domain makes contacts with the catalytic subunits.

E. coli ATCase shows positive homotropic interactions between catalytic sites for binding of the substrate L-aspartate, a phenomenon which is explained by the transition from a T state (referred to as *Ec*ATC$_T$) having a low affinity for aspartate to an R state (*Ec*ATC$_R$) with a high affinity for this substrate. The crystal structures of these two extreme conformations have been solved at 2.6 Å and 2.1 Å resolution, respectively.[14,15] Analysis of these structures indicates that the holoenzyme undergoes substantial conformational changes upon T–R transition. Changes at the quaternary level include an expansion by 12 Å along the molecular 3-fold axis, along with rotations of the catalytic subunits about the 3-fold axis and rotations of the regulatory subunits about their respective 2-fold axes. Changes at the tertiary level include the closure of the aspartate and carbamoyl phosphate domains to form the active site, as well as the rearrangement of the 80s and 240s loop regions. The enzyme also exhibits heterotropic interactions between catalytic and regulatory sites. Its activity is synergistically feedback inhibited by the two end products CTP and UTP, and is stimulated by ATP. These effectors bind to the same regulatory sites, at a distance of 60 Å from the catalytic sites.[16]

The enzymatic properties of the ATCase from *Sulfolobus acidocaldarius*, a thermoacidophilic archaeon of the *Crenarchaeota* group, were investigated previously.[17] This enzyme was shown to be highly thermostable and to display cooperativity between catalytic sites for the binding of its substrate aspartate. It exhibits, however, a very different regulation pattern: the holoenzyme is strongly activated by all nucleoside triphosphates, while *E. coli* ATCase is inhibited by pyrimidine nucleotides and activated only by ATP. It was shown furthermore that, in contrast to *E. coli* and other hyperthermophilic ATCases,[18,19] the catalytic trimer is not intrinsically stable and the stability of the enzyme is achieved by the association of the catalytic and the regulatory subunits.

Previously, 3.5 Å and 1.8 Å resolution structures of the catalytic trimer of *Methanococcus jannaschii* and *Pyrococcus abyssi* ATCase were described,[20,21] but the structure of a complete dodecameric enzyme to test the hypotheses from the *E. coli* ATCase has been lacking. We now report the first hyperthermophilic ATCase structure of the intact (unliganded) holoenzyme ([$2c_3 : 3r_2$]) at 2.6 Å resolution. We compare the tertiary and quaternary structure with the *Ec*ATC$_T$ and *Ec*ATC$_R$ structures in terms of the overall architecture and possible implications for its high thermostability. Special attention was given to the different interfaces between the different subunits because of their important role in the allosteric behavior and the thermostability of this enzyme.

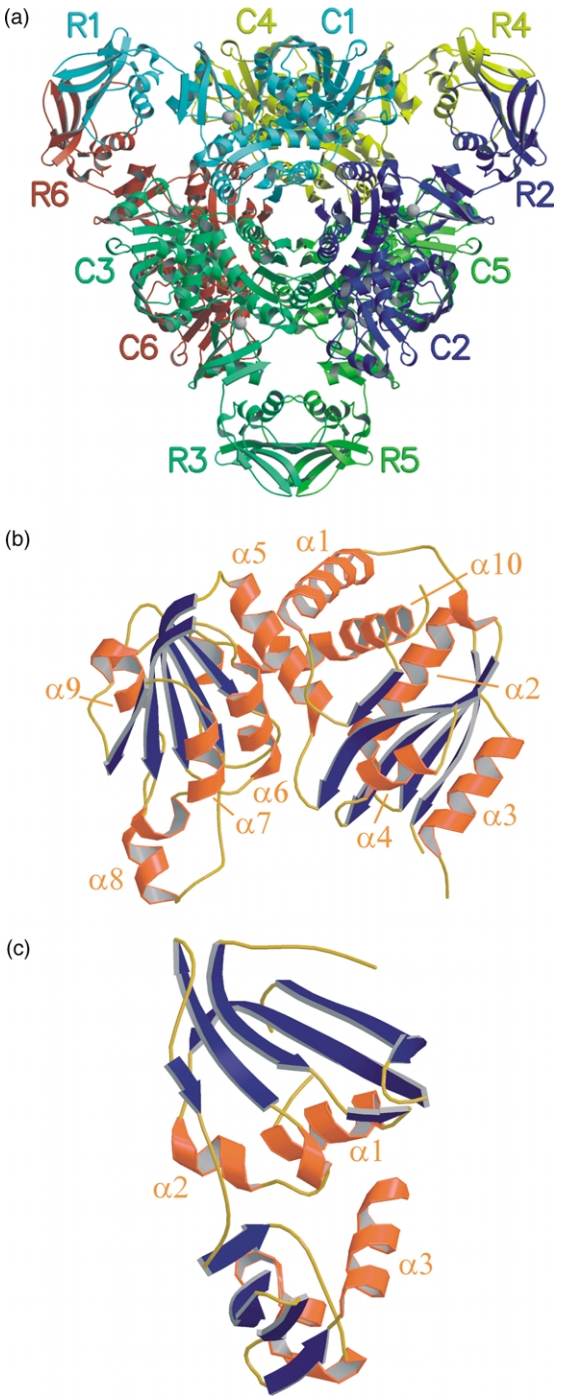


Figure 1. (a) Overview of the *S. acidocaldarius* ATCase holoenzyme viewed along its molecular 3-fold axis. Each Cn–Rn dimer is depicted in its own color. (b) and (c) Ribbon diagrams of the *Sa*ATC$_T$ catalytic and regulatory subunit structures. α-Helical regions are numbered in accordance with Figure 7.

Results

Overall structure

The catalytic and regulatory monomers of the *S. acidocaldarius* ATCase form a dodecamer with D_3 point group symmetry. The enzyme crystallizes in space group *P*6322 with one catalytic and one regulatory subunit in the asymmetric unit. The heterododecameric complex (Figure 1(a)) is generated by the crystallographic symmetry operators of the *P*6322 space group, resulting in two dodecameric enzymes per unit cell. The statistics of the structure determination can be found in Table 1. The overall quaternary structure resembles that of a T state ATCase.[12] We will therefore refer to this structure as *Sa*ATC$_T$.

Catalytic monomer structure

The *Sa*ATC$_T$ catalytic subunit has 299 residues and shares 43% sequence identity with the *E. coli* ATCase catalytic subunit. It is composed of two domains: the carbamoyl phosphate domain (or polar domain) and the aspartate domain (or equatorial domain), linked by two α-helices. Each domain consists of a β-sheet of five parallel strands surrounded by α-helices (Figure 1(b)). Superposition leads to root-mean-square deviation (rmsd) values of 1.4 Å for *Ec*ATC$_R$ and 1.6 Å for *Ec*ATC$_T$ (based on 272 C$^\alpha$ atoms of Protein Data Bank (PDB) entries 1D09 and 6AT1, respectively). The similarity with *Ec*ATC$_R$ is mainly due to several α-helical regions (helix numbers α_{1-7}). Furthermore, the polar and equatorial domains of the *Sa*ATC$_T$ catalytic chain are in closer contact compared to the *E. coli* T state. This is illustrated by the total number of interdomain ion pairs and hydrogen bonds, the *S. acidocaldarius* ATCase totaling 14 of them, *E. coli* T state only ten. More specifically, the E50–R167 and E50–R234 salt bridges are characteristic and necessary interactions for the *E. coli* R state.[22,23] In *Sa*ATC$_T$ the first ion pair is present (E47–R162), indicating yet another agreement with *Ec*ATC$_R$.

The region that undergoes the largest changes upon the T to R transition in *E. coli*, namely the 240s loop region, is closer to the *E. coli* T state (Figure 2(a)). In region 227–233 an extra Val is inserted (V228), and maximal distances between C$^\alpha$ atoms of T state *S. acidocaldarius* and *E. coli* T and R state ATCase of 2.3 Å and 11.8 Å, respectively, are found. In region 236–238 the maximal distances are 7.7 Å and 5.0 Å.

Other important differences between the *S. acidocaldarius* and *E. coli* C$^\alpha$ traces can be found mostly at the surface of the monomer. The N-terminal region 7–43 shows considerable differences with the corresponding region of the *E. coli* enzyme. Helix α_1 is shifted about 2 Å along its longitudinal axis. Loop region 33–37 (residues 36–41 in *E. coli*) has moved 1.5 Å away from its position at the C1–C2 interface in *E. coli*. The E37(C1)–K40(C2) salt bridge interaction between the catalytic subunits in the *E. coli* ATCase is thereby absent. Loop region 175–177 (residues 180–182 in *E. coli*) has a different conformation in *S. acidocaldarius*. An equivalent of the R17–D180 interdomain salt link of *E. coli* ATCase is absent.

The recently determined structure of the *P. abyssi* ATCase catalytic subunit in complex with PALA was found to be very similar to the *E. coli* R state catalytic subunit (rmsd 1.2 Å, based on 301 C$^\alpha$ atoms), except for the 240s region and the same α_1-helical region.[21] Apart from this region, the *S. acidocaldarius* catalytic subunit shows more similarity with the *P. abyssi* catalytic subunit than with that from *E. coli* in regions 110–116, 202–212 and 269–272.

All important active site residues of *E. coli* ATCase[15,24] are conserved in *S. acidocaldarius*. Most of them show good positional agreement. Exceptions are R221 and Q223, which are located further away from the active site (Figure 3). Since they move closer to the active site on forming the high affinity, high activity R state in *E. coli*, one could expect a larger conformational change taking

Table 1. Data collection and refinement statistics

Crystals	
Space group	*P*6322
Unit-cell parameters (Å, deg.)	$a = b = 132.85$, $c = 140.28$ $\alpha = \beta = 90°$, $\gamma = 120°$
Packing density V_M (Å^3 Da^{-1})	3.5
Solvent content (%)	64.3
Data collection	
Resolution range (Å)	20.0–2.59
Unique/measured reflections	23,141/201,357
Completeness (%)	99.7 (98.9)
R_{sym}[a] (%)	10.2 (42.1)
Average $I/\sigma(I)$	15.76 (2.15)
Refinement statistics	
R_{work} (%)	19.3
R_{free} (%)	23.9
Reflections in the test set (%)	5
No. of protein atoms	3499
No. of zinc atoms	1
No. of water molecules	93
Mean B-factor (Å^2)	
Catalytic subunit (main-chain/side-chain)	25.4/28.2
Regulatory subunit (main-chain/side-chain)	29.1/32.5
Solvent atoms	40.1
Rms bond length deviations (Å)	0.025
Rms bond angle deviations (deg.)	2.45
Rms dihedrals (deg.)	7.64
Ramachandran plot	
Most favored regions (%)	88.9
Additional allowed regions (%)	10.9
Generously allowed regions (%)	0.3
Disallowed regions (%)	0.0

Values between parentheses reflect the data in the 2.64–2.59 Å resolution shell.

[a] $R_{sym} = \sum_h \sum_i |I(h,i) - \langle I(h)\rangle| / \sum_h \sum_i I(h,i)$, where $I(h,i)$ is the intensity of the ith measurement of reflection h and $\langle I(h)\rangle$ is the average value over multiple measurements.

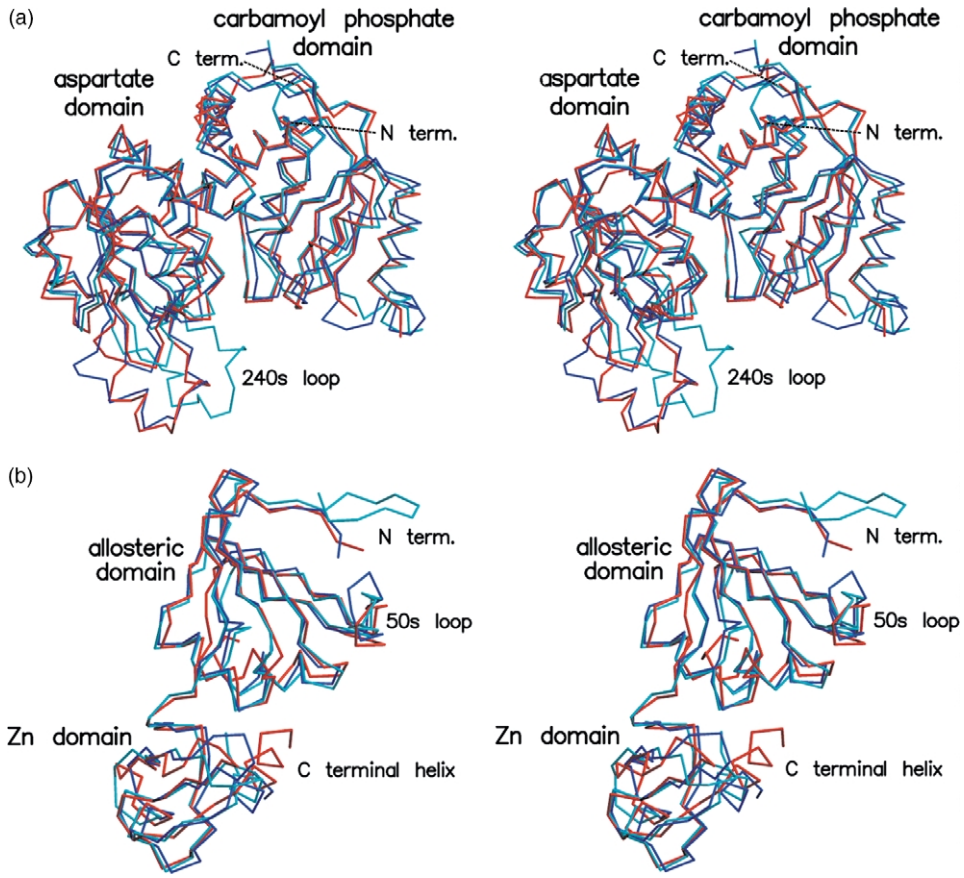


Figure 2. Stereoview of the superposition of C^{α} main-chain atoms of T state ATCase of *S. acidocaldarius* (red), *E. coli* (dark blue) and R state ATCase of *E. coli* (light blue): (a) catalytic chains; (b) regulatory chains.

place in *S. acidocaldarius* ATCase. However, the positioning of residue K126 of *Sa*ATC_T is not compatible with the large displacement of R234 seen in *E. coli* ATCase.

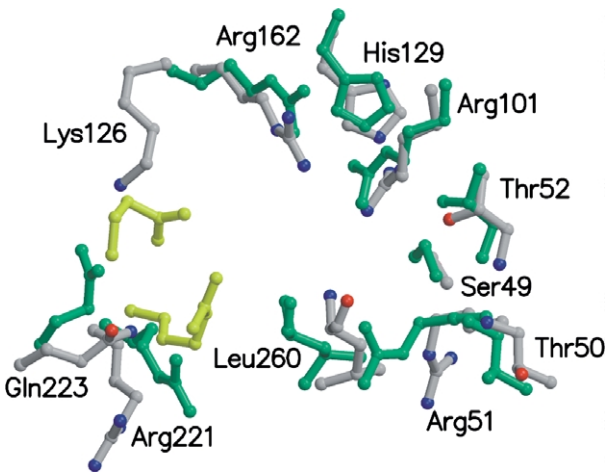


Figure 3. Superposition of the active site residues of the *S. acidocaldarius* (color-coded according to atom type) and *E. coli* (all atoms colored green) T state and residues R229 and Q231 (all atoms colored yellow) of the *E. coli* R state ATCase catalytic subunit. The 80s loop residues involved in catalysis were not included because of their absence in the *S. acidocaldarius* model.

Regulatory monomer structure

The *Sa*ATC_T regulatory subunit has 164 residues and shares 30% sequence identity with the *E. coli* ATCase regulatory subunit and consists mainly of β-sheets (Figure 1(c)). It is composed of an N-terminal allosteric domain (effector-binding domain) and a C-terminal Zn-binding domain. The core of each domain consists of an anti-parallel β-sheet with five and four β-strands, respectively. The Zn ion is coordinated by four Cys residues. Superposition leads to rmsd values of 1.7 Å for *Ec*ATC_T and 2.0 Å for *Ec*ATC_R (based on 142 C^{α} atoms of PDB entries 6AT1 and 1D09, respectively), showing that, in contrast with the catalytic chain, the regulatory subunit resembles more *Ec*ATC_T. The largest differences are situated in loop regions at the surface of the molecule and in a prolonged C-terminal helix which is shifted 4 Å towards the allosteric domain (Figure 2(b)).

The effector binding sites are well preserved between the two enzymes[14,17] (Figure 4). Although not directly involved in effector binding, residues 52–56 of *E. coli* (50s loop) define portions of the triphosphate and ribose subsites of the nucleotide binding site. The 50s loop changes conformation upon T to R transition and has been put forward as one of the ways in which the effector binding sites are coupled to the allosteric-Zn domain

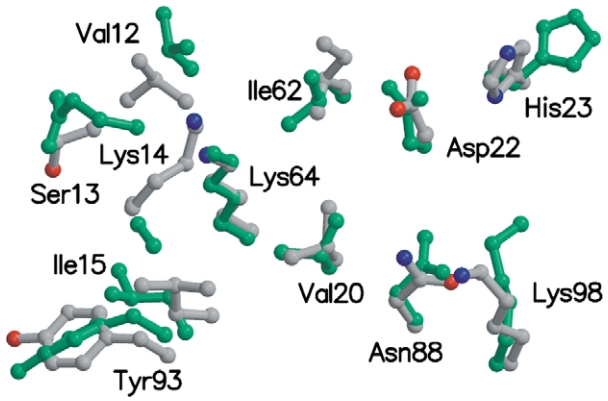


Figure 4. Superposition of the effector binding site residues of the *S. acidocaldarius* (color-coded according to atom type) and *E. coli* (all atoms colored green) T state ATCase regulatory subunit.

interface.[25] Residues 53–60 of the 50s region of $SaATC_T$ resemble $EcATC_R$ more closely (Figure 2(b)).

Quaternary structure

The *S. acidocaldarius* ATCase has the subunit composition $[2c_3 : 3r_2]$ (Figure 1(a)). The catalytic subunits are arranged as two sets of trimers (c_3) in complex with three sets of regulatory dimers (r_2). The dimension along the 3-fold axis is 6 Å longer than for $EcATC_T$ and 6 Å shorter than for $EcATC_R$. Despite the various resemblances between the $SaATC_T$ and $EcATC_R$, the overall subunit orientation is in good agreement with $EcATC_T$, so we can confidently consider the structure presented as a T state ATCase (Figure 5).

Upon comparison of the subunit interfaces, $SaATC_T$ has six extra C1–R2 types of interface, never observed in either $EcATC_R$ or $EcATC_T$.[12] All other interfaces previously described in *E. coli* are still present, but important deviations occur. What follows is a detailed description of the various types of interfaces. For the nomenclature of the interfaces, see Figure 1(a).

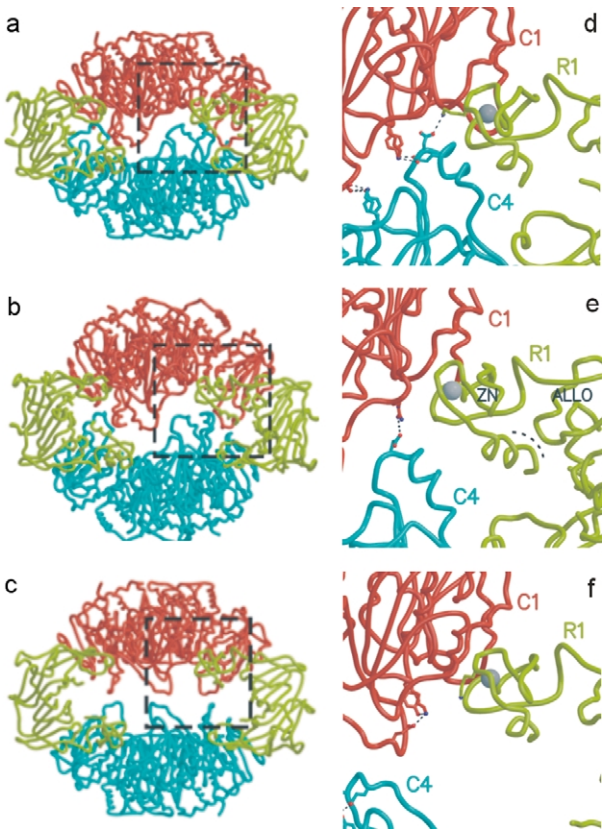


Figure 5. (a)–(c) Overview of *E. coli* T state, *S. acidocaldarius* T state and *E. coli* R state ATCase, respectively, seen along a molecular 2-fold axis with one regulatory dimer left out for clarity. The catalytic trimers are colored red and cyan, the regulatory dimers are in yellow. (d)–(f) Detailed view of important interactions situated at the C1–C4 and R1–C4 type interfaces of $EcATC_T$, $SaATC_T$ and $EcATC_R$ (shown boxed in (a)–(c)). The expanded hydrophobic interaction area between the Zn and allosteric domains of the regulatory chain of *S. acidocaldarius* ATCase is indicated by an arc. Residues implicated in important polar interactions (indicated by broken lines) are depicted with their side-chains. Zinc ions are represented by a grey sphere. Interactions indicated in (d): E239(C4/1)–K164(C1/4), E239(C4/1)–Y165(C1/4) and D236(C4)–K143(R1); (e): D229(C1/4)–R190(C4/1); (f) K164–E239 and K164–Y165 (both intrasubunit interactions).

C1–R2 type interfaces

A unique feature of the *S. acidocaldarius* ATCase is the presence of a C1–R2 type interface. The surface area buried in it is small (Table 2) and polar because of a salt bridge between residues R268(C1) and E125(R2). This salt link is part of an extensive ionic network of six residues, involved in 13 salt links and located in three interfaces (C1–C2; C1–R1; C1–R2, see Figure 6).

C1–C2 type interfaces

This type of interface, which connects two catalytic subunits from one trimer, contains the catalytic site and involves both the aspartate and carbamoyl phosphate binding domains of one monomer and the 80s loop of the adjacent catalytic monomer. In comparison with $EcATC_T$ and $EcATC_R$, the surface area buried in this interface is about 28% smaller (Table 2). In this calculation we took the flexible region encompassing residues 74–80 into account. This large difference can in part be attributed to the different conformation of the loop region following helix α_1.

While the $EcATC_R$ C1–C2 interface has fewer (14) polar interactions (salt bridges and hydrogen bonds) than $EcATC_T$ (19), $SaATC_T$ has only seven. However, five of these interactions are part of a six-residue ion pair network (Figure 6), whereas in *E. coli*, only a three-residue network is found here. It has been observed that the same network is

Table 2. Comparative analysis of the surface area buried in the different subunit and domain interfaces of the *S. acidocaldarius*, *E. coli* and *P. abyssi* ATCases

Interface type	*E. coli* T state	*E. coli* R state	*S. acidocaldarius*	*P. abyssi*
C1–R1 type interfaces				
Total (Å^2)	1333	1617	1356	
Apolar (Å^2)	651	795	769	
	(48.9)	(49.2)	(56.7)	
C1–C2 type interfaces				
Total (Å^2)	2176	2186	1357	2360
Apolar (Å^2)	1318	1271	827	1403
	(60.6)	(58.1)	(60.9)	(59.5)
R1–R6 type interfaces				
Total (Å^2)	1657	2105	1700	
Apolar (Å^2)	948	1251	1072	
	(57.2)	(59.4)	(63.1)	
C1–C4 type interfaces				
Total (Å^2)	717	310	500	
Apolar (Å^2)	246	247	246	
	(34.3)	(79.5)	(49.3)	
C1–R4 type interfaces				
Total (Å^2)	632	0	339	
Apolar (Å^2)	404	0 (0)	188	
	(63.9)		(55.4)	
C1–R2 type interfaces				
Total (Å^2)	0	0	177	
Apolar (Å^2)	0 (0)	0 (0)	56	
			(31.5)	
Catalytic chain–domain interfaces				
Total (Å^2)	2337	2587	2743	2601
Apolar (Å^2)	1205	1342	1540	1436
	(51.6)	(51.9)	(56.2)	(55.2)
Regulatory chain–domain interfaces				
Total (Å^2)	954	930	1021	
Apolar (Å^2)	704	674	835	
	(73.7)	(72.5)	(81.7)	

Values between parentheses are percentages of apolar buried surface area compared to total buried surface area.

extended to four residues in the *P. abyssi* ATCase structure.[21]

C1–R1 type interfaces

Most of the interactions between the catalytic and regulatory chains occur at the C1–R1 type interface, which is mainly composed of carbamoyl phosphate and Zn-binding domain regions. The amount of surface area buried in this interface is almost equal to that of *Ec*ATC$_T$; the hydrophobic content of this area is 18% higher (Table 2). Four aromatic residues (Y104(C), F110(C), Y117(R) and Y144(R)) are present in the *Sa*ATC$_T$ interface, compared to only one (Y140(R)) in *Ec*ATC$_T$. In contrast to a previous modeling study[17] in which fewer polar interactions were inferred in the C1–R1 interface, a total of 14 polar interactions was found compared with 12 of these in *Ec*ATC$_T$. The unique R state interactions between Zn-binding and aspartate binding domains are absent from the *Sa*ATC$_T$ C1–R1 interface; however, these domains are located closer to each other than in the *Ec*ATC$_T$ structure and salt link E292(C1)–R141(R1) is present.

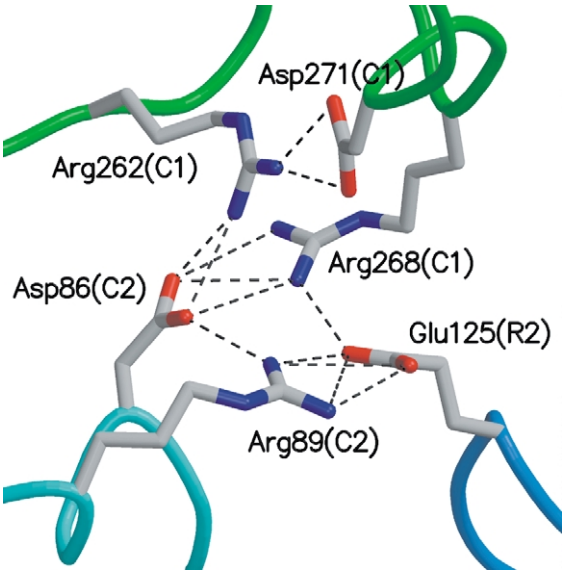


Figure 6. Six-residue ion pair network connecting three types of interfaces (C1–C2, C2–R2 and C1–R2) of the *S. acidocaldarius* ATCase.

R1–R6 type interfaces

The five-stranded β-sheets of the allosteric domains of two neighboring regulatory chains form a ten-stranded antiparallel β-sheet. Most of the polar interactions involved in this interface are hydrogen bonds connecting two β-strands. In contrast to *Ec*ATC$_T$ and *Ec*ATC$_R$ there are no salt bridges present.

C1–C4 type interfaces

The C1–C4 type interface in *E. coli* ATCase is a small interaction area that undergoes drastic changes upon the T to R transition and has major influence on the stability of the T and R state.[26–28] Although structurally more comparable to *Ec*ATC$_T$ than to *Ec*ATC$_R$, *Sa*ATC$_T$ shows a different organization. Large conformational differences are known to occur in the 240s loop region, which contains residues important in stabilizing *Ec*ATC$_T$. The *E. coli* E239(C1)–K164(C4) and E239(C1)–Y165(C4) interactions are replaced in *S. acidocaldarius* by D229(C1)–R190(C4) interactions (Figure 5). While residues E239 and D229 are located at almost equivalent positions, residues K164 (K159 in *S. acidocaldarius*) and R190 are not positionally related. There is, however, a structural connection between the regions that contain residues K159 and R190, namely a hydrogen bond between the main-chain atoms of L158 and L187 and a three-residue ionic network between R169, E192 and E196, not present in *Ec*ATC$_{T/R}$.

R1–C4 type interfaces

The R1–C4 interface is observed in *Ec*ATC$_T$, but not in *Ec*ATC$_R$. It is considered important in stabilizing the T state.[29] The R1–C4 type interface is also present in *Sa*ATC$_T$ but it is about 50% smaller (Table 2) and the number of apolar van der Waals contacts is reduced. An important salt bridge in *Ec*ATC$_T$ (D236(C4)–K143(R1)) links the 240s loop region with the C-terminal region of the regulatory chain.[29,30] This salt bridge, as well as other polar interactions, is absent from this interface (Figure 5).

Analysis of thermostability

Amino acid composition

Of the 20 amino acids, the residues Asn, Gln, Met and Cys can be classified as thermolabile due to their tendency to undergo deamidation or oxidation at high temperatures.[31] A comparison of the amino acid composition of *S. acidocaldarius* and *E. coli* ATCase shows that only Gln occurs significantly less in the thermostable ATCase. As reported for the *P. abyssi* catalytic chain,[21] more charged residues are found in *S. acidocaldarius* ATCase, especially residues Lys and Glu. In accordance with the general tendency,[32] most of these extra charges are found at the surface of *Sa*ATC$_T$. The catalytic chain has significantly more Tyr residues. As discussed above, several of them are found in the C1–R1 type interface of *Sa*ATC$_T$. Although Pro is known to potentially hamper thermal unfolding by conferring more rigidity to surface loops,[33–35] there are no important differences in proline content between the two enzymes.

Secondary structure elements

A small difference in α-helical content can be found for the regulatory subunits (Table 3). This higher α-helical content can be attributed to the extended C-terminal helix. In the catalytic monomers there are only small local differences in secondary structure.

α-Helix distortion can be a factor in protein stability,[36] e.g. the presence of proline residues or β-branched residues (Val, Ile, Leu) can distort α-helix geometry. The numbers of these residues present in α-helices are comparable in *E. coli* and *S. acidocaldarius*. Helix dipole stabilization does not seem to be responsible for the higher thermal stability of *S. acidocaldarius* ATCase either.

Loops and N and C termini are usually the regions with the highest thermal factors in a protein crystal structure and it has been stated that they are likely to unfold first during thermal denaturation.[37] In the *Sa*ATC$_T$ catalytic chain, the N and C-terminal regions (located on the surface of the molecule), are six and four residues shorter, respectively, than in *E. coli*. The N and C-terminal regions of the regulatory chain are three and six residues shorter in *E. coli*. In contrast to the *P. abyssi* catalytic subunit,[21] loop shortening (compared to the *E. coli* ATCase) is not a general feature of the *S. acidocaldarius* ATCase (Figure 7).

Table 3. Comparative analysis of *S. acidocaldarius*, *E. coli* and *P. abyssi* ATCase subunit structures

	E. coli			
	T state	R state	*S. acidocaldarius*	*P. abyssi*
Catalytic subunit				
Secondary structure				
α-Helices (%)	35	31	37	35
Pro content	2	1	1	2
Ile, Leu, Val content	29	28	28	28
β-Strands (%)	15	15	16	15
Hydrogen bonds				
Total	290	295	269	278
Per residue	0.93	0.95	0.90	0.90
Salt bridges				
Total	14	31	26	40
Per residue	0.045	0.100	0.087	0.130
Aromatic pairs	2	2	5	4
VDW-volume (Å^3)	31,870	31,730	31,280	32,480
Packing density	0.74	0.75	0.73	0.71
Total ASA (Å^2)	13,271	12,839	13,358	13,651
Apolar ASA (Å^2)	7370 (55.5)	7106 (55.2)	6973 (52.2)	7143 (52.3)
Regulatory subunit				
Secondary structure				
α-Helices (%)	12	8	15	
Pro content	0	0	0	
Ile, Leu, Val content	6	4	7	
β-Sheets (%)	33	33	37	
Hydrogen bonds				
Total	111	94	116	
Per residue	0.73	0.61	0.71	
Salt bridges				
Total	14	15	15	
Per residue	0.092	0.098	0.091	
Aromatic pairs	0	0	0	
VDW-volume (Å^3)	15,110	15,880	15,580	
Packing density	0.75	0.76	0.75	
Total ASA (Å^2)	8558	9596	8901	
Apolar ASA (Å^2)	4448 (52.0)	5316 (55.4)	5095 (57.2)	

ASA, accessible surface area; values between parentheses are percentages of total ASA.

Packing density and cavities

The packing densities of *Sa*ATC$_T$ catalytic and regulatory monomers were found to be very similar to those of *Ec*ATC$_T$, *Ec*ATC$_R$, and *P. abyssi* ATCase (Table 3). The search of probe-accessible cavities with the program VOIDOO led to the

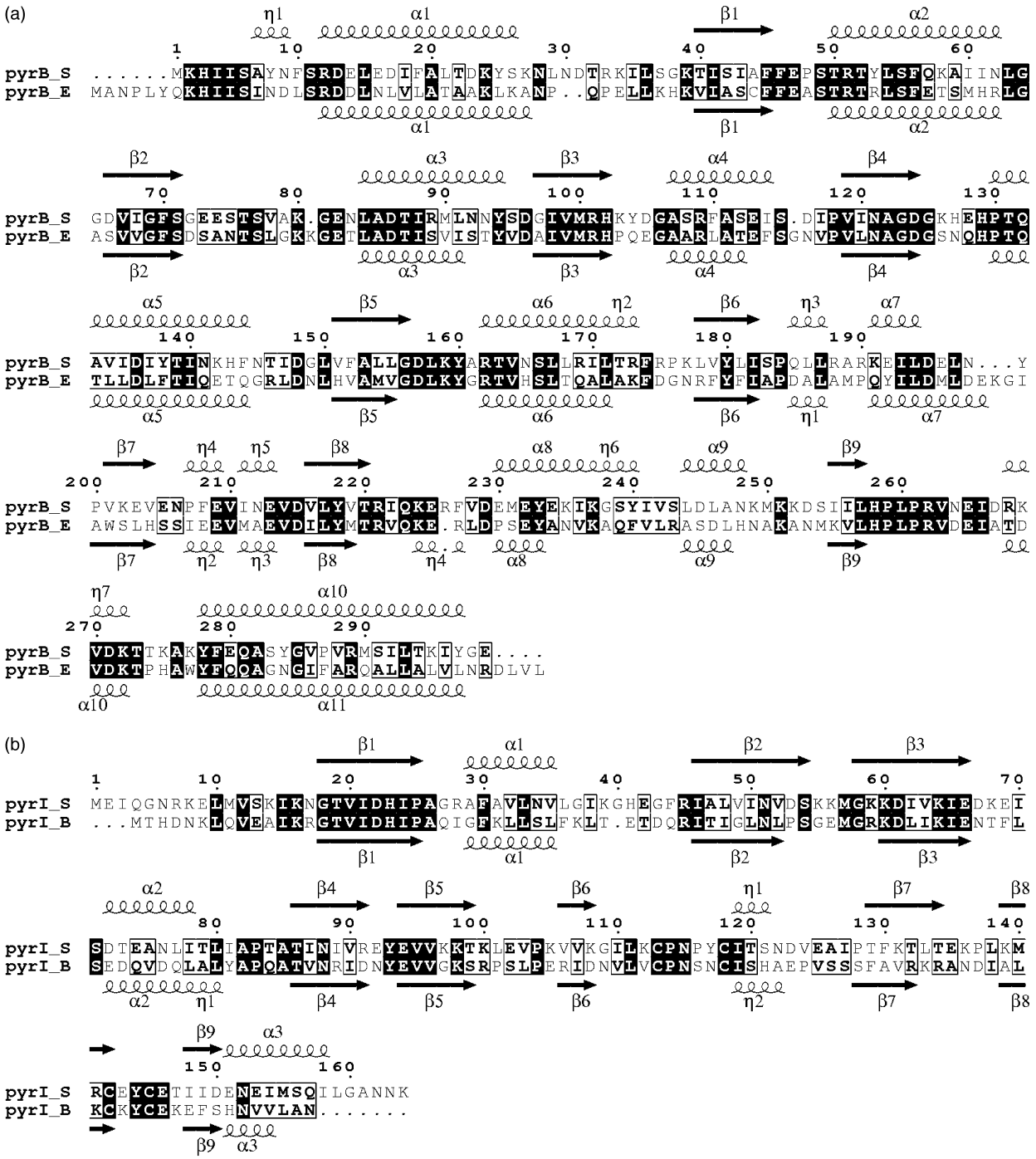


Figure 7. Sequence alignment of *S. acidocaldarius* and *E. coli* ATCase. Residues in boxes are conserved, residues shaded in black are strictly conserved. The secondary structure is indicated: β-strands are shown as arrows, 3_{10}-helices are represented by thinner coils than α-helices and labeled separately. (a) Catalytic chains of *S. acidocaldarius* (pyrB_S.) and *E. coli* (pyrB_E.). (b) Regulatory chains of *S. acidocaldarius* (pyrB_S.) and *E. coli* (pyrB_E.).

identification of one cavity of 38.5 Å^3 in the 240s region of *Ec*ATC$_T$ (24.1 Å^3 in *Ec*ATC$_R$) and a cavity with a volume of 47 Å^3 at the C1–C4 type interface of *Ec*ATC$_T$. In *Sa*ATC$_T$ a different C$^\alpha$ trace of residues 236–238, combined with the replacement by residues having bulkier side-chains (V160L, V230I, L235F, S238M, V243I and V248I, *Ec*ATC numbering), results in denser packing and thus in the absence of the first cavity. The cavity volume calculated for the *P. abyssi* catalytic trimer is similar to that of *Ec*ATC$_T$ and *Ec*ATC$_R$. Nevertheless, closer examination reveals that the cavity cited above is absent from *P. abyssi* ATCase (its Y248 side-chain is pointing into the cavity). The second cavity mentioned is lined by the residues that form the typical interactions of the C1–C4 interface in *Ec*ATC$_T$. Due to the altered C1–C4 interface in *Sa*ATC$_T$, this cavity is absent.

Hydrogen bonds, ion pairs and aromatic interactions

The total number of hydrogen bonds of the *S. acidocaldarius* and *E. coli* ATCase is not significantly different. Whereas the regulatory subunits contain an equal number of ion pairs, the catalytic subunits contain twice as many ion pairs[26] compared to *Ec*ATC$_T$,[14] but still significantly less than *Ec*ATC$_R$.[31] More important, however, is the presence of ion pair networks.[38,39] In the *Sa*ATC$_T$ dodecamer, 68% of the salt links are part of ion pair networks, in contrast with only 19% for *Ec*ATC$_T$ and 59% for *Ec*ATC$_R$.

Five interacting aromatic pairs are detected in the *Sa*ATC$_T$ catalytic monomer. In comparison, the *Ec*ATC$_T$ and *Ec*ATC$_R$ and *P. abyssi* catalytic subunits contain two and four aromatic pairs, respectively (Table 3).

Accessible surface area (ASA)

A detailed comparison of the buried (total and apolar) surface areas of the different subunit interfaces can be found in Table 2. In order to obtain a better estimate of the degree of interaction between the domains within each subunit, an approximating calculation of the buried surface area between the domains was made. The result points out that the aspartate and carbamoyl phosphate domains have a more intimate contact compared to *E. coli* T state ATCase. Furthermore, the *S. acidocaldarius* ATCase has 19% more buried apolar surface area between its regulatory chain domains than the unliganded *E. coli* ATCase. This can be attributed to the extended C-terminal helix interacting with the allosteric domain instead of the R1–C4 type interface. The intimate contact between the allosteric and Zn-binding domains is further emphasized by the average crystallographic *B*-factors, which are more comparable for the two domains in *Sa*ATC$_T$ (32.4 Å^2 and 28.3 Å^2 compared to 45.9 Å^2 and 26.5 Å^2 for *Ec*ATC$_T$).

Discussion

In a previous study[17] the kinetic behavior of *S. acidocaldarius* ATCase, including Michaelian saturation kinetics for carbamoyl phosphate and homotropic cooperative interactions for aspartate, was found to be similar to that of the *E. coli* ATCase. In *E. coli* ATCase, two main structural and functional states are generally accepted to be present: a low affinity, low activity T state (*Ec*ATC$_T$) and a high affinity, high activity R state (*Ec*ATC$_R$).[40] Based upon the overall quaternary structure and the absence of ligands known to induce the R state, the crystal structure of the unliganded ATCase of *S. acidocaldarius* described here is thought to represent the functional T state of *Sa*ATC$_T$.

Comparison of the *Sa*ATC$_T$ and *Ec*ATC$_T$ structures shows that their quaternary organization is similar, i.e. [$2c_3 : 3r_2$]. Important structural differences occur in the N-terminal and 240s regions of the catalytic chains and the C-terminal region of the regulatory chains. The differences in the N-terminal region of the former chains are partially responsible for the weakening of the interactions between the subunits within the catalytic trimers (c_3), explaining why *S. acidocaldarius* ATCase catalytic chains mostly remain monomeric in the absence of regulatory chains. This distinguishes the enzyme from the *E. coli* ATCase. A considerable transition to the trimeric state is observed in the presence of carbamoyl phosphate.[17] In our crystal structure the 80s loop could not be modeled due to a lack of electron density. Possibly, carbamoyl phosphate can move the 80s loop in an active conformation at the interface between monomers, thus stabilizing the monomer–monomer interactions.

Upon the T to R state transition in *E. coli*, the two domains of the catalytic chains move closer to each other, bringing the active site residues in their high affinity, high activity conformation. It has been shown for *E. coli* ATCase that weakening of the C1–C2 type interface promotes this domain closure.[41] The C1–C2 type interface is weakened in the *Sa*ATC$_T$ structure, bringing the two domains closer to each other. Therefore, even though the overall quaternary structure is consistent with a T state ATCase, the *Sa*ATC$_T$ catalytic chain conformation is intermediate between *Ec*ATC$_T$ and *Ec*ATC$_R$.

Large conformational differences between the 240s regions of *S. acidocaldarius* and *E. coli* ATCase have marked effects on the C1–C4 and R1–C4 type interfaces. It has been proposed that because of steric hindrance, domain closure in one catalytic chain in *E. coli* cannot occur without some change in the quaternary structure of the enzyme,[6] making the elongation of the enzyme a necessary feature of the allosteric transition. This elongation requires the breaking of the T state intersubunit interactions involving D236(C) and E239(C). When the side-chain of either of these residues is replaced by uncharged substitutes, the net results are the inability of the *E. coli* enzyme to exist in the T state, and the formation of a new structural state that is neither T nor R.[27] In *Sa*ATC$_T$ the important interactions E239(C1)–K164(C4), E239(C1)–Y165(C4) and D236(C4)–K143(R1) are absent and although a shift can be seen at the C1–C4 type interfaces towards the *E. coli* R state, the *S. acidocaldarius* holoenzyme still shows homotropic cooperativity between the catalytic sites, which can be explained by the existence of two distinct functional states. The unique D229(C1)–R190(C4) salt bridges therefore are likely to be major contributors to the stabilization of the low affinity, low activity state of *S. acidocaldarius* ATCase. A unique feature observed in this crystal structure is the existence of a C1–R2 type of interface which is a potentially stabilizing factor for the T state of *S. acidocaldarius* ATCase. An extensive ion pair network connects

this novel interface with three other interface types (C1–C2, C1–R1 and C1–R2).

The R1–C4 type of interface is important for homotropic cooperativity in *E. coli* ATCase.[29,30] More specifically, it is thought that the intersubunit link involving D236(C) is important for the stabilization of the 240s loop in its tight-state position. In *Sa*ATC$_T$, this link is not present due to structural changes in the C-terminal region of the regulatory chain. This may again explain why the 240s loop in *Sa*ATC$_T$ cannot be held in the same position as in *Ec*ATC$_T$, and why the catalytic chain domains are located closer to each other in the unliganded enzyme.

Although similar in kinetic behavior, *S. acidocaldarius* ATCase exhibits a different regulation pattern compared to the *E. coli* enzyme. All nucleoside triphosphate effectors activate the *Sulfolobus* holoenzyme, whereas *E. coli* ATCase is synergistically feedback inhibited by the end products CTP and UTP, and activated only by ATP.[42–45] Studies on hybrid enzymes built from regulatory and catalytic polypeptides of ATCases with divergent functional characteristics established that the regulatory subunits determine the type of allosteric regulatory properties exhibited by the holoenzyme.[46–50] For a signal to propagate from the nucleotide binding site to the active site, it must pass through different interfaces. Mutagenesis studies have assessed the importance of different regions in these interfaces in the modulation of the heterotropic effects, but no clear signal transduction pathways have been described.[51–55] When comparing the structures of *Sa*ATC$_T$ and *Ec*ATC$_{T/R}$, considerable differences are found in all of these regions, especially in the allosteric-zinc domain interface and the contact regions between the regulatory and catalytic chains (C1–R1 and R1–C4 type interfaces). It has been described that replacing D236(C) with an alanine in *E. coli* ATCase results in the elimination of heterotropic interactions, which suggests that the D236(C)–K143(R) interaction at the R1–C4 interface may function to transmit heterotropic signals.[29] Furthermore, it has been suggested that the interactions between residues 146 to 149 of the regulatory chain and the 240s loop region of the catalytic chain might be important in the modulation of the ATP effect.[51,52] Since these interactions are weakened or lost in *Sa*ATC$_T$, important differences in effector signaling pathways are to be expected. Especially, the novel C1–R2 type interface could have an important role.

Any direct signal from the nucleotide binding site to the active site must pass through the allosteric-zinc domain interface. In *Sa*ATC$_T$ and *Ec*ATC$_T$, different residues are interacting in this mainly hydrophobic interface. A particular deletion mutant of *E. coli* ATCase (rΔ[152–153]) is not inhibited by CTP, and it has been suggested that this is due to the formation of a new intrachain interaction between the C-terminal region and the allosteric domain of the regulatory chain. CTP could thus inhibit the reorientation of the two domains of the regulatory chain that allows the formation of this interaction. In the mutant, the effect of CTP would be pre-empted as a result of the deletion.[52,56,57] In *Sa*ATC$_T$ the weakening of the R1–C4 interface in favor of the interactions between the C-terminal helical region and the allosteric domain could produce a similar effect, pre-empting or even reversing the heterotropic effect of CTP.

Thermostability

ATCases are generally thermostable enzymes; *E. coli* ATCase loses only 25% activity after incubation for six hours at 60 °C, although the catalytic trimer is less thermostable and loses 80% of its activity.[58] Since *S. acidocaldarius* is a hyperthermophilic archaeon thriving at temperatures around 80 °C, its ATCase is highly thermostable.[17] Despite the high thermostability, we find that the C1–C2 type is drastically weakened. This contrasts with the recently reported findings on the catalytic subunit trimer structure of *P. abyssi* ATCase, where this interface is stronger than in *E. coli*. Therefore, the stability of the holoenzyme has to be achieved through the association of the catalytic and the regulatory subunits whose main interaction area is situated at the C1–R1 type interfaces. A comparison of this type of interface in the *Ec*ATC$_T$ and *Sa*ATC$_T$ shows that the apolar fraction of the buried surface area is higher in *Sa*ATC$_T$, in agreement with a stabilized interface.[59–62] In *Ec*ATC$_{T/R}$, the regulatory chains are considerably less thermostable than the catalytic subunits.[58] In comparison with the catalytic chains, the regulatory chains are indeed more loosely organized in terms of tertiary structure. Whereas the catalytic chain domains are connected by two polypeptide chains and several non-covalent interactions, the regulatory chain domains are connected by one polypeptide chain and contain considerably less non-covalent interdomain interactions. Consequently, thermo-adaptation of ATCases with regulatory chains should have required more substantial modifications of the regulatory chains than of the catalytic chains.[18] The *Sa*ATC$_T$ and *Ec*ATC$_T$ regulatory chains are indeed more divergent than their catalytic chains.

The regulatory chain C terminus of *Sa*ATC$_T$ is extended in comparison to *Ec*ATC$_T$. This region is stabilized in *Sa*ATC$_T$ due to the presence of extra secondary and tertiary structural features at the allosteric-zinc domain interface, including an increase in buried hydrophobic surface. Due to this rearrangement of the C terminus, the R1–C4 interface is substantially weakened compared to *Ec*ATC$_T$. This does not necessarily reduce the overall thermostability, because weakening of the R1–C4 interface has been shown to be possible without decreasing the overall thermostability of *E. coli* ATCase.[29] Overall, we can therefore hypothesize that the C1–R1, the allosteric-zinc and the novel C1–R2 type interfaces are important for thermostability, whereas C1–C2 and R1–C4 interfaces can even be weaker.

We further analyzed a number of other factors

considered to be important in general thermostability of enzymes. Every enzyme seems to have a unique combination of strategies,[63–68] and the same is true for the *Sa*ATC$_T$ enzyme. In particular, there is a decreased occurrence of "thermolabile" Gln residues, and a higher occurrence of charged residues,[37,39] leading to an increase in ion pair networks.[38,39] *Ec*ATC$_R$, but not *Ec*ATC$_T$, has more ion pairs than *Sa*ATC$_T$, but the ion pairs of *Sa*ATC$_T$ are more frequently involved in the formation of networks, stressing the importance of ion pair networks compared to individual pairs.[38,39] The catalytic subunit structure of *P. abyssi* ATCase demonstrates an even more prominent role for charge–charge interactions.[21] Further adaptions seem to be an increased number of tyrosine residues resulting in an increased number of aromatic pairs, and the absence of cavities in *Sa*ATC$_T$.

Because of the complexity of the interplay between its structure and function, adaptation to extreme temperatures might not have been compatible with the maintenance or rather the emergence of the mechanisms of action present in *E. coli* ATCase. Significant structural differences were found in regions important for its allosteric behavior, located often at interfaces between domains or subunits. Some of these regions require a certain degree of flexibility, a property that is counterproductive in adaptation of proteins to high temperatures. It is therefore possible that the adaptation of *S. acidocaldarius* ATCase resulted in the observed shift of the unliganded T state enzyme towards the R state and the differences in regulatory behavior. Elucidation of the crystal structures of the R state and effector-liganded ATCases of *S. acidocaldarius*, together with thermostability studies on mutant enzymes, may provide further answers.

Materials and Methods

Crystallization

Protein purified following a method described earlier[17] was used to grow crystals with the hanging-drop, vapour-diffusion technique at 293 K. Protein solution (1 μl) (10 mg/ml in 100 mM sodium phosphate buffer (pH 7.3), 2 mM β-mercaptoethanol, 0.2 mM EDTA) was mixed with 1 μl of reservoir solution (100 mM sodium acetate (pH 4), 30% saturated ammonium sulfate). Within three to four days large crystals (0.6 mm × 0.3 mm) with the shape of a hexagonal prism appeared, which often had smoothed edges that gave the crystals an egg-like appearance. Crystals obtained at a higher pH diffracted to lower resolution.

Data collection

Selected crystals were transferred to a solution containing 20% (v/v) glycerol in mother liquor prior to flash-cooling in liquid nitrogen. The X-ray diffraction data were collected on a MAR CCD (MAR Research) detector system on the XRD synchrotron beamline at the ELETTRA synchrotron in Trieste using radiation with a wavelength of 1.00 Å. Data collection was performed at 100 K on a single crystal. A high-resolution dataset was collected to 2.5 Å, followed by a low-resolution dataset to 2.8 Å. Data processing was carried out using the programs DENZO and SCALEPACK.[69] The diffraction intensities were scaled and merged to 2.6 Å resolution using the program SCALEPACK. Intensities were converted into amplitudes using the program TRUNCATE from the CCP4 suite.[70] The full data collection statistics are presented in Table 1.

Structure determination

One catalytic subunit of the *E. coli* ATCase (PDB 8AT1) served as a starting model for solving the structure by molecular replacement. Rotation and translation searches have been performed using the program AMoRe[71] and the best solution calculated had a correlation coefficient of 54.9 and an *R*-factor of 49.8% in the resolution range 20–3.5 Å. Cycles of rigid body refinement were performed using the program Refmac of the CCP4 package.[70] Visual inspection of the resulting maps revealed interpretable electron density for most secondary structure elements of the catalytic subunit. Alternating rounds of simulated annealing refinement according to the slow-cooling protocol of CNS[72] and model building using TURBO-FRODO[73] resulted in maps with a clear density in the regions corresponding to the ATCase regulatory chain. In the final stages of refinement, Refmac was used for ARP-water cycling to refine the solvent structure, in combination with TLS refinement.[74] Although the density is good for both subunits, exceptions occur in solvent-exposed loops. Specifically the electron density in residues 28–31 and 225–226 of the catalytic subunit and in residues 11–14 and 38–44 of the regulatory subunit is weak. Residues 74–80 of the catalytic subunit and residues 1–10, 40–41 and 161–164 of the regulatory subunit were not modeled due to absence of electron density.

Structure analysis

PDB entries 6AT1 and 1D09 were chosen as T and R state models of the *E. coli* ATCase for molecular comparison with *S. acidocaldarius* ATCase. PDB entry 1ML4 was used for molecular comparison with the *P. abyssi* ATCase catalytic subunit. Structure validation was performed with the program PROCHECK;[75] secondary structure was assigned using DSSP.[76] Hydrogen bonds between protein atoms were calculated with the program HBPLUS[77] with the default parameters for distance and angles. The presence of salt-bridges was inferred when Asp or Glu side-chain carbonyl oxygen atoms were found to be within a 4.0 Å distance from the nitrogen atoms in Arg, Lys and His side-chains. We defined an interacting aromatic pair as one for which the distance between phenyl ring centroids was less than 7 Å and with dihedral angles within 30° of 90°.[78] Accessible surface areas were calculated with the programs ACCESS and ACCFMT.[79] Probe-accessible cavity volumes were searched with the program VOIDOO[80] using a 1.2 Å probe radius and a minimum volume of 10 Å^3, corresponding to "real" cavities as described.[81] VOIDOO was also used for van der Waals volume determination of the proteins. The real volumes were calculated by the Voronoi procedure,[82] using the programs ACCESS and VOLUME.[79] The packing density was calculated as the ratio of the van der Waals volume to the real volume.

Figures were prepared with MOLSCRIPT[83] and rendered with Raster3D.[84]

Protein Data Bank accession number

The coordinates of the *S. acidocaldarius* T state ATCase have been deposited in the RCSB Protein Data Bank, with accession code 1PG5.

Acknowledgements

This research has been supported by the Fund for Scientific Research-Flanders (grants 3G006896 and 3G044899). D. De V. is a research fellow of the same institution. We acknowledge the use of the XRD beamline at the ELETTRA synchrotron, Trieste.

References

1. Jones, M. E., Spector, L. & Lipmann, F. (1955). Carbamyl phosphate. The carbamyl donor in enzymatic citrulline synthesis. *J. Am. Chem. Soc.* **77**, 819–820.
2. Reichard, P. & Hanshoff, G. (1956). Aspartate carbamyl transferase from *Escherichia coli*. *Acta Chem. Scand.* **10**, 548–560.
3. Allewell, N. M. (1989). *Escherichia coli* aspartate transcarbamoylase: structure, energetics, and catalytic and regulatory mechanisms. *Annu. Rev. Biophys. Biochem. Chem.* **18**, 71–92.
4. Davidson, J. N., Chen, K. C., Jamison, R. S., Musmanno, L. A. & Kern, C. B. (1993). The evolutionary history of the 1st 3 enzymes in pyrimidine biosynthesis. *Bioessays*, **15**, 157–164.
5. Hervé, G. (1989). Aspartate transcarbamylase from *Escherichia coli*. In *Allosteric Enzymes* (Hervé, G., ed.), pp. 62, CRC Press, Boca Raton. FL.
6. Kantrowitz, E. R. & Lipscomb, W. N. (1990). *Escherichia coli* aspartate transcarbamylase: the molecular basis for a concerted allosteric transition. *Trends Biochem. Sci.* **15**, 53–59.
7. Lipscomb, W. N. (1994). Aspartate transcarbamylase from *Escherichia coli*: activity and regulation. *Advan. Enzymol. Relat. Areas Mol. Biol.* **68**, 67–151.
8. Stevens, R. C., Chook, Y. M., Cho, C. Y., Lipscomb, W. N. & Kantrowitz, E. R. (1991). *Escherichia coli* aspartate carbamoyltransferase: the probing of crystal structure analysis via site-specific mutagenesis. *Protein Eng.* **4**, 391–408.
9. Konigsberg, W. H. & Henderson, L. (1983). Amino acid sequence of the catalytic subunit of aspartate transcarbamoylase from *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **80**, 2467–2471.
10. Hoover, T. A., Roof, W. D., Foltermann, K. F., O'Donovan, G. A., Benconi, D. A. & Wild, J. R. (1983). Nucleotide sequence of the structural gene (*pyrB*) that encodes the catalytic polypeptide of aspartate transcarbamoylase of *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **80**, 2462–2466.
11. Schachman, H. K., Pauza, C., Navre, M., Karels, M. J., Wu, L. & Yang, Y. R. (1984). Location of amino acid alterations in mutants of aspartate transcarbamoylase: structural aspects of interallelic complementation. *Proc. Natl Acad. Sci. USA*, **81**, 115–119.
12. Krause, K. L., Volz, K. W. & Lipscomb, W. N. (1987). The 2.5 Å structure of aspartate carbamoyltransferase complexed with the bisubstrate analogue *N*-phosphonacetyl-L-aspartate. *J. Mol. Biol.* **193**, 527–553.
13. Meighen, E. A., Pigiet, V. & Schachman, H. K. (1970). Hybridization of native and chemically modified enzymes, III: the catalytic subunits of aspartate transcarbamylase. *Proc. Natl Acad. Sci. USA*, **67**, 234–241.
14. Stevens, R. C., Gouaux, J. E. & Lipscomb, W. N. (1990). Structural consequences of effector binding to the T state of aspartate carbamoyltransferase: crystal structures of the unligated and ATP- and CTP-complexed enzymes at 2.6 Å resolution. *Biochemistry*, **29**, 7691–7701.
15. Jin, L., Stec, B., Lipscomb, W. N. & Kantrowitz, E. R. (1999). Insights into the mechanisms of catalysis and heterotropic regulation of *Escherichia coli* aspartate transcarbamylase based upon a structure of the enzyme complexed with the bisubstrate analogue *N*-phosphonacetyl-L-aspartate at 2.1 Å. *Proteins: Struct. Funct. Genet.* **37**, 729–742.
16. Honzatko, R. B. & Lipscomb, W. N. (1982). Interactions of phosphate ligands with *Escherichia coli* aspartate carbamoyltransferase in the crystalline state. *J. Mol. Biol.* **160**, 265–286.
17. Durbecq, V., Thia-Toong, T.-L., Charlier, D., Villeret, V., Roovers, M. *et al.* (1999). Aspartate carbamoyltransferase from the thermoacidophilic archaeon *Sulfolobus acidocaldarius*: cloning, sequence analysis, enzyme purification and characterization. *Eur. J. Biochem.* **264**, 233–241.
18. Purcarea, C., Hervé, G., Ladjimi, M. M. & Cunin, R. (1997). Aspartate transcarbamylase from the deep-sea hyperthermophilic archaeon *Pyrococcus abyssi*: genetic organization, structure and expression in *Escherichia coli*. *J. Bacteriol.* **179**, 4143–4157.
19. Hack, E. S., Vorobyova, T., Sakash, J. B., West, J. M., Macol, C. P., Hervé, G. *et al.* (2000). Characterization of the aspartate transcarbamoylase from *Methanococcus jannaschii*. *J. Biol. Chem.* **275**, 15820–15827.
20. Vitali, J., Vorobyova, T., Webster, G. & Kantrowitz, E. R. (2000). Crystallization and structure determination of the catalytic trimer of *Methanococcus jannaschii* aspartate transcarbamoylase. *Acta Crystallog. sect. D*, **56**, 1061–1063.
21. Van Boxstael, S., Cunin, R., Khan, S. & Maes, D. (2003). Aspartate transcarbamylase from the hyperthermophilic archaeon *Pyrococcus abyssi*: thermostability and 1.8 Å resolution crystal structure of the catalytic subunit complexed with the bisubstrate analogue *N*-phosphonacetyl-L-aspartate. *J. Mol. Biol.* **326**, 203–216.
22. Middleton, S. A., Stebbins, J. W. & Kantrowitz, E. R. (1989). A loop involving catalytic chain residues 230–245 is essential for the stabilization of both allosteric forms of *Escherichia coli* aspartate transcarbamoylase. *Biochemistry*, **28**, 1617–1626.
23. Newton, C. J. & Kantrowitz, E. R. (1990). The importance of domain closure for the allosteric transition in *Escherichia coli* aspartate transcarbamoylase. *Biochemistry*, **29**, 1444–1454.
24. Ke, H.-M., Lipscomb, W. N., Cho, Y. & Honzatko, R. B. (1988). Complex of *N*-phosphonacetyl-L-aspartate with aspartate carbamoyltransferase; X-ray refinement, analysis of conformational changes and catalytic and allosteric mechanisms. *J. Mol. Biol.* **204**, 725–747.
25. Gouaux, J. E., Stevens, R. C. & Lipscomb, W. N.

(1990). Crystal structures of aspartate carbamoyltransferase ligated with phosphonoacetamide, malonate and CTP or ATP at 2.8 Å resolution and neutral pH. *Biochemistry*, **29**, 7702–7715.
26. Ladjimi, M. M. & Kantrowitz, E. R. (1988). A possible model for the concerted allosteric transition in *Escherichia coli* aspartate transcarbamylase as deduced from site-directed mutagenesis studies. *Biochemistry*, **27**, 276–283.
27. Tauc, P., Vachette, P., Middleton, S. A. & Kantrowitz, E. R. (1990). Structural consequences of the replacement of Glu239 by Gln in the catalytic chain of *Escherichia coli* aspartate transcarbamylase. *J. Mol. Biol.* **214**, 327–335.
28. Hsuanyu, Y., Wedler, F. C., Middleton, S. A. & Kantrowitz, E. R. (1989). Kinetic consequences of site-specific mutation of Glu239···Gln in *E. coli* aspartate transcarbamylase: comparison with catalytic subunits and Phe240 mutant enzyme. *Biochem. Biophys. Acta*, **995**, 54–58.
29. Newton, C. J. & Kantrowitz, E. R. (1990). The regulatory subunit of *Escherichia coli* aspartate carbamoyltransferase may influence homotropic cooperativity and heterotropic interactions by a direct interaction with the loop containing residues 230–245 of the catalytic chain. *Proc. Natl Acad. Sci. USA*, **87**, 2309–2313.
30. Eisenstein, E., Markby, D. W. & Schachman, H. K. (1990). Heterotropic effectors promote a global conformational change in aspartate transcarbamoylase. *Biochemistry*, **29**, 3724–3731.
31. Russell, R. J. M., Ferguson, J. M. C., Haugh, D. W., Danson, M. J. & Taylor, G. I. (1997). The crystal structure of citrate synthase from the hyperthermophilic archaeon *Pyrococcus furiosus* at 1.9 Å resolution. *Biochemistry*, **36**, 9983–9994.
32. Cambillau, C. & Claverie, J. M. (2000). Structural and genomic correlates of hyperthermostability. *J. Biol. Chem.* **275**, 32383–32386.
33. Haney, P., Konisky, J., Koretke, K. K., Luthey-Schulten, Z. & Wolynes, P. G. (1997). Structural basis for thermostability and identification of potential active site residues for adenylate kinases from the archaeal genus *Methanococcus*. *Proteins: Struct. Funct. Genet.* **28**, 117–130.
34. Watanabe, K., Hata, Y., Kizaki, H., Katsube, Y. & Suzuki, Y. (1997). The refined crystal structure of *Bacillus cereus* oligo-1,6-glucosidase at 2.0 Å resolution: structural characterization of proline-substitution sites for protein thermostabilization. *J. Mol. Biol.* **269**, 142–153.
35. Bogin, O., Peretz, M., Hacham, Y., Korkhin, Y., Frolow, F., Kalb(Gilboa), A. J. & Burstein, Y. (1998). Enhanced thermal stability of *Clostridium beijerinckii* alcohol dehydrogenase after strategic substitution of amino acid residues with prolines from the homologous thermophilic *Thermoanaerobacter brockii* alcohol dehydrogenase. *Protein Sci.* **7**, 1156–1163.
36. Erra-Pujada, M., Debeire, P., Duchiron, F. & O'Donohue, M. J. (1999). The type II pullulanase of *Thermococcus hydrothermalis*: molecular characterization of the gene and expression of the catalytic domain. *J. Bacteriol.* **181**, 3284–3287.
37. Vieille, C. & Zeikus, G. J. (2001). Hyperthermophilic enzymes: sources, uses and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.* **65**, 1–43.
38. Yip, K. S., Stillman, T. J., Britton, K. L., Artymiuk, P. J., Baker, P. J., Sedelnikova, S. E. *et al.* (1995). The structure of *Pyrococcus furiosus* glutamate dehydrogenase reveals a key role for ion-pair networks in maintaining enzyme stability at extreme temperatures. *Structure*, **3**, 1147–1158.
39. Yip, K. S., Britton, K. L., Stillman, T. J., Lebbink, J., de Vos, W. M., Robb, F. T. *et al.* (1998). Insights into the molecular basis of thermal stability from the analysis of ion-pair networks in the glutamate dehydrogenase family. *Eur. J. Biochem.* **255**, 336–346.
40. Howlett, G. J., Blackburn, M. N., Compton, J. G. & Schachman, H. K. (1977). Allosteric regulation of aspartate transcarbamoylase. Analysis of the structural and functional behavior in terms of a two-state model. *Biochemistry*, **16**, 5091–5099.
41. Baker, D. P., Fetler, L., Keiser, R. T., Vachette, P. & Kantrowitz, E. R. (1995). Weakening of the interface between adjacent catalytic chains promotes domain closure in *Escherichia coli* aspartate transcarbamoylase. *Protein Sci.* **4**, 258–267.
42. Yates, R. A. & Pardee, A. B. (1956). Control of pyrimidine biosynthesis in *Escherichia coli* by a feed-back mechanism. *J. Biol. Chem.* **221**, 757–770.
43. Wild, J. R., Loughrey-Chen, S. J. & Corder, T. S. (1989). In the presence of CTP, UTP becomes an allosteric inhibitor of aspartate transcarbamoylase. *Proc. Natl Acad. Sci. USA*, **86**, 46–50.
44. Gerhart, J. C. & Pardee, A. B. (1962). The enzymology of control by feedback inhibition. *J. Biol. Chem.* **237**, 891–896.
45. Bethell, M. R., Smith, K. E., White, J. S. & Jones, M. E. (1968). Carbamyl phosphate: an allosteric substrate for aspartate transcarbamylase of *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **60**, 1442–1449.
46. Shanley, M. S., Foltermann, K. F., O'Donovan, G. A. & Wild, J. R. (1984). Properties of hybrid aspartate transcarbamoylase formed with native subunits from divergent bacteria. *J. Biol. Chem.* **259**, 12672–12677.
47. Foltermann, K. F., Beck, D. A. & Wild, J. R. (1986). *In vivo* formation of hybrid aspartate transcarbamoylases from native subunits of divergent members of the family Enterobacteriaceae. *J. Bacteriol.* **167**, 285–290.
48. Kedzie, K. (1987). Characterization of the *pyrB* gene of *Serratia marcescens* and hybrid formation with the *pyrB* gene of *Escherichia coli*, leading to the production of chimeric ATCase. PhD dissertation, Texas A&M University, College Station, TX.
49. Beck, D. A., Kedzie, K. M. & Wild, J. R. (1989). Comparison of the aspartate transcarbamoylases from *Serratia marcescens* and *Escherichia coli*. *J. Biol. Chem.* **264**, 16629–16637.
50. Wild, J. R. & Wales, M. E. (1990). Molecular evolution and genetic engineering of protein domains involving aspartate transcarbamoylase. *Annu. Rev. Microbiol.* **44**, 193–218.
51. De Staercke, C., Van Vliet, F., Xi, X.-G., Swarupa Rani, C., Ladjimi, M., Jacobs, A. *et al.* (1995). Intramolecular transmission of the ATP regulatory signal in *Escherichia coli* aspartate transcarbamylase: specific involvement of a clustered set of amino acid interactions at an interface between regulatory and catalytic subunits. *J. Mol. Biol.* **246**, 132–143.
52. Cunin, R., Wales, M. E., Van Vliet, F., De Staercke, C., Scapozza, L., Swarupa Rani, C. & Wild, J. R. (1996). Allosteric regulation in a family of enterobacterial aspartate transcarbamylases: intramolecular transmission of regulatory signals in chimeric enzymes. *J. Mol. Biol.* **262**, 258–269.
53. Van Vliet, F., Xi, X.-G., De Staercke, C., De Wannemaeker, B., Jacobs, A., Cherfils, J. *et al.* (1991).

Heterotropic interactions in aspartate transcarbamoylase: turning allosteric ATP activation into inhibition as a consequence of a single tyrosine to phenylalanine mutation. *Proc. Natl Acad. Sci. USA*, **88**, 9180–9183.
54. Stebbins, J. W. & Kantrowitz, E. R. (1989). The importance of the link between Glu204 of the catalytic chain and Arg130 of the regulatory chain for the homotropic and heterotropic properties of *Escherichia coli* aspartate transcarbamoylase. *J. Biol. Chem.* **264**, 14860–14864.
55. Stevens, R. C. & Lipscomb, W. N. (1992). A molecular mechanism for pyrimidine and purine nucleotide control of aspartate transcarbamoylase. *Proc. Natl Acad. Sci. USA*, **89**, 5281–5285.
56. Xi, X. G., Van Vliet, F., Ladjimi, M. M., De Wannemaeker, B., De Staercke, C., Piérard, A. *et al.* (1990). Co-operative interactions between the catalytic sites in *Escherichia coli* aspartate transcarbamoylase. Role of the C-terminal regions of the regulatory chains. *J. Mol. Biol.* **216**, 375–384.
57. Kerbiriou, D. & Hervé, G. (1972). Biosynthesis of an aspartate transcarbamylase lacking cooperative interactions. Disconnection of homotropic and heterotropic interaction under influence of 2-thiouracil. *J. Mol. Biol.* **64**, 379–392.
58. Knapp, S., de Vos, W. M., Rice, D. & Ladenstein, R. (1997). Crystal structure of glutamate dehydrogenase from the hyperthermophilic eubacterium *Thermotogamaritima* at 3.0 Å resolution. *J. Mol. Biol.* **267**, 916–932.
59. Auerbach, G., Ostendorp, R., Prade, L., Korndörfer, I., Dams, T., Huber, R. & Jaenicke, R. (1998). Lactate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*: the crystal structure at 2.1 Å resolution reveals strategies for intrinsic protein stabilization. *Structure*, **6**, 769–781.
60. Singleton, M., Isupov, M. & Littlechild, J. (1999). X-ray structure of pyrrolidone carboxyl peptidase from the hyperthermophilic archaeon *Thermococcus litoralis*. *Struct. Fold. Des.* **7**, 237–244.
61. Bogan, A. A. & Thorn, K. S. (1998). Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* **280**, 1–9.
62. Clackson, T. & Wells, J. A. (1995). A hot spot of binding energy in a hormone–receptor interface. *Science*, **267**, 383–386.
63. Legrain, C., Villeret, V., Roovers, M., Tricot, C., Clantin, B., Van Beeumen, J. *et al.* (2001). Ornithine transcarbamylase from *Pyrococcus furiosus*. *Methods Enzymol.* **331**, 227–235.
64. Jaenicke, R. (1991). Protein stability and molecular adaptation to extreme conditions. *Eur. J. Biochem.* **202**, 715–728.
65. Jaenicke, R. (1998). What ultrastable globular proteins teach us about proteins stabilization. *Biochemistry*, **63**, 312–321.
66. Jaenicke, R. & Böhm, G. (1998). The stability of proteins in extreme environments. *Curr. Opin. Struct. Biol.* **8**, 738–748.
67. Maes, D., Zeelen, J. P., Thanki, N., Beaucamp, N., Alvarez, M. & Dao Thi, M. H. (1999). The crystal structure of triosephosphate isomerase (TIM) from *Thermotoga maritima*: a comparative thermostability structural analysis of ten different TIM structures. *Proteins: Struct. Funct. Genet.* **37**, 441–453.
68. Maes, D. & Backmann, J. (1999). Protein thermostability: a case study on triosephosphate isomerase. *Rec. Res. Dev. Biochem.* **1**, 105–116.
69. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
70. Collaborative Computational Project Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
71. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallog.* **50**, 157–163.
72. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
73. Roussel, A. & Cambillau, C. (1992). Turbo-Frodo. Biographics, AFMB, Marseille, France.
74. Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). Use of TLS parameters to model anisotripic displacements in macromolecular refinement. *Acta Crystallog. sect. D*, **57**, 122–133.
75. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemistry quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
76. Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–2637.
77. McDonald, I. & Thornton, J. (1994). Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238**, 777–793.
78. Burley, S. K. & Petsko, G. A. (1985). Aromatic–aromatic interaction: a mechanism of protein structure stabilization. *Science*, **229**, 23–28.
79. Richards, F. M. (1977). Areas; volumes, packing density and protein structure. *Annu. Rev. Biophys. Bioeng.* **6**, 151–175.
80. Kleywegt, G. J. & Jones, T. A. (1994). Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Crystallog. sect. D*, **50**, 178–185.
81. Hubbard, S. J. & Argos, P. (1995). Detection of internal cavities in globular proteins. *Protein Eng.* **8**, 1011–1015.
82. Voronoi, G. F. (1908). Nouvelles applications des paramètres continus à la théorie des formes quadratiques. *J. Reine Angew. Math.* **82**, 1–14.
83. Kraulis, P. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
84. Meritt, E. A. & Bacon, D. J. (1997). Raster3D photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524.

Edited by R. Huber

(Received 4 November 2003; received in revised form 24 March 2004; accepted 26 March 2004)

5 Expression, purification, crystallization and preliminary X-ray crystallographic studies of a cold-adapted aspartate carbamoyltransferase from *Moritella profunda*

Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Expression, purification, crystallization and preliminary X-ray crystallographic studies of a cold-adapted aspartate carbamoyltransferase from *Moritella profunda*

Dirk De Vos, Paco Hulpiau, Bjorn Vergauwen, Savvas N. Savvides and Jozef Van Beeumen*

Laboratorium voor Eiwitbiochemie en Eiwitengineering, Universiteit Gent, K. L. Ledeganckstraat 35, Ghent, Belgium

Correspondence e-mail: jozef.vanbeeumen@ugent.be

Received 2 December 2004
Accepted 26 January 2005
Online 12 February 2005

Aspartate carbamoyltransferase (ATCase) catalyzes the carbamoylation of the α-amino group of L-aspartate by carbamoyl phosphate (CP) to yield *N*-carbamoyl-L-aspartate and orthophosphate in the first step of *de novo* pyrimidine biosynthesis. Apart from its key role in nucleotide metabolism, the enzyme is generally regarded as a model system in the study of proteins exhibiting allosteric behaviour. Here, the successful preparation, crystallization and diffraction data collection of the ATCase from the psychrophilic bacterium *Moritella profunda* are reported. To date, there is no structural representative of a cold-adapted ATCase. The structure of *M. profunda* ATCase is thus expected to provide important insights into the molecular basis of allosteric activity at low temperatures. Furthermore, through comparisons with the recently reported structure of an extremely thermostable ATCase from *Sulfolobus acidocaldarius*, it is hoped to contribute to general principles governing protein adaptation to extreme environments. A complete native data to 2.85 Å resolution showed that the crystal belongs to space group $P3_221$, with unit-cell parameters a = 129.25, b = 129.25, c = 207.23 Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$, and that it contains three catalytic and three regulatory subunits per asymmetric unit. The three-dimensional structure of the *Escherichia coli* ATCase was sufficient to solve the structure of the *M. profunda* ATCase *via* the molecular-replacement method and to obtain electron density of good quality.

1. Introduction

Aspartate carbamoyltransferase (ATCase; EC 2.1.3.2) catalyzes the carbamoylation of the α-amino group of L-aspartate by carbamoyl phosphate (CP) to yield *N*-carbamoyl-L-aspartate and orthophosphate in the first step of *de novo* pyrimidine biosynthesis. Apart from its key role in nucleotide metabolism, this enzyme is generally regarded as a model system in the study of proteins exhibiting allosteric behaviour. In this regard, the ATCase from *Escherichia coli* has been the subject of extensive structural and biochemical studies (reviewed in Allewell, 1989; Lipscomb, 1994; Helmstaedt *et al.*, 2001). The *E. coli* ATCase consists of catalytic chains encoded by the *pyrB* gene and regulatory chains encoded by the *pyrI* gene. The catalytic chains, organized into trimers (c_3), and the regulatory chains, organized into dimers (r_2), combine to form a dodecameric complex with the subunit composition $2(c_3):3(r_2)$. ATCase exists in different quaternary structures in different organisms. For example, in *Bacillus subtilis* the enzyme exists as a trimer of catalytic chains without any regulatory chains (Lerner & Switzer, 1986).

The X-ray structure of the dodecameric enzyme from *E. coli* has been determined in both the absence and presence of substrates/inhibitors/products (Kim *et al.*, 1987; Ke *et al.*, 1988; Gouaux & Lipscomb, 1988, 1990; Gouaux *et al.*, 1990; Jin *et al.*, 1999; Huang & Lipscomb, 2004). Furthermore, a low-resolution C^α trace of the *B. subtilis* enzyme is also available (Stevens *et al.*, 1991). Recently, we reported the first structure of the intact (unliganded) dodecameric ATCase from a hyperthermophilic organism (De Vos *et al.*, 2004). To date, however, the three-dimensional structure of a cold-adapted ATCase is an important missing component in our understanding of extremophilic proteins. The structure of a psychrophilic ATCase would thus lead to a better understanding of cold activity and thermolability of enzymes in general and in particular in relation to

allosteric regulation. Furthermore, structural characterization of psychrophilic enzymes may provide clues on how to improve their stability at higher temperatures while maintaining their flexibility in a colder environment, a feature that may be useful in many biotechnological applications (Gerday *et al.*, 2000).

Here, we report the successful expression, purification, crystallization and structure determination of the ATCase from *Moritella profunda*, a psychropiezophilic eubacterium isolated from deep Atlantic sediments. The *pyrB* and *pyrI* genes of *M. profunda* were isolated previously by complementation of an *E. coli pyrBI* mutant; they show sequence identities of 74 and 52%, respectively, with their *E. coli* counterparts (Xu, 2002). Based on this, tertiary and quaternary structures similar to the *E. coli* ATCase were expected for the *M. profunda* enzyme. We have now been able to determine the structure of *M. profunda* ATCase using maximum-likelihood methods as implemented in the recently launched program *PHASER* v.1.2 (Storoni *et al.*, 2004). We have identified three copies of the catalytic chain and three copies of the regulatory chain in the asymmetric unit. Model building and refinement of the structure are under way.

2. Materials and methods

2.1. Protein preparation

The cloning of the *pyrB* and *pyrI* genes of *M. profunda* by complementation of an *E. coli pyrBI* mutant has been described previously (Xu, 2002). These genes were subsequently inserted into the expression plasmids pET-Duet and pACYC-Duet (Novagen), which are designed to coexpress two target genes using separate promoters. The *pyrB* gene was cloned for production of the catalytic chain with an N-terminal 6×His tag containing the sequence MGSSHHHHHHSQDPNS. When transformed into *E. coli* BL21(DE3), both constructs resulted in substantial amounts of recombinant protein. SDS–PAGE analysis revealed that similar amounts of catalytic and regulatory chain were present. Comparison of the protein expressed in the native host with the recombinant protein shows that they display very similar functional properties (work to be published). A 10 ml volume of an overnight 301 K starter culture was inoculated into 1 l Luria–Bertani medium (Sambrook *et al.*, 1989) at the same temperature. When the culture medium reached an optical density (at 600 nm) of 0.5, isopropyl-β-D-thiogalactoside (IPTG) was added to a final concentration of 0.5 m*M*, after which the culture was incubated overnight at 288 K. The cell pellet obtained after centrifugation was resuspended in 50 m*M* Tris–HCl containing 300 m*M* NaCl at a final pH of 8.0 (buffer *A*). The cells were lysed by sonication and the cell-free supernatant was loaded onto an Ni–NTA column (Qiagen). After washing the column with buffer *A* containing 10 m*M* imidazole, the protein was eluted with buffer *A* containing 250 m*M* imidazole. The protein was then further purified to homogeneity on a Superdex-200 size-exclusion column (Amersham) and the quality of the purification was evaluated with SDS–PAGE (Fig. 1). The high NaCl concentration in buffer *A* served mainly to prevent precipitation of the recombinant ATCase on the Ni–NTA column.

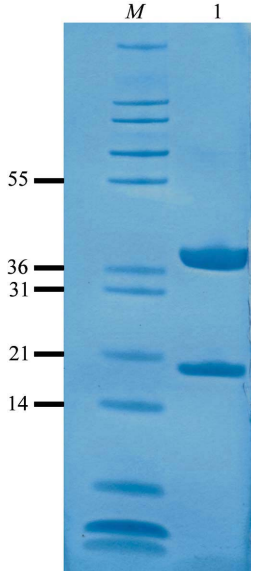


Figure 1
SDS–PAGE analysis of recombinant *M. profunda* ATCase. Lane 1, purified enzyme. Molecular standards (kDa) are indicated by *M.*

2.2. Crystallization and data collection

Owing to the temperature lability of the enzyme, all crystallization trials were performed at 277 K. Because the protein precipitates at lower temperatures, it was stored at this temperature at a concentration of 10 mg ml^{-1} in 5 m*M* Tris–HCl with 30 m*M* NaCl at a final pH of 8.0. Under these conditions, the protein is stable and retains its activity for at least six months. Crystallization experiments were performed using the hanging-drop vapour-diffusion method. Drops were prepared by mixing equal volumes of protein solution and precipitant solution and were allowed to equilibrate against 500 µl reservoir solution.

Selected crystals were transferred to a solution containing 20%(*v*/*v*) glycerol in mother liquor prior to flash-cooling by plunging directly into liquid nitrogen. From crystals tested on our in-house FR591 rotating-anode generator (Bruker–Nonius), those diffracting the best were selected for further testing with synchrotron radiation. X-ray diffraction data were collected at a wavelength of 0.92 Å at beamline BW7A (DESY, EMBL Hamburg Outstation, Germany) using a MAR CCD (MAR Research) detector system. A native data set was collected from a single crystal. The crystal-to-detector distance was 220 mm. A total of 240 rotation images were collected with an oscillation angle of 0.5°. Data processing was carried out using the programs *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). The diffraction intensities were indexed using the program *DENZO* and were scaled and merged to 2.85 Å resolution using the program *SCALEPACK*. Intensities were converted into structure-factor amplitudes using the program *TRUNCATE* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994).

Figure 2
Trigonal crystals of the *M. profunda* ATCase. The crystals grew as rhombohedra with typical dimensions of 0.3 × 0.3 × 0.3 mm.

Table 1
X-ray data-collection statistics.

Values in parentheses refer to the highest resolution shell.

Space group	$P3_221$
Unit-cell parameters (Å, °)	$a = b = 129.25$, $c = 207.23$, $\alpha = \beta = 90$, $\gamma = 120$
Packing density V_M (Å^3 Da^{-1})	3.1
Solvent content (%)	60.5
Resolution (Å)	50.0–2.85 (2.93–2.85)
No. of unique reflections	45685 (3817)
Redundancy	6.7 (6.8)
R_{sym}† (%)	14.9 (53.6)
Average $I/\sigma(I)$	9.9 (3.3)
Completeness (%)	96.1 (97.3)

† $R_{sym} = \sum_h \sum_i |I(h, i) - \langle I(h)\rangle| / \sum_h \sum_i I(h, i)$, where $I(h, i)$ is the intensity of the ith measurement of reflection h and $\langle I(h)\rangle$ is the average value over multiple measurements.

3. Results and discussion

From an initial screening of crystallization conditions with Structure Screen 1 (Molecular Dimensions Ltd), five conditions were found that produced crystals (conditions 27, 30, 32, 44 and 45). Except for condition 45 (4.0 *M* sodium formate), all other conditions contained lithium or ammonium sulfate as the precipitant. Starting from condition 32 (0.1 *M* Tris pH 8.5, 2.0 *M* ammonium sulfate), the buffer pH and precipitant concentration were optimized. Lowering the ammonium sulfate concentration to 1.5 *M* was sufficient to provide crystals that grew overnight to dimensions of 0.3 × 0.3 × 0.3 mm (Fig. 2). Crystals could be obtained within the pH range 7.0–8.5 and their quality did not vary in a pH-dependent manner. The relative ease with which the *M. profunda* ATCase crystallizes challenges the general notion that cold-adapted proteins are difficult to crystallize because of inherent flexibility. Similar cases have been reported in the past (Villeret *et al.*, 1997; Van Petegem *et al.*, 2002).

A complete data set was collected to 2.85 Å resolution using a single crystal grown at pH 8.0. The full data-collection statistics are presented in Table 1. Assuming a subunit molecular weight for the catalytic and regulatory chains of 36 152 and 16 889 Da, respectively, and assuming three copies of each chain in the asymmetric unit, the solvent content of the crystal was 60.5%, with a volume-to-weight ratio of 3.1 Å^3 Da^{-1}. These values are within the frequently observed ranges for protein crystals (Matthews, 1968).

One subunit of the *E. coli* ATCase catalytic chain (Stevens *et al.*, 1990; PDB code 6at1) served as the starting model for solving the structure by molecular replacement using maximum-likelihood methods implemented in the program *PHASER* v.1.2 (Storoni *et al.*, 2004). In a fully automated rotational and translational search, three copies of the catalytic chain with reasonable packing could be located in the asymmetric unit. Three copies of the smaller regulatory subunits could also be located. This was performed by refeeding *PHASER* with the catalytic chain after a first round of model refinement and searching with the *E. coli* ATCase regulatory chain (PDB code 6at1). Furthermore, we were able to resolve the ambiguity between the enantiomorphic space groups $P3_121$ and $P3_221$ in favour of the second. After applying the crystallographic symmetry operations of space group $P3_221$, we found that the positions of the catalytic and regulatory chains correspond well with their position in the *E. coli* ATCase dodecameric complex. Combining this with the result from size-exclusion chromatography and with our finding that the regulatory chains were co-purified on Ni–NTA in equal amounts with the histidine-tagged catalytic chains suggests that the *M. profunda* ATCase quaternary organization is in agreement with that of *E. coli*, *i.e.* $2(c_3):3(r_2)$. Initial electron-density maps revealed several unique structural features of *M. profunda* ATCase compared with the search models and thus confirmed the correctness of our solution. Model building and refinement of this structure are under way.

We gratefully acknowledge access to EMBL beamline BW7A at the DORIS storage ring, DESY, Hamburg. DDV is a Research Fellow of the Research Foundation-Flanders (FWO-Vlaanderen).

References

Allewell, N. M. (1989). *Annu. Rev. Biophys. Chem.* **18**, 71–92.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
De Vos, D., Van Petegem, F., Remaut, H., Legrain, C., Glansdorff, N. & Van Beeumen, J. J. (2004). *J. Mol. Biol.* **339**, 887–900.
Gerday, C., Aittaleb, M., Bentahir, M., Chessa, J. P., Claverie, P., Collins, T., D'Amico, S., Dumont, J., Garsoux, G., Georlette, D., Hoyoux, A., Lonhienne, T., Meuwis, M. A. & Feller, G. (2000). *Trends Biotechnol.* **18**, 103–107.
Gouaux, J. E. & Lipscomb, W. N. (1988). *Proc. Natl Acad. Sci. USA*, **85**, 4205–4208.
Gouaux, J. E. & Lipscomb, W. N. (1990). *Biochemistry*, **29**, 389–402.
Gouaux, J. E., Stevens, R. C. & Lipscomb, W. N. (1990). *Biochemistry*, **29**, 7702–7715.
Helmstaedt, K., Krappmann, S. & Braus, G. H. (2001). *Microbiol. Mol. Biol. Rev.* **65**, 404–421.
Huang, J. & Lipscomb, W. N. (2004). *Biochemistry*, **43**, 6422–6426.
Jin, L., Stec, B., Lipscomb, W. N. & Kantrowitz, E. R. (1999). *Proteins Struct. Funct. Genet.* **37**, 729–742.
Ke, H.-M., Lipscomb, W. N., Cho, Y. & Honzatko, R. B. (1988). *J. Mol. Biol.* **204**, 725–747.
Kim, K. H., Pan, Z., Honzatko, R. B., Ke, H.-M. & Lipscomb, W. N. (1987). *J. Mol. Biol.* **196**, 853–875.
Lerner, C. G. & Switzer, R. L. (1986). *J. Biol. Chem.* **261**, 11156–11165.
Lipscomb, W. N. (1994). *Adv. Enzymol.* **68**, 67–152.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning. A Laboratory Manual*, 2nd ed. New York: Cold Spring Harbor Laboratory Press.
Stevens, R. C., Gouaux, J. E. & Lipscomb, W. N. (1990). *Biochemistry*, **29**, 7691–7701.
Stevens, R. C., Reinisch, K. M. & Lipscomb, W. N. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 6087–6091.
Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). *Acta Cryst.* D**60**, 432–438.
Van Petegem, F., Collins, T., Meuwis, M.-A., Gerday, C., Feller, G. & Van Beeumen, J. (2002). *Acta Cryst.* D**58**, 1494–1496.
Villeret, V., Chessa, J. P., Gerday, C. & Van Beeumen, J. (1997). *Protein Sci.* **6**, 2462–2464.
Xu, Y. (2002). PhD thesis. Vrije Universiteit Brussel, Brussels, Belgium.

6 Structural basis for cold activity and regulation

Biochemical properties and crystal structure of aspartate carbamoyltransferase from the psychrophilic bacterium *Moritella profunda*

STRUCTURAL BASIS FOR COLD ACTIVITY AND REGULATION: BIOCHEMICAL PROPERTIES AND CRYSTAL STRUCTURE OF ASPARTATE CARBAMOYLTRANSFERASE FROM THE PSYCHROPHILIC BACTERIUM *MORITELLA PROFUNDA**

Dirk De Vos[a§], Bjorn Vergauwen[a§], Paco Hulpiau[a], Ying Xu[b], Nicolas Glansdorff[b] and Jozef J. Van Beeumen[a$]

[a]From the Laboratory of Protein Biochemistry and Protein Engineering, Ghent University, K.L. Ledeganckstraat 35, 9000 Gent, Belgium

[b]From the Laboratory of Microbiology, Free University of Brussels (VUB) and J.M. Wiame Research Institute 1, ave E. Gryzon, B-1070, Brussels, Belgium

Running title: Psychrophilic ATCase of *Moritella profunda*

[§]These authors contributed equally to this manuscript.

[$]Address correspondence to: Jozef J. Van Beeumen, Laboratory of Protein Biochemistry and Protein Engineering, Ghent University, K.L. Ledeganckstraat 35, 9000 Gent, Belgium; Tel. +32 (0)9 264 51 27; Fax. +32 (0)9 264 53 38; E-mail: Jozef.Vanbeeumen@Ugent.be

Aspartate carbamoyltransferase (EC 2.1.3.2) is extensively studied as a model for cooperativity and allosteric regulation. The structure of the *Escherichia coli* enzyme has been thoroughly analyzed by X-ray crystallography, and recently the crystal structures of two hyperthermophilic ATCases of the same structural class have been characterized. We here report the detailed functional and structural investigation of the ATCase from the psychrophilic deep sea bacterium *Moritella profunda*. Our analysis indicates that the enzyme conforms to the *E. coli* model in that two allosteric states exist that are influenced by similar homotropic interactions. The heterotropic properties differ in that CTP and UTP inhibit the holoenzyme, but ATP seems to exhibit a dual regulatory pattern. The crystal structure of the unliganded *M. profunda* ATCase structure shows resemblance to a more extreme T state reported previously for an *E. coli* ATCase mutant. A detailed molecular analysis reveals potential features of adaptation to cold activity and cold regulation. Finally, structural and functional comparison of ATCases across the full physiological temperature range demonstrates the important, but fundamentally different role for electrostatics in protein adaptation at both extremes.

Aspartate carbamoyltransferase (ATCase[1]; E.C. 2.1.3.2) is one of the most widely studied allosteric enzymes (1-4). ATCase catalyzes the first reaction in the pyrimidine biosynthetic pathway, the reaction of carbamoyl phosphate with aspartate to form phosphate and *N*-carbamoyl-L-aspartate (5). Since Yates and Pardee (6) observed that the enzyme from *Escherichia coli* is subject to feedback inhibition, there has been an intense effort to determine and understand the regulatory mechanism of this enzyme. In addition to heterotropic control, the enzyme also exhibits positive homotropic interactions between catalytic sites (7,8).

The enzyme is a dodecamer of 310 kDa, comprising six catalytic chains (34 kDa each) which assemble to form two catalytic trimers, and six regulatory chains (17 kDa each) which assemble to form three regulatory dimers. The X-ray structure of the enzyme reveals that both the catalytic and the regulatory chains are composed of two folding domains. The aspartate and carbamoyl phosphate domains of the catalytic chain are responsible for binding the substrates aspartate and carbamoyl phosphate, respectively, while the allosteric and zinc domains of the regulatory chain are responsible for binding the regulatory nucleotides and a structural zinc, respectively. The quaternary structure of the enzyme undergoes alteration upon the binding of substrates, or substrate analogues, such as *N*-phosphonacetyl-L-aspartate (PALA), from a low-activity low-affinity T state to a high-activity high-affinity R state. This quaternary conformational change involves an elongation of the molecule by 11 Å along the molecular 3-fold axis, along with rotations about the molecular 3-fold and 2-fold axes. Changes at the tertiary level include the closure of the carbamoyl phosphate and aspartate domains to form the active site, as well as the rearrangement of the 80s and 240s loops (9).

A number of structure-function studies have been carried out on B-class (10,11) ATCases, mostly programmed around the paradigm enzyme from *E. coli*. In two cases, other B-class ATCases were concurrently analyzed structurally and functionally, with both enzymes originating from (hyper)thermophilic organisms (12-14). We here report the first crystal structure of an ATCase from a cold-adapted organism, *Moritella profunda*, in combination with a characterization of its steady-state kinetics and allosteric behavior.

M. profunda is a psychrophilic γ-proteobacterium isolated from deep sea sediments, with an optimal growth temperature of ~2 °C (15). *E. coli* and *M. profunda* ATCases display high amino acid sequence identities both for the catalytic and regulatory chains (73 % and 53 %, respectively), thus providing a good starting point to elucidate the molecular determinants of cold-activity which are still poorly understood (16). Recent interest in the isolation and understanding of psychrophilic enzymes has been spurred by their biotechnological potential (17). Indeed, to provide their hosts with sufficient catalytic power, enzymes from psychrophiles must be highly active at low temperatures. However, to survive such extreme growth conditions, these enzymes must also retain their regulatory properties. Both functional characteristics are shown here to be valid for the *M. profunda* ATCase. Given the complex allosteric behavior of the *E. coli* enzyme (cfr. effector signals influencing remote sites at distances of over 60 Å), the question indeed arises how this dynamic system is able to sustain its intricate regulatory characteristics at low temperatures.

EXPERIMENTAL PROCEDURES

Biochemical reagents

All biochemical reagents used in this study were from Sigma-Aldrich (St-Louis, MO), unless indicated otherwise.

Cloning, expression and purification

The *M. profunda pyrBI* genes were cloned and recombinantly expressed in *E. coli* as described previously (18), which also describes the purification scheme for the recombinant His-tagged ATCase. Purified ATCase was stored in 50 mM Tris-HCl buffer, pH 7.5, containing 300 mM NaCl, at 4°C. When stored under these conditions, the enzyme did not lose activity for at least 6 months. Protein concentration was determined using the Bradford (19) colorimetric protein assay of Bio-Rad Laboratories (Hercules, CA), using bovine serum albumin (BSA) as the standard.

Bacterial strains, culture conditions, and preparation of native ATCase

M. profunda LMG 21259 (15) was obtained from the BCCM Culture Collection of the Laboratory of Microbiology at the Ghent University, Ghent, Belgium.
M. profunda was grown in rich Marine broth 2216 (Difco, BD Biosciences, Cockeysville, MD), pH 7.6, at 8 °C, in a 7-L fermentor vessel (Applitek, Nazareth, Belgium) at an aeration rate of about 100 ml min^{-1} $liter^{-1}$, until late exponential phase. Cells were collected by centrifugation (15,000 x g, 15 min, 4 °C), resuspended in buffer A (50 mM Tris-HCl, pH 7.5, 300 mM NaCl) and disrupted by sonication (Branson sonifier, model 450) using a 2-s pulses at 2-s intervals routine at 30% of the maximum intensity. The crude cell lysate was centrifuged for 30 min, at 15,000 x g and 4 °C, after which the supernatant was applied onto a desalting column (HiPrepTM 26/10 Desalting Column (Amersham Pharmacia Biotech, Uppsala, Sweden); 10 ml/min), pre-equilibrated with buffer A. The protein eluate was collected, quantitated using the Bio-Rad Bradford protein assay, and used immediately as a source of native ATCase in enzyme assays.

Enzyme assays

ATCase activity was measured using the colorimetric method of Prescott and Jones (20). Absorptions at 2.5 min intervals for 10 min were measured at 466 nm using an Uvikon 943 UV-visible double beam spectrophotometer (Kontron Instruments, Watford, UK). Progress curves were linear in this time frame. Both recombinant His-tagged and native ATCase aggregate in the absence of NaCl. Unless otherwise indicated, assays were therefore performed in 50 mM Tris-HCl buffer, pH 7.5, containing 300 mM NaCl, which creates an ionic strength that should closely resemble that of the interior of marine organisms (21). L-Asp saturations were performed in the presence of 5 mM CP. CP saturations were performed in the presence of 300 mM L-Asp. Saturation curves were fitted to

the Hill equation using a non-linear curve fitting routine (Graphpad PRISM 4.0 software package). The experiments were performed in duplicate; mean values for kinetic parameters are reported with errors representing the standard error of the mean (SEM).

One enzyme unit (U) is defined as the amount of protein that converts 1 µmol substrate to product per h.

Structure refinement

The crystallization, data collection and structure solution of the unliganded *Mp*ATC$_T$ have been described previously (18). The structure of the unliganded *Moritella profunda* ATCase (referred to as *Mp*ATC) was refined starting from the C1 and R1 chains of the unliganded *E. coli* ATCase (referred to as *Ec*ATC, pdb entry 6AT1), using the program REFMAC (22) from the CCP4 suite (Collaborative Computational Project Number 4, 1994 (23)) by monitoring the convergence of R_{work} and R_{free} (Table 2). Inspection of electron density maps and model refinement were carried out with TURBO-FRODO (created by Roussel and Cambillau (1992); http://www.afmb.univ-mrs.fr/-TURBO-) and side chains with missing electron density were not modelled. In the final stages TLS refinement, as implemented in Refmac, was used to further improve the electron density in ill-defined regions (with 9 TLS groups corresponding to three catalytic chains, three zinc domains and three allosteric domains). After refinement of the protein, ligand and water molecules were built into the remaining difference density of the $F_o - F_c$ electron density maps, followed by further refinement and editing of the waters. Refinement statistics (shown in Table 2) indicate high average crystallographic B factors. The quality of the electron density, however, in general is as can be expected for a crystal structure at this resolution. Residues missing from the final model are the N-terminal affinity tag including the first residue of the catalytic chains, and regions with less well-defined electron density for side chains are the 80s and 240s loops. The electron density of the N-terminal allosteric domains of the regulatory chains is of lower quality than can be expected at this resolution (esp. chain F) and were therefore excluded from certain analyses of thermal stability. Missing from the regulatory chain model are 11 N-terminal residues, loop region 48-54 and residues 152 and 153 at the C terminal end for chain D. In addition, loop regions 128-131 of chains E and F are missing from the model. Three extra N terminal residues, one extra C terminal residue and residue 48 are additionally missing from chain F. Structure validation was performed with the program PROCHECK (24).

Structure analysis

Unless stated differently, PDB entry 6AT1 was chosen as T state model of the *Ec*ATC for molecular comparison with *Mp*ATC. Structural superpositions were performed using LSQKAB (Collaborative Computational Project Number 4, 1994 (23)). Secondary structure was assigned using DSSP (25). Hydrogen bonds between protein atoms were calculated using the program HBPLUS (26) with the default parameters for distance and angles. The presence of salt-bridges was inferred when Asp or Glu side-chain carbonyl oxygen atoms were found to be within a 4.0 Å distance from the nitrogen atoms in Arg, Lys and His side-chains. Accessible surface areas were calculated with the programs ACCESS and ACCFMT (27). Probe-accessible cavity volumes were searched with the program VOIDOO (28) using a 1.2 Å probe radius and a minimum volume of 10 Å^3, corresponding to 'real' cavities according to Hubbard and Argos (29). VOIDOO was also used for van der Waals volume determination of the proteins. The real volumes were calculated by the Voronoi procedure (30), using the programs ACCESS and VOLUME (27). The packing density was calculated as the ratio of the van der Waals volume to the real volume. Figures were prepared with the program PYMOL (31). Electrostatic surface potential representations were generated by PYMOL from pdb entries 2BE7 for *Mp*ATC, 6AT1 for *Ec*ATC and 1PG5 for *Sa*ATC, after adding missing side chains with the program DEEPVIEW/SWISS-PDBVIEWER 3.7 (32).

Protein Data Bank accession codes

Atomic coordinates and related structure factors have been deposited in the RCSB Protein Data Bank, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/) with identification code 2BE7.

RESULTS

Sequence analysis

Cloning and expression of *Moritella profunda* *pyrB* and *PyrI* genes have been described previously (18). Comparison of deduced amino acid sequences clearly shows homology to ATCases from mesophilic and thermophilic bacteria, from hyperthermophilic archaea and from eukaryotes. High sequence identities were seen on comparison with enterobacterial ATCases, *E. coli* in particular (identities of 73 % and 53 % for *pyrB* and *pyrI* gene products, respectively; Fig. 1). Sequence identities with the ATCases from hyperthermophilic archaea that also share the *pyrBI* pattern of genetic organisation were about 20 % lower, but this still gives 45-50 % identity for the catalytic chain (e.g. 45 % for the *S. acidocaldarius* ATCase) and about 25-35 % identity for the regulatory chain (e.g. 28 % for *S. acidocaldarius* ATCase).

Resistance to thermoinactivation

The protein stability of the psychrophilic recombinant *M. profunda* holoenzyme ATCase was measured by determining the residual activity at 4 °C after incubation for increasing periods at temperatures ranging from 50 to 65 °C (Fig. 2). More than 50 % activity remained after incubating the enzyme for one hour at 55 °C. However, 10 min at 65 °C almost completely inactivated the enzyme. For a comparison, a 40 min incubation period of the *Ec*ATC holoenzyme at 66 °C resulted in half maximal velocity (33). These results therefore indicate that the *M. profunda* enzyme is characterized by a lowered resistance to thermoinactivation compared to the *E. coli* ATCase.

Temperature dependence of enzyme activity

Under laboratory conditions, the maximal growth rate of *M. profunda* is at 2 °C or less, and no growth has been observed above 13 °C (15). At 4 °C, the psychrophilic ATCase was ~5.6 times more active than the *E. coli* enzyme (77.9 µmol min^{-1} mg^{-1} vs. 14.0 µmol min^{-1} mg^{-1}), while both enzymes are comparably active at 30 °C.

Steady-state saturation kinetics

Table 1 depicts the steady-state kinetic parameters for the recombinant *Mp*ATC holoenzyme, which were determined under physiologically relevant conditions (i.e. 50 mM Tris-HCl buffer, 300 mM NaCl, pH 7.5, and 4 °C). For comparison, Table 1 also includes kinetic parameters for a mesophilic (*E. coli*) and for a hyperthermophilic ATCase (*Pyrococcus abyssi*). For the *Ec*ATC holoenzyme, cooperativity is apparent for both CP and L-Asp; however, L-Asp saturation curves are more sigmoidal (n_H between 1.7 and 2.2) than carbamoyl phosphate (CP) saturation curves (n_H of ~1.2). These differences in cooperativity reflect the ordered BiBi binding of substrates (34); CP binds before L-Asp, and carbamoyl-aspartate is released before inorganic phosphate (35). The L-Asp and CP saturation curves for the psychrophilic ATCase of *M. profunda* are shown in Fig. 3; both curves are S-shaped, with $n_{H,Asp} > n_{H,CP}$, indicating an *E. coli*-like mechanism of substrate binding. While the $S_{0.5}$-values for CP are more comparable, those for L-Asp increase in the following order: thermophile (2.3 mM) < mesophile (7.5 mM) < psychrophile (101 ± 10 mM).

The foregoing indications for cooperative rather low-affinity L-Asp binding are corroborated by the *N*-phosphonacetyl-L-aspartate (PALA) inhibition titration curves shown in Fig. 4. In the presence of lower than half-saturating L-Asp concentrations, PALA stimulates the activity of the *E. coli* ATCase: while PALA blocks the sites which it occupies, it converts the remaining sites to a more active state, resulting in an increase in activity (36). Whereas at the $S_{0.5}$-concentration of 100 mM only direct inhibition by PALA can be observed, at 1/5 of this L-Asp-concentration the *M. profunda* ATCase reaction was stimulated more than 100%, indicating *E. coli*-like cooperativity towards L-Asp. The concentration range (10-80 µM) at which PALA stimulated the psychrophilic reaction was 10-fold that at which the *E. coli* enzyme is activated by the inhibitor (1-8 µM), which is in agreement with the (~10-fold) difference in the enzymes' $S_{0.5}$-values for L-Asp.

Saturation by the nucleotide effectors

*Ec*ATC is inhibited by CTP and UTP and activated by ATP, all nucleotides acting in the high micromolar range. Titration curves by these nucleotide effectors are shown in Fig. 5 for the *Mp*ATC and the results are summarized and compared to the *E. coli* and *P. abyssi* values in Table 1. The titration studies were conducted at 20 mM L-Asp, a concentration corresponding to 1/5 the $S_{0.5}$-value, and in the presence of $S_{0.5}$-concentrations of CP (0.5 mM). CTP and UTP inhibited the psychrophilic enzyme. As shown

for the *E. coli* enzyme, inhibition was never complete; CTP maximally inhibited 83% (cf. 60-80% for the *E. coli* enzyme), while UTP inhibited no more than 40% (cf. no more than 10% for the *E. coli* enzyme). The CTP concentration at which half the maximal effect (i.e. the $E_{50,\ CTP}$-value) is observed was calculated to be 3.39 ± 0.50 μM, which is about two orders of magnitude lower than the $E_{50,\ CTP}$-value of ~ 206 μM (37,38) reported for *Ec*ATC. UTP inhibited the psychrophilic ATCase less effectively; 1.65 ± 0.43 mM UTP is required to obtain half the maximal effect.

Interestingly, the ATP-titration curve showed a dual pattern; stimulation was seen up to 3 mM, while higher concentrations of ATP inactivated the reaction. At concentrations exceeding 8 mM, ATP inhibited the reaction completely. The relative activity maximally increased ~2.5-fold in the presence of about 1.25 mM ATP and the $E_{50,ATP}$-value was calculated to be 145 ± 34 μM.

M. profunda ATCase in the native host

The specific activity, determined at 4°C, of ATCase in *M. profunda* crude lysate, prepared from late exponential phase cultures grown in rich Marine broth at 8 °C, was 0.83 U (mg protein)$^{-1}$. Saturation kinetics, cooperativity, as well as temperature and nucleotide dependence of activity were found to be similar for native enzyme in crude *M. profunda* extract and for His-tagged recombinant protein expressed in *E. coli* (Table 1). These results suggest that the recombinant enzyme is a suitable molecular form for the study of structure-function relations of this psychrophilic ATCase.

X-ray structure refinement and validation (Table 2)

The enzyme crystallizes in space group $P3_221$ with three catalytic (chains A, B, C) and three regulatory monomers (chains D, E, F) in the asymmetric unit. The dodecameric [$2c_3$:$3r_2$] complex typical for class-B ATCases is generated by the crystallographic twofold axis of space group $P3_221$, in accordance with the molecular weight estimated by gel filtration chromatography (18) (Fig. 6). With the exception of L267, all non-glycine residues display main-chain dihedral angles that are localized within allowed or additional allowed regions of the Ramachandran plot, as defined by the program PROCHECK (24). In accordance with all other ATCase structures, the peptide link between L267 and P268 is in a *cis* conformation (the Ramachandran analysis assumes a *trans* conformation). Also included in the model are three sulphate ions located at the active sites and 36 solvent molecules.

Structural characterization

The main violations against the 3-fold non-crystallographic symmetry relating the catalytic-regulatory chain heterodimers within the asymmetric unit are situated at the 80s and 240s loop regions of the catalytic chains. Superposition of the catalytic chains leads to rmsds for C_α's of 0.40 Å for A and B, 0.46 Å for A and C, and 0.35 Å for B and C. Superposition of the regulatory chains gives rmsds of 0.39 Å for D and E, 0.48 Å for D and F and 0.51 Å for E and F. Comparison with the unliganded T state *E. coli* catalytic and regulatory chains (rmsds of 1.0 – 1.1 Å) indicates a good overall fit in agreement with strong similarities in secondary and tertiary structure. The (sole) deletion at the position of G120 in *Ec*ATC only induces a local structural effect. However, closer examination clearly indicates that the two catalytic chain domains are located further from each other in accordance with a domain opening relative to $EcATC_T$. Furthermore, the 240s loop region, which typically in *Ec*ATC undergoes the largest movement upon the domain closure taking place at the T to R transition, has undergone a shift (maximum C_α distance of 6 Å) in the opposite direction. Interestingly, similar shifts were previously reported for the crystal structures of single site mutant Y165F of *Ec*ATC (catalytic chain) which, hence, has been described as an 'extreme T state' model (38). Strikingly, the similarity with this mutant extends beyond a mere superficial resemblance.

Similar as for the Y165F mutant, a collapse (relative to $EcATC_T$) is observed of the hydrophobic residues Val 230, Phe 235 and Phe 247 onto P189 and L192 (*M. profunda* numbering) which appears to drive changes in several hydrogen bonds and salt bridges, so that the positions of numerous groups within the loop are altered (Fig. 7). Backbone atoms from R234 to E239 shift, as a rigid body, ~2.0 Å away from the CP binding domain (compared to 2.4 Å for Y165F), making the active site more open to solvent. The interactions among R229, Q231, and E272, observed in $EcATC_T$, are eliminated. The region between E237 to F247 is also notably different from that in $EcATC_T$. As observed in

Y165F and chain C6 of other known T state structures (39), this segment of the main chain is tilted and the short helix from 237 to 242 now extends to residue 247. However, in contrast to Y165F, the side chain of residue Q246 is no longer capable of forming stabilizing interactions, as it is replaced by a glycine. This replacement, together with P237E, could increase the flexibility of the 240s loop. Residue K244 is the most affected by this rearrangement and its N_ζ atom has moved by 12.4 Å and is involved in a *Mp*ATC specific hydrogen bond with Q241 (alanine in EcATC). In addition, the *Ec*ATC T state stabilizing Y240-D271 hydrogen bond of 2.9 Å is weakened (4.2 Å).

In *Ec*ATC_T, R167 from the Asp binding domain is close to E50 from the CP binding domain. R229 is held in its T state location by a charge-charge interaction with E272. When substrates bind, R167 interacts with the α-carboxylate group of bound aspartate and E50, while R229 interacts with the β-carboxylate group and E233. As in the Y165F mutant, R167 of *Mp*ATC_T adopts a totally different conformation (Fig. 8). Instead of pointing toward E50, it moves out of the active site. However, in contrast to the situation in the Y165F mutant, it is not close enough to Y195 to form a hydrogen bond; instead it interacts with the Ser131 carbonyl oxygen. The side chain of R229 points towards the active site, yet is located too far to directly affect binding. These changes in the active site make it more open than the wild-type *Ec*ATC_T and hence would significantly alter interactions with bound substrate. Illustrative for this is the sulphate ion modelled into the electron density at the active site of the (three crystallographically unique) catalytic chains of *Mp*ATC. Their position is shifted (1.2-1.7 Å) relative to that of the phosphonate moiety of the phosphonacetamide in complex with the *Ec*ATC_T (pdb code 3AT1), which seems to be caused by an outward shift of the side chain of R55 (R54 in *Ec*ATC).

A region showing major differences with *Ec*ATC_T (maximum C_α distances of ~8 Å) is the 80s loop region. Interestingly, the first part (residues 75-78) of this loop shows more resemblance to the *Ec*ATC_R backbone conformation and is characterized by the lowest sequence homology. In contrast, the rest of the 80s loop region shows almost perfect sequence conservation and, indeed, its main-chain conformation, apart from a rigid-body like movement, is quite similar to the unliganded *Ec*ATC_T (Fig. 9). F103 seems particularly important in governing the rigid-body shift, firstly because it is the only drastic substitution in region 79-120, and secondly because its bulky side chain induces a tilt that is amplified by α-helix 3 (residues 89-99) towards the 80s loop region. Furthermore, the presence of G78 (Ala in *Ec*ATC) may ensure that S81 is still optimally positioned to interact with the phosphate of CP as observed in the corresponding complex of *Ec*ATC_T. The side chains of residues K84 and K85 (K83 and K84 of *E. coli*, resp.) point into different directions than in *Ec*ATC_T. The exact conformation of their side chains, however, is less certain due to poor electron density.

The regulatory chains share a good overall fit with the *Ec*ATC_T. Loop region 128-131, which was modelled only for chain D due to poor electron density, does not correspond to this region in the R1 chain of *Ec*ATC_T but has moved towards the C1-R1 interface, as previously described for the R6 chain of *Ec*ATC_T. At the allosteric domain-zinc domain interface, Y77 (*E. coli* numbering) has been replaced by a leucine (L78). The substitution of this residue, shown to be critical for the heterotropic effects in *Ec*ATC (38), however, appears to be (at least partially) compensated by the substitution of Phe for L105. The substitution of Pro for V107 could furthermore aid in the correct positioning of F106, thus mirroring the role of Y77 which fits as a pole in a hydrophobic pocket (Fig. 10). These substitutions clearly result in a (maximum 1.2 Å) shift of the backbone of region 106-113. More importantly this may explain the slight compression of the regulatory chain domains relative to *Ec*ATC_T. This relative compression or domain closure, however, is not accompanied by a large-scale movement of the catalytic trimers as seen upon binding of CTP to *Ec*ATC, which can be due to a small compensatory shift in the position of the C1-R1 interface. Consequently, the allosteric domains of the regulatory dimer fit to those of *Ec*ATC_T. Finally, the effector-binding residues are conserved well in sequence and conformation, which is also reflected in the similar effector binding mode in low-resolution complexes obtained by soaking with ATP (unpublished results).

Quaternary structure

The quaternary organization and the dimensions closely match those of the

unliganded *Ec*ATC$_T$. Due to the compression of the regulatory chain domains, the holoenzyme has a slightly smaller cross-section in the plane perpendicular to the pseudo threefold axis. Only subtle alterations are observed at the interfaces. At the C1-C2 type interfaces the K40-E37 (*E. coli* numbering) is replaced by an intramolecular salt bridge between the same residues (K41-E38 in *Mp*ATC). Furthermore in *Mp*ATC, the 80s loops are pairwize connected by a D75-N79 hydrogen bond (D75 is replaced by S74 in *Ec*ATC). The CP and Zn domains have shifted compared to *Ec*ATC$_T$, resulting in a slightly different location of the C1-R1 type interfaces without significantly altering the majority of interactions at these interfaces. However, the E204c-R130r salt bridge observed at the C6-R6 interface of *Ec*ATC$_T$ is no longer present in *Mp*ATC$_T$, firstly because the Asp domain is now located further from the C1-R1 type interface (due to the opening of the catalytic chain domains) and secondly by a more local shift observed for the α-helical region of residues 196-204 to a position that is similar as in the extreme T state mutant Y165F. Although formation of the typical *Ec*ATC$_R$ E200c-R130r interaction is not possible due to the R130T substitution (*E. coli* numbering), lysine 130 could replace the positive charge in *Mp*ATC$_R$. Furthermore, an extra negative charge (E201) is present that enhances the potential for charge-charge interactions during the allosteric transition.

R1-R6, the third large interface type, is also found to be very similar. One salt link is lacking due to the substitution of K for D39 (*E. coli*); yet, because of the absence of electron density for some charged side chains it is not clear if this loss is compensated for. Two smaller interfaces shown to be crucial for homotropic and heterotropic effects are the C1-C4 and R1-C4 type interfaces. At the C1-C4 type interface equivalent, shifts in main-chain positions of 240s and 160s loops prevent the loss of the E239-K164 intersubunit salt links. The presence of a bulkier Phe side chain at position 235 (*M. profunda*) forces the side chain of E239 into an alternative position, resulting in the loss of E239-Y165 at one C1-C4 interface (involving chain A). The most drastically altered by the changes in the 240s loop region is interface type R1-C4, with the rearrangement of 242-247 leading to a loss of both polar and apolar interactions. The important D236-K143 salt bridge (40), however, is still present although its distance is above the cut-off value of 4.0 Å at one of these interfaces (involving chain C). The extra salt bridges between E237 (equivalent to P237 in *Ec*ATC) and S111 (of the regulatory chain) may compensate for the weakening of this type of interface.

Thermostability

Composition analysis of the catalytic chain shows that significantly more negatively charged residues are present in *Mp*ATC than in *Ec*ATC (41 vs. 35), which is mainly due to the increased number of glutamate residues (24 vs. 14). The majority of these extra negative charges are located at the surface and are not directly involved in salt links as can be seen from the more negative surface potential compared to *Ec*ATC$_T$ (Fig. 11). The significantly higher ratio of Ile/Leu residues (20/31 vs. 15/38) for *Mp*ATC is noted as well. The number of residues known as thermolabile (Asn, Gln, Cys and Met) is not higher for *Mp*ATC: more cysteines (3 vs. 1) and less methionines (5 vs. 9) are present in *Mp*ATC compared to *Ec*ATC. Moreover, the number of Gly and Pro does not significantly differ. Finally, unlike the mesophilic ATCase from *E. coli* and the thermophilic enzyme from *Sulfolobus acidocaldarius*, the *Mp*ATC$_T$ has no solvent-accessible tunnel along the (pseudo-) threefold axis (Fig. 11).

Concerning the composition of the regulatory chain, a somewhat smaller number of negatively charged residues is observed, mainly due to the lower Asp content of *Mp*ATC (6 vs. 9 in *Ec*ATC). However, a remarkable observation is the significantly smaller number of Arg (1 vs. 8) which appears to be compensated for by the higher number of Lys residues (17 vs. 11). More polar residues are found in *Mp*ATC regulatory chain, mainly because of the significantly higher number of Thr (14 vs. 7). Finally, the Ile/Leu ratio is also somewhat higher (13/11 vs. 12/15).

The results of our analysis of molecular interactions and surface areas are summarized in Tables 3 and 4. Notable is the reduced number of intramolecular hydrogen bonds for the catalytic chains (the analysis is not presented for the regulatory chains because of missing side chains for the allosteric domains). The number of ion pairs is similar, as is the percentage apolar surface area. We have found, furthermore, that the ion pair networks are mostly conserved and that their number and size are not significantly different. From Table 4 it follows that the surface area buried in the different types of interfaces is

smaller for the C1-C2, R1-C4 and C1-C4 type interfaces. For the latter this has resulted mainly from a reduction in polar buried surface area. The number of salt links is smaller for the C1-C2 type interfaces.

DISCUSSION

Upon investigating the structural features of cold adaptation it is important to minimize the effects of random mutational drift and varying selective pressure by studying enzymes with highly homologuous sequences (41). In this study we have characterized the (pure) recombinant ATCase from a strict psychrophile, i.e. *M. profunda* (doubling time ~5.5 h at 2 °C, under atmospheric pressure), which is highly related to its B-class mesophilic counterpart from *E. coli*; the *pyrB* and *pyrI* genes from these proteobacterial backgrounds are 73 % and 53 % identical at the amino acid level. Sequence comparison indicates a strict conservation of all active site residues that contact CP or L-Asp. With the exception of the Y89F substitution, all 12 residues proposed to be involved in effector binding, and the four cysteine residues involved in binding the Zn^{2+} ion in *Ec*ATC are strictly conserved as well. ATCase sequence analysis therefore is consistent with the phylogenetically close relationship between the mesophile *E. coli* and the psychrophile *M. profunda,* rationalizing our choice to use the *M. profunda* ATCase as a model to study the structural basis for cold activity and cold regulation.

Measurements of the enzyme's temperature dependence of activity indicate typical behaviour for a cold-adapted enzyme. At 4 °C, *M. profunda* ATCase is considerably more active than its *E. coli* counterpart, while both enzymes are comparably active at 30 °C.

The currently accepted hypothesis (16,42) suggests that psychrophilic enzymes have to increase their flexibility in order to perform catalysis at low temperatures, the enhanced flexibility being generated by the generally lower stability of the protein structure. Investigation of time-dependent thermo-inactivation reveals that *Mp*ATC is notably less stable than *Ec*ATC. Interestingly, the class-A ATCase (in dialyzed cell-free extracts) of the Antarctic strain *TAD*1 was shown to be as stable as the *E. coli* enzyme (43). For example, its activity in crude *TAD*1 extracts did not show any decrease after one hour of preincubation at 60 °C. It was proposed that the high activity of the *TAD*1 enzyme has resulted from a discrete modification, localized at the catalytic site, not affecting the global thermostability of the entire protein. Conversely, in the case of *Mp*ATC, our kinetic experiments indicate that the high specific activities at low temperatures may have been attained at the cost of a lower global stability of the protein. Accordingly, it is expected that a mobile and flexible active site binds it substrate weakly, and indeed, most psychrophilic enzymes have higher K_m values than their mesophilic counterparts (16). This also seems to be the case for *Mp*ATC with a slightly (~2.7 times) higher $S_{0.5}$-concentration for carbamoyl phosphate and a significantly (~13.5 times) higher $S_{0.5}$-concentration for aspartate than *Ec*ATC. The *TAD*1 ATCase exhibits a higher K_m for carbamoyl phosphate, yet it was demonstrated to potentially result from a higher catalytic constant. Our finding that the concentration range at which the bisubstrate analogue PALA activates the ATCase reaction is about tenfold higher for *Mp*ATC than for *Ec*ATC indicates that the high $S_{0.5}$-value originates from a lower substrate affinity *per se*, and not from mechanistic differences. It has been suggested (44) that a significant increase in overall rather than local flexibility would lead to enzymes with a weak affinity for substrate. Alternatively, in a two-state allosteric system, a relative stabilization of the low-affinity (low-activity) T state would cause the same effect. However, a higher cooperativity for aspartate binding would be expected, as expressed by the Hill coefficient, which is not observed for *Mp*ATC. That this cooperativity reflects the existence of T and R states with different affinity and/or catalytic activity is, indeed, clearly demonstrated by the stimulation of activity by PALA at subsaturating aspartate concentrations. Sedimentation velocity experiments furthermore indicate that this T-R transition is accompanied by a global expansion of the molecule similar to that documented for *Ec*ATC (unpublished results).

In general, the unliganded *Mp*ATC_T exhibits a good overall fit with the tertiary and quaternary structural features of the unliganded *Ec*ATC_T. However, a more scrutinized analysis reveals a number of similarities with the structure of the unliganded *Ec*ATC mutant Y165F, which has been coined as an extreme T state (39). The relative domain opening of the catalytic chain domains and the conformational changes at the 240s region are the most prominent features of

Y165F mutant and are indeed shared by *Mp*ATC$_T$.
Furthermore, this mutation resulted in a high $S_{0.5}$-value (90 mM), ~16 fold higher than the WT *Ec*ATC value (45), which was attributed to the structural differences of this extreme T state, and more specifically to the opening of the active site. Although Y165 and relevant interaction partners are conserved in *Mp*ATC, it is possible that the high [$S_{0.5}$] for aspartate may be caused by this extreme T state rather than by a higher overall degree of flexibility. Indeed, it is possible that the 240s region adopts this conformation because of a local loss of interactions, like in the Y165F mutant, or a local increase in flexibility, thereby causing an opening of the catalytic domains. The P237E and Q246G substitutions are potential mediators of increased flexibility of the 240s loop. The latter substitution also results in a loss of interactions compared to the Y165F T state. It is important to note that the T state currently described for *Mp*ATC probably represents the true unliganded T state of this enzyme and is considered to be an extreme version only on comparison with the structure of *E. coli* T state. E.g. it is even more sensitive to CTP inhibition than *Ec*ATC.

It may be tempting to infer a high flexibility from the overall high absolute temperature factors (B-factors), yet their significance is generally blurred by the data quality and refinement procedures followed. The highest relative B factors can be found in loop regions, specifically the 80s and 240s loops for the catalytic chains. The zinc domains are comparable with the catalytic chains, yet the electron density maps (and relative B factors) were significantly worse for the allosteric domains. Although neither of the regulatory chain domains is involved in short-range crystal contacts, this does indicate that the allosteric-zinc interface presents a rather loose association (e.g. in contrast with the stabilized allosteric-zinc interface in the *S. acidocaldarius* T state ATCase structure). The breakdown of the threefold symmetry (as present in the *E. coli* and *S. acidocaldarius* holoenzymes) for *Mp*ATC, because of conformational heterogeneity, may yet be another indication for a higher flexibility. Indeed, Endrizzi *et al.* (46) have found that the release of constraints imposed by the holoenzyme resulted in a breakdown of the threefold symmetry in the isolated and unliganded catalytic trimer. Conformational heterogeneity was manifested as varying interdomain hinge angles and a lack of interpretable electron density for the 80s and 240s loop regions, and therefore (on comparing with *Mp*ATC$_T$) can be considered as the extreme and limiting case of the loosening of steric constraints of the holoenzyme. Furthermore, from our analyses of the different subunit interfaces of the *Mp*ATC$_T$ it follows that the C1-C2, C1-C4 and R1-C4 interfaces have a decreased buried surface area. The C1-C2 interface has less salt links. The E204-R130 salt link typical for the C6-R6 type interface in the unliganded and CTP-liganded *Ec*ATC is missing from the Cn-Rn type interfaces of *Mp*ATC because of an arginine to threonine substitution. Interestingly, this is one of the few important violations of the twofold non-crystallographic symmetry in the *E. coli* holoenzyme and may explain why this twofold symmetry is maintained in *Mp*ATC. A site-directed mutagenesis study by Stebbins and Kantrowitz (47), furthermore, has shown that eliminating this salt link causes the temperature for dissociation of the holoenzyme into catalytic and regulatory subunits to drop by 6 °C.

Ionpair networks, proposed to be major thermostabilizing elements in both hyperthermophilic ATCase crystal structures (13,14), do not seem to be reduced in the cold active enzyme. We did find a potential role for electrostatics in that the higher glutamate content of the catalytic chains results in more (negative) surface charges that are not involved in salt links. An excess of negative surface charges has been reported earlier on psychrophilic enzymes (16,48) and results in this case in an interesting picture of thermo-adaptation across the full physiological temperature range (Fig. 11), i.e. the mesophilic ATCase (middle) is flanked by a psychrophilic ATCase (top) with a more negative surface potential and a hyperthermophilic ATCases (bottom) with more positive surface potential. This may reflect the different role for electrostatics in heat and cold adaptation: in the hyperthermophile the formation of ion pairs and ion pair networks contributes to the integrity of the subunits and their interfaces, whereas in the psychrophile electrostatics may enhance protein solvation (negative charges of carboxylates in proteins are more strongly hydrated) and active site flexibility (49).

When considering the adaptation to cold it has to be noted that *M. profunda* (originally isolated from deep-sea sediments at a depth of 2815 m), has been characterized as a moderately

piezophile (growth optimum 20-24 MPa), meaning its ATCase in principle could have undergone adaptation to high pressure. Pressure denaturation commonly requires pressures above 100 MPa and is not expected to be a significant stress phenomenon (50). Pressure-induced dissociation of oligomeric proteins and inhibition of enzyme activity is observed well within the biologically relevant range of pressures. Decreased k_{cat} values were suggested to be an adaptive trait of enzymes to the high pressure of the deep sea (51), possibly resulting from a higher conformational rigidity, as thermostability and piezostability often seem to be linked (52). In the case of *Mp*ATC it is thus quite possible that the adaptation to the cold might have been hampered under the influence of pressure, leading to more modest increases in k_{cat}, and explaining the reduced catalytic efficiency. Investigation of steady-state kinetics under high pressure is needed to clarify whether *Mp*ATC does exhibit piezophilic properties. Indeed, the T-R equilibrium may well be disturbed by high pressure, as was shown for the *E. coli* ATCase where it was shifted towards the R state (the maximal effect being observed at 120 MPa) ((53)).

*Mp*ATC is inhibited by CTP and UTP and exhibits a dual response towards ATP: activation below 3 mM and inhibition above this concentration, which is practically complete at 8 mM. Competitive inhibition by all nucleotides for carbamoyl phosphate has been documented for *Ec*ATC and predicts a binding strength that is approximately one order of magnitude lower compared to carbamoyl phosphate (54). Therefore, it usually does not manifest itself in nucleotide saturation curves because of the high carbamoyl phosphate concentrations used. Nonetheless, we have observed this effect even in the presence of saturating concentrations of carbamoyl phosphate (and near saturating aspartate). Furthermore, similar UTP concentrations do clearly not yield the same inhibition at similar concentrations: more than 30 mM UTP is needed to reach similar inhibition. It appears then that the ATP inhibitory effect is heterotropic, an observation not previous reported for B-class ATCases. Heterotropic ATP inhibition has been reported for the engineered Y77F mutant of the *Ec*ATC regulatory chain by Van Vliet *et al.* (38). Strikingly, the otherwise well-conserved allosteric-zinc domain interface of *Mp*ATC presents a mirror image of this situation, with a leucine as the equivalent of Y77 and a phenylalanine at the opposite (zinc domain-) side replacing a leucine. Analysis of the *Ec*ATC mutant suggests that the mutation promotes an increased hydrophobic packing at the allosteric-zinc interface, resulting in a compression of the regulatory chain domains. A compression relative to *Ec*ATC_T is indeed observed in *Mp*ATC_T, yet the enzyme is not devoid of homotropic interactions as is the Y77F mutant. Furthermore, the remarkably low electron density for the allosteric domains is indicative for a rather loose arrangement at the allosteric-zinc interface. The weakening of hydrophobic interactions at cold temperatures may explain this observation. Irrespective of the origin of this apparently higher flexibility, it may well have the important function of more efficiently transmitting effector signals at low temperatures. Indeed, when considering proteins as complex statistical ensembles of conformational states, a redistribution of these states resulting in functional changes (as induced by ligand binding in functional cooperativity) is likely favored for regions with high conformational heterogeneity (or thus flexibility). A similar reasoning has formed the basis for mapping allosteric regions on proteins by statistical thermodynamics computer simulations (55). Interestingly, the effector binding strength (indicated by the *Mp*ATC E_{50} values) is not attenuated by cold adaptation (cf. Table 1).

In summary, we have found that the *M. profunda* ATCase is able to maintain (most of) the catalytic and regulatory properties of its mesophilic counterpart at its physiological temperature, at the cost of a lower thermal stability. Our study confirms the importance of subunit association and of surface electrostatics in determining the stability of ATCases (12, 14, 56, 57). A general weakening of subunit interfaces may also lower the kinetic barrier for the T to R transition, whereas surface electrostatics may play a role in conferring flexibility required for high activity at low temperatures. Our analysis of the *Mp*ATC_T crystal structure finally suggests that the concept of flexibility for cold activity can be extended towards cold regulation.

REFERENCES

1. Allewell, N. M. (1989) *Annu Rev Biophys Biophys Chem* **18**, 71-92
2. Hervé, G. (1989) *Aspartate transcarbamylase from Escherichia coli*. Allosteric enzymes (Hervé, G., Ed.), CRC Press, Boca Raton
3. Lipscomb, W. N. (1994) *Adv Enzymol Relat Areas Mol Biol* **68**, 67-151
4. Wild, J. R., and Wales, M. E. (1990) *Annu Rev Microbiol* **44**, 193-218
5. Reichard, P., and Hanshoff, G. (1956) *Acta Chem. Scand.* **10**, 548-560
6. Pardee, A. B., and Yates, R. A. (1956) *J Biol Chem* **221**, 757-770
7. Gerhart, J. C., and Schachman, H. K. (1968) *Biochemistry* **7**, 538-552
8. England, P., and Herve, G. (1994) *Biochemistry* **33**, 3913-3918
9. Ke, H. M., Lipscomb, W. N., Cho, Y. J., and Honzatko, R. B. (1988) *J Mol Biol* **204**, 725-747
10. Labedan, B., Xu, Y., Naumoff, D. G., and Glansdorff, N. (2004) *Mol Biol Evol* **21**, 364-373
11. Bethell, M. R., and Jones, M. E. (1969) *Arch Biochem Biophys* **134**, 352-365
12. Vitali, J., Vorobyova, T., Webster, G., and Kantrowitz, E. R. (2000) *Acta Crystallogr D Biol Crystallogr* **56 (Pt 8)**, 1061-1063
13. Van Boxstael, S., Cunin, R., Khan, S., and Maes, D. (2003) *J Mol Biol* **326**, 203-216
14. De Vos, D., Van Petegem, F., Remaut, H., Legrain, C., Glansdorff, N., and Van Beeumen, J. J. (2004) *J Mol Biol* **339**, 887-900
15. Xu, Y., Nogi, Y., Kato, C., Liang, Z., Ruger, H. J., De Kegel, D., and Glansdorff, N. (2003) *Int J Syst Evol Microbiol* **53**, 533-538
16. Feller, G., and Gerday, C. (2003) *Nat Rev Microbiol* **1**, 200-208
17. Gerday, C., Aittaleb, M., Bentahir, M., Chessa, J. P., Claverie, P., Collins, T., D'Amico, S., Dumont, J., Garsoux, G., Georlette, D., Hoyoux, A., Lonhienne, T., Meuwis, M. A., and Feller, G. (2000) *Trends Biotechnol* **18**, 103-107
18. De Vos, D., Hulpiau, P., Vergauwen, B., Savvides, S. N., and Van Beeumen, J. (2005) *Acta Crystallogr F Structural Biology and Crystallization Communications* **61**, 279-281
19. Bradford, M. M. (1976) *Anal Biochem* **72**, 248-254
20. Prescott, L. M., and Jones, M. E. (1969) *Anal Biochem* **32**, 408-419
21. Gilboa, H., Kogut, M., Chalamish, S., Regev, R., Avi-Dor, Y., and Russell, N. J. (1991) *J Bacteriol* **173**, 7021-7023
22. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr D Biol Crystallogr* **53**, 240-255
23. (1994) *Acta Crystallogr D Biol Crystallogr* **50**, 760-763
24. Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *Journal of Applied Crystallography* **26**, 283-291
25. Kabsch, W., and Sander, C. (1983) *Biopolymers* **22**, 2577-2637
26. Mcdonald, I. K., and Thornton, J. M. (1994) *J Mol Biol* **238**, 777-793
27. Richards, F. M. (1977) *Annual Review of Biophysics and Bioengineering* **6**, 151-176
28. Kleywegt, G. J., and Jones, T. A. (1994) *Acta Crystallograph D Biol Crystallogr* **50**, 178-185
29. Hubbard, S. J., and Argos, P. (1995) *Protein Engineering* **8**, 1011-1015
30. Voronoi, G. F. (1908) *J. Reine Angew. Meth.* **82**, 1-14
31. DeLano, W. L. (2004) *Abstracts of Papers of the American Chemical Society* **228**, U313-U314
32. Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. (2003) *Nucleic Acids Research* **31**, 3381-3385

33. Baker, D. P., Stebbins, J. W., DeSena, E., and Kantrowitz, E. R. (1994) *J Biol Chem* **269**, 24608-24614
34. England, P., Leconte, C., Tauc, P., and Herve, G. (1994) *Eur J Biochem* **222**, 775-780
35. Wedler, F. C., and Gasser, F. J. (1974) *Arch Biochem Biophys* **163**, 57-68
36. Collins, K. D., and Stark, G. R. (1971) *J Biol Chem* **246**, 6599-6605
37. Liu, L., Wales, M. E., and Wild, J. R. (2000) *Arch Biochem Biophys* **373**, 352-360
38. Van Vliet, F., Xi, X. G., De Staercke, C., de Wannemaeker, B., Jacobs, A., Cherfils, J., Ladjimi, M. M., Herve, G., and Cunin, R. (1991) *Proc Natl Acad Sci U S A* **88**, 9180-9183
39. Ha, Y., and Allewell, N. M. (1998) *Proteins* **33**, 430-443
40. Chan, R. S., Sakash, J. B., Macol, C. P., West, J. M., Tsuruta, H., and Kantrowitz, E. R. (2002) *J Biol Chem* **277**, 49755-49760
41. Sheridan, P. P., Panasik, N., Coombs, J. M., and Brenchley, J. E. (2000) *Biochim Biophys Acta* **1543**, 417-433
42. Zavodszky, P., Kardos, J., Svingor, and Petsko, G. A. (1998) *Proc Natl Acad Sci U S A* **95**, 7406-7411
43. Sun, K., Camardella, L., Di Prisco, G., and Herve, G. (1998) *FEMS Microbiol Lett* **164**, 375-382
44. D'Amico, S., Claverie, P., Collins, T., Georlette, D., Gratia, E., Hoyoux, A., Meuwis, M. A., Feller, G., and Gerday, C. (2002) *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **357**, 917-924
45. Wales, M. E., Hoover, T. A., and Wild, J. R. (1988) *J Biol Chem* **263**, 6109-6114
46. Endrizzi, J. A., Beernink, P. T., Alber, T., and Schachman, H. K. (2000) *Proc Natl Acad Sci U S A* **97**, 5077-5082
47. Stebbins, J. W., and Kantrowitz, E. R. (1989) *J Biol Chem* **264**, 14860-14864
48. Russell, N. J. (2000) *Extremophiles* **4**, 83-90
49. Kumar, S., and Nussinov, R. (2004) *Chembiochem* **5**, 280-290
50. Gross, M., and Jaenicke, R. (1994) *Eur J Biochem* **221**, 617-630
51. Somero, G. N. (1992) *Annu Rev Physiol* **54**, 557-577
52. Mombelli, E., Shehi, E., Fusi, P., and Tortora, P. (2002) *Biochim Biophys Acta* **1595**, 392-396
53. Herve, G., Schmitt, B., and Serre, V. (2004) *Cell Mol Biol (Noisy-le-grand)* **50**, 347-352
54. Porter, R. W., Modebe, M. O., and Stark, G. R. (1969) *J Biol Chem* **244**, 1846-1859
55. Freire, E. (2000) *Proc Natl Acad Sci U S A* **97**, 11680-11682
56. Edge, V., Allewell, N. M., and Sturtevant, J. M. (1988) *Biochemistry* **27**, 8081-8087
57. Durbecq, V., Thia-Toong, T. L., Charlier, D., Villeret, V., Roovers, M., Wattiez, R., Legrain, C., and Glansdorff, N. (1999) *Eur J Biochem* **264**, 233-241
58. Beernink, P. T., Yang, Y. R., Graf, R., King, D. S., Shah, S. S., and Schachman, H. K. (2001) *Protein Sci* **10**, 528-537
59. Wild, J. R., Loughrey-Chen, S. J., and Corder, T. S. (1989) *Proc Natl Acad Sci U S A* **86**, 46-50
60. Van Boxstael, S., Maes, D., and Cunin, R. (2005) *Febs J* **272**, 2670-2683

FOOTNOTES

* This research has been supported by the Fonds voor Wetenschappelijk Onderzoek Vlaanderen (FWO grants 3G006896 and 3G044899). D. De Vos is a research fellow of the same institution. We gratefully acknowledge access to beamline BW7A at the EMBL/DESY Hamburg Outstation.

[1]The abbreviations used are: ATCase, Aspartate carbamoyltransferase; *Ec*ATC, *Mp*ATC and *Sa*ATC stand for *E. coli*, *M. profunda* and *S. acidocaldarius* ATCase respectively, and, when present, subscripts "T" and "R" denote their corresponding T and R states; CP, carbamoyl phosphate; E_{50}, effector concentration at which half the maximal effect is observed; n_H, Hill coefficient; PALA, *N*-phosphonacetyl-L-aspartate; rmsd, root-mean-square deviation; SEM, standard error of the mean.

TABLES AND FIGURES

Table 1. Summary of the kinetic parameters and of the nucleotide response of the *M. profunda* psychrophilic ATCase. The counterpart values for the mesophilic enzyme of *E. coli* (58,59) and the thermophilic enzyme of *P. abyssi* (13,60) are given for comparison.

Table 2. Diffraction data processing and refinement statistics for the unliganded *M. profunda* ATCase.

Table 3. Comparative analysis of *E. coli* and *M. profunda* T-state ATCase subunit structures.

Table 4. Comparative analysis of different types of subunit interfaces of the *M. profunda* and *E. coli* T-state ATCases.

Figure 1. Amino acid sequence alignment of *M. profunda* and *E. coli* ATCase. Residues in boxes are conserved, residues shaded in red are strictly conserved. Top secondary structures are based on chain C of *Mp*ATC, while bottom secondary structures are derived from chain A of *Ec*ATC; β-strands are shown as arrows, 3_{10} helices are represented by thinner coils than α-helices and are labeled separately. (A) Catalytic chains of *Mp*ATC (M._profunda_c) and *Ec*ATC (E._coli_c). (B) Regulatory chains of *Mp*ATC (M._profunda_r) and *Ec*ATC (E._coli_r) .

Figure 2. Thermostability of the *M. profunda* ATCase holoenzyme. As a function of time, the residual activity was measured at 4°C under saturating standard assay conditions after incubation at either 50 °C (squares), 55 °C (triangles), 60 °C (inverted triangles), or 65 °C (diamonds). Residual activity defined as A/A_0*100, where A is the activity of the preincubated sample and A_0 the activity of untreated sample.

Figure 3. Substrate saturation curves of the *M. profunda* ATCase. The activity was measured in the presence of increasing concentrations of L-Asp (squares) or CP (triangles). The L-Asp saturation curve was determined in the presence of 5 mM CP. The CP saturation curve in the presence of 300 mM L-Asp; the obtained kinetic parameters for CP should thus be considered as apparent values.

Figure 4. Effects of varying the concentration of the bisubstrate analog, PALA, on the relative activity of the *M. profunda* ATCase. PALA titration curves were measured in the presence of half-saturating [L-Asp] (100 mM (triangles)) and under low L-Asp saturation (20 mM (squares)). In all assays, the CP concentration was 0.5 mM. Relative activity defined is as A/A_0*100; where A is the activity in the presence of PALA and A_0 the activity in its absence.

Figure 5. Effect of nucleotides on the activity of *M. profunda* ATCase. Reaction mixtures containing low-saturating (20 mM) L-Asp and half-saturating (0.5 mM) CP-concentrations were titrated as a function of [ATP] (squares), [UTP] (diamonds), or [CTP] (triangles). Relative activity defined as A/A_0*100; where A is the activity in the presence of the nucleotide and A_0 the activity in its absence.

Figure 6. Overview of *Mp*ATC$_T$ along the pseudo-threefold symmetry axis. Each Cn-Rn dimer is depicted in a separate colour.

Figure 7. Stereoview of the superposed C$_\alpha$-traces of *Mp*ATC$_T$ (black), *Ec*ATC$_T$ (green) and *E. coli* Y165F mutant (cyan) 240s regions. Side chains are indicated (as sticks) for

residues 189, 192, 230, 235 and 247 (*M. profunda* numbering) and are further discussed in the text.

Figure 8. Superposition of the active site residues of *Mp*ATC$_T$ (colour-coded according to atom type) and *Ec*ATC$_T$ (all atoms coloured green) and of residue R167 of the *E. coli* Y165F mutant (cyan). Lysine residues 84 and 85 were omitted for clarity. The active site sulphate ion of *Mp*ATC and the phosphonacetamide of the corresponding complex of *Ec*ATC$_T$. are depicted as sticks and colour-coded according to atom type, with yellow denoting phosphorus and magenta denoting sulfur.

Figure 9. Stereoview of the superposed C$_\alpha$-traces of *Mp*ATC$_T$ (black), *Ec*ATC$_T$ (green) and *Ec*ATC$_R$ (red) 80s regions. Side chains are indicated (as sticks) for residues 84, 85 and 103 (*M. profunda* numbering), H3 stands for α-helix 3. These structural elements are further discussed in the text.

Figure 10. Superposition of *Mp*ATC$_T$ (blue) and *Ec*ATC$_T$ (green) regulatory chains. Important residues of the allosteric-zinc interdomain interface are depicted as sticks. (A) Overview. (B) Detailed view.

Figure 11. Electrostatic surface potentials (contoured at +/- 70 kT/e) of ATCases at different temperature levels. Positive potentials are shown in blue, negative potentials are shown in red. At the left: top view (along the crystallographic threefold axis) of the catalytic trimers of *Mp*ATC$_T$ (A), *Ec*ATC$_T$ (B), *Sa*ATC$_T$ (C). At the right (D,E,F): side view (along the crystallographic twofold axis) of the respective holoenzymes.

Table 1

Variable Substrate	*M. profunda* holoenzyme (pH 7.5; 4°C)[a,b]			*E. coli* holoenzyme (pH 7.0; 30 °C)			*P. abyssi* holoenzyme (pH 8.0; 55 °C)[e]		
	V_{max} (U)	$S_{0.5}$ (mM) n_H	$V_{max}/S_{0.5}$	V_{max} (U)[c]	$S_{0.5}$ (mM)[c] n_H[c]	$V_{max}/S_{0.5}$[c]	V_{max} (U)	$S_{0.5}$ (mM) n_H	$V_{max}/S_{0.5}$
Asp	5.20 ± 0.17	101 ± 10 (85 ± 5) 1.85 ± 0.12 (2.21 ± 0.16)	0.05	10.8	7.5 1.7	1.44	9.9	2.70 1.70	3.67
CP	4.87 ± 0.15	0.53 ± 0.04 (0.35 ± 0.04) 1.26 ± 0.09 (1.27 ± 0.15)	9.19	10.8	0.2 ND	54	3.2[f]	0.005[f] 1.0[f]	640[f]
NTP effector[a]		Relative activity (%)	E_{50} (μM)		Relative activity (%)[d]	E_{50} (μM)[d]		Relative activity (%)	E_{50} (μM)
ATP		~260 (~300)	145 ± 34 (165 ± 25)		186	800		350	250
CTP		~20 (~20)	3.39 ± 0.50 (4.01 ± 0.60)		31	~150		50	3.0
UTP		~60 (~ND)[g]	1650 ± 430 (ND)		90-95	NR[h]		65	600

[a] The data are derived from colorimetric assays as described in the Materials section. For the determination of the kinetic parameters, either CP was kept constant at 5 mM, or L-Asp was kept constant at 300 mM, meaning that the kinetic parameters obtained for CP saturation should be considered as apparent values. The parameters were determined by using a non-linear least squares routine using the Hill equation. For the determination of the nucleotides effector responses, L-Asp and CP were kept constant at low-saturating concentrations (20 mM) and half-saturating concentrations (0.5 mM), respectively. Relative activity is defined as A/A_0*100; where A is the activity in the presence of the nucleotide and A_0 the activity in its absence. All the data are the average of two independent determinations.

[b] Values between parentheses are those obtained for the native holoenzyme in cell free crude *M. profunda* lysate under identical conditions as in [a].

[c] data were taken from ref. (58).

[d] data were taken from ref. (59).

[e] data were taken from ref. (60).

[f] values obtained at 37 °C (59).

[g] ND: not determined.

[h] NR: not reported.

Table 2

	***Mp*ATC**
Diffraction data	
Resolution (Å)	20-2.85
Space group	$P3_221$
Unit cell parameters	a = 129.2
	b = 129.2
	c = 207.2
Refinement	
R_{work}/R_{free} (%)	21.1/25.7
Rms bond length deviations (Å)	0.011
Rms bond angle deviations (°)	1.33
No. atoms	
Protein (catalytic/regulatory chains)	7149/2759
Solvent	36
Sulphate	15
Average B-factor (Å^2)	
Protein (catalytic/regulatory chains)	53.9/56.2
Solvent	37.4
Sulphate	87.8
Ramachandran plot	
% most favoured regions	86.2
% additional allowed	13.6
% disallowed	0.2

Table 3

	E. coli	*M. profunda*
Catalytic subunit		
Secondary structure content		
α-helices (%)	34	36
β-strands (%)	15	14
Hydrogen bonds		
Total	278	268
Per residue	0.89	0.86
Salt bridges		
Total	18	20
Per residue	0.058	0.065
Total ASA (Å^2)	13,477	13,039
Apolar ASA (Å^2)	7,437 (55)	7,187 (55)
VDW-volume (Å^3)	31,850	31,400
Packing density	0.74	0.74
Cavities	1	0.33
Regulatory subunit		
Secondary structure content		
α-helices (%)	12	11
β-sheets (%)	33	29
Hydrogen bonds		
Total	101	
Per residue	0.66	
Salt bridges		
Total	11.5	
Per residue	0.075	
Total ASA (Å^2)	8,492	
Apolar ASA (Å^2)	4,453 (52)	
VDW-volume (Å^3)	15,120	
Packing density	0.76	0.76
Cavities	0	0

All values are averages of all subunits of that specific type; ASA: accessible surface area; values between parentheses are percentages of total ASA.

Table 4

Interface type	***E. coli***	***M. profunda***
C1-R1 type interfaces		
Buried surface area		
Total (Å^2)	1362	1300
Apolar (Å^2)	648 (48)[a]	644 (50)
Hydrogen bonds	10.5	11.3
Ion pairs	3.5	3
C1-C2 type interfaces		
Buried surface area		
Total (Å^2)	2365	2000
Apolar (Å^2)	1420 (60)	1140 (57)
Hydrogen bonds	8	8
Ion pairs	10	7.7
C1-C4 type interfaces		
Buried surface area		
Total (Å^2)	718	608
Apolar (Å^2)	247 (34)	278 (46)
Hydrogen bonds	3	3.3
Ion pairs	3	4
C1-R4 type interfaces		
Buried surface area		
Total (Å^2)	666	576
Apolar (Å^2)	397 (60)	360 (63)
Hydrogen bonds	2	2.7
Ion pairs	1	1

[a] Values between parentheses are percentages of apolar buried surface area compared to total buried surface area.

Figure 1

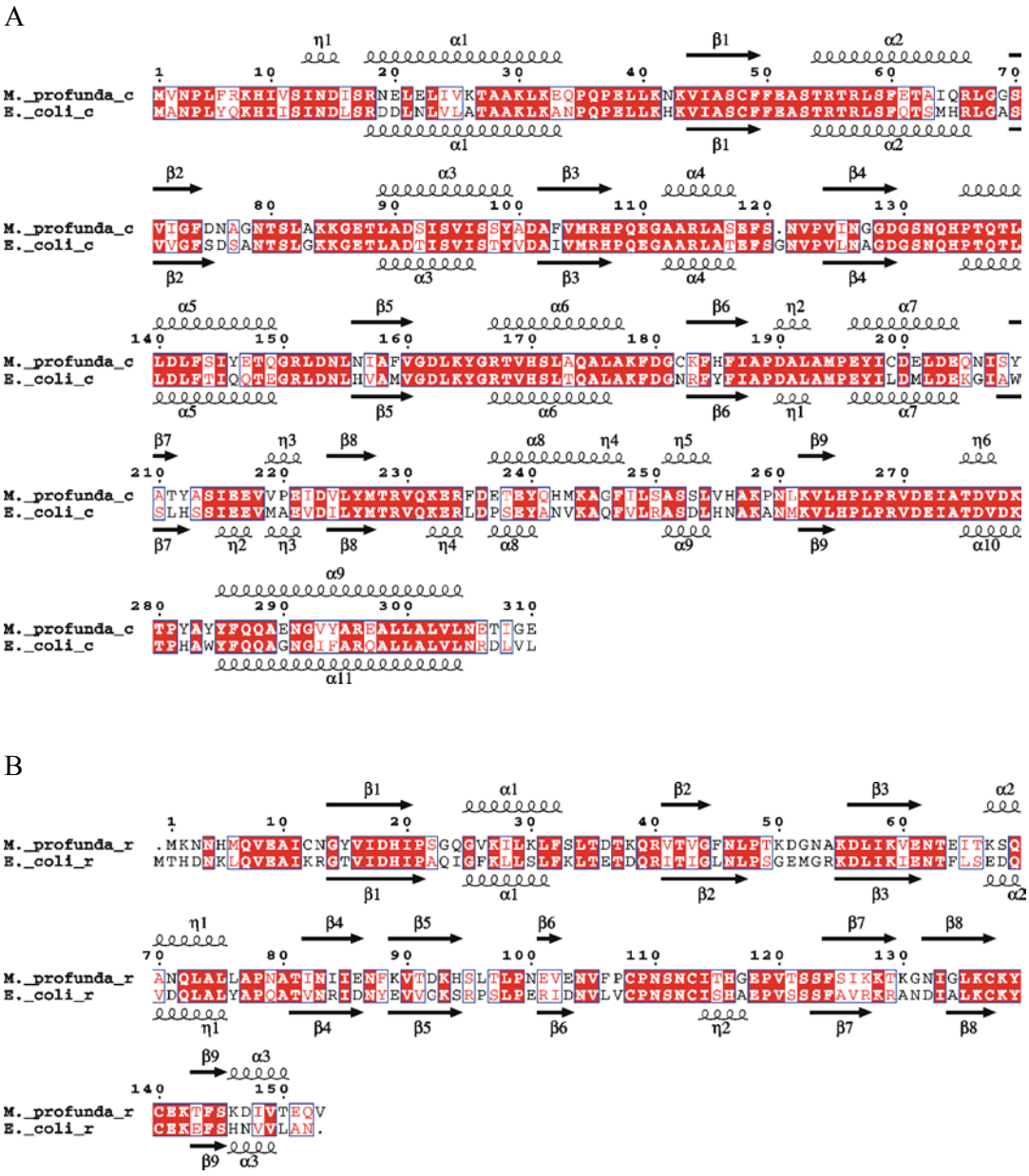


Figure 2

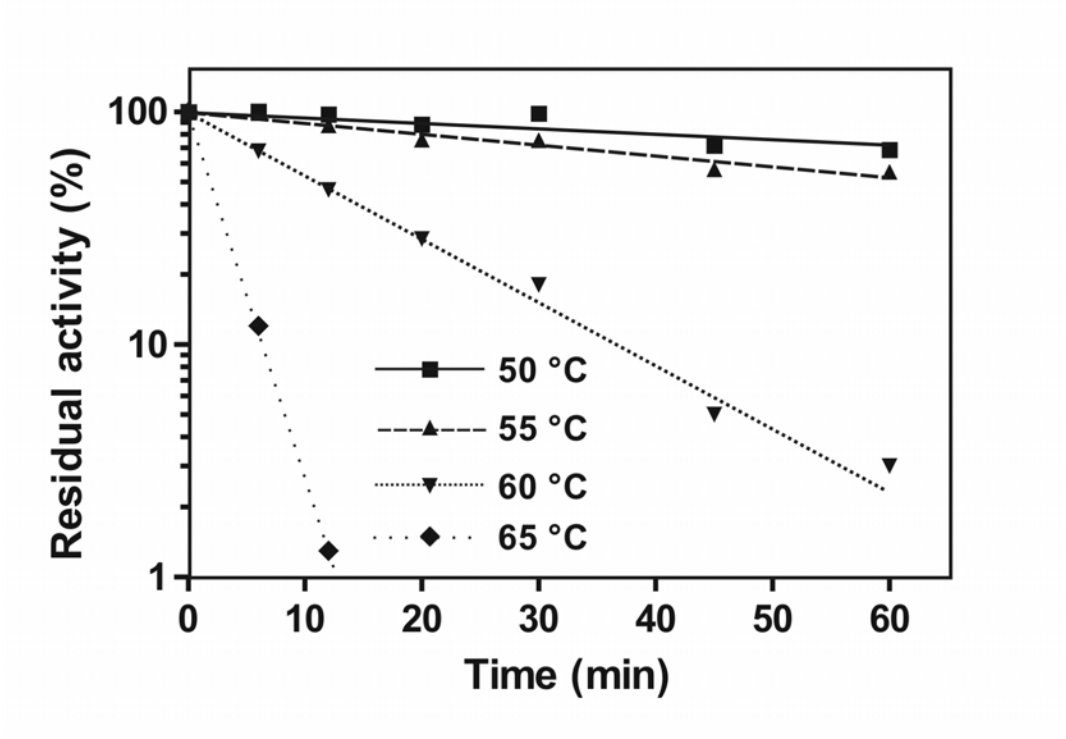

Residual activity (%)
100
10
1
0
10
20
30
40
50
60
Time (min)
50 °C
55 °C
60 °C
65 °C

Figure 3

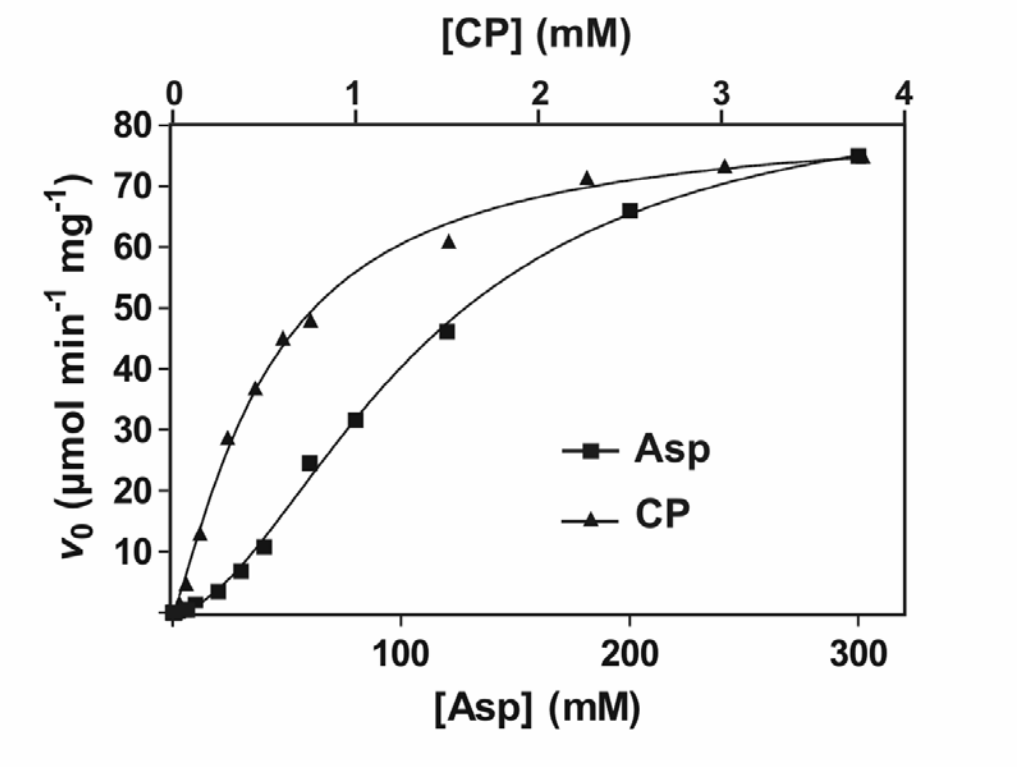

[CP] (mM)
0
1
2
3
4
80
70
60
50
40
30
20
10
v_0 (µmol min^{-1} mg^{-1})
Asp
CP
100
200
300
[Asp] (mM)

Figure 4

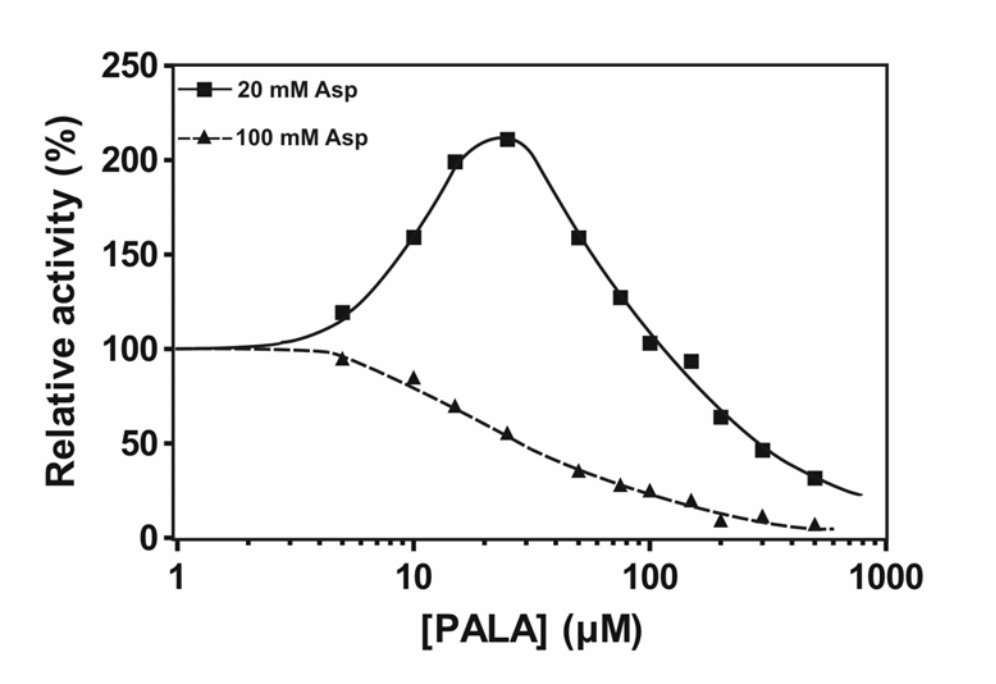

250
200
150
100
50
0
Relative activity (%)
20 mM Asp
100 mM Asp
1
10
100
1000
[PALA] (µM)

Figure 5

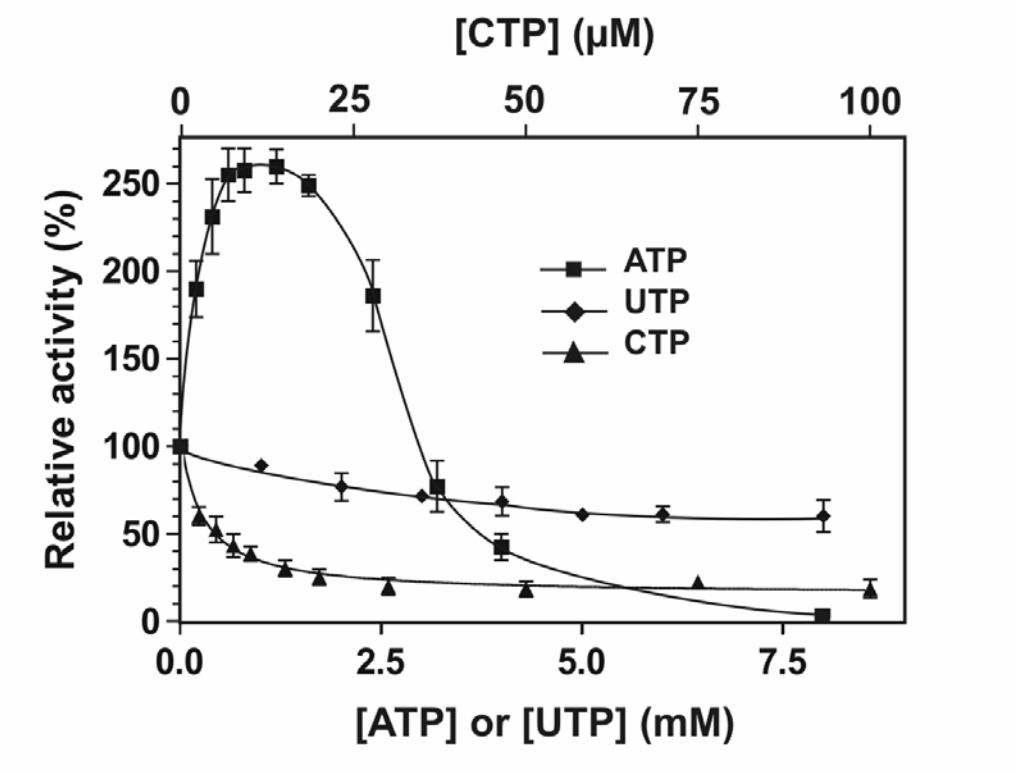

[CTP] (μM)
0
25
50
75
100
Relative activity (%)
250
200
150
100
50
0
ATP
UTP
CTP
0.0
2.5
5.0
7.5
[ATP] or [UTP] (mM)

Figure 6

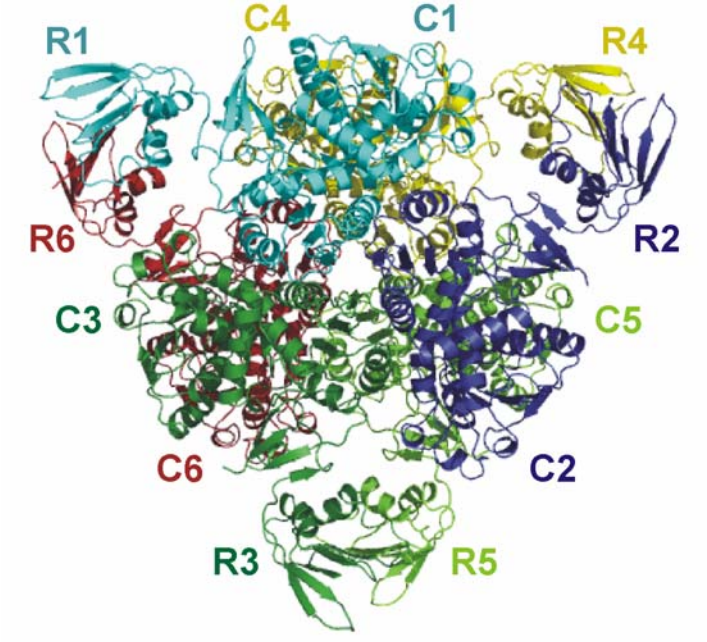


Figure 7

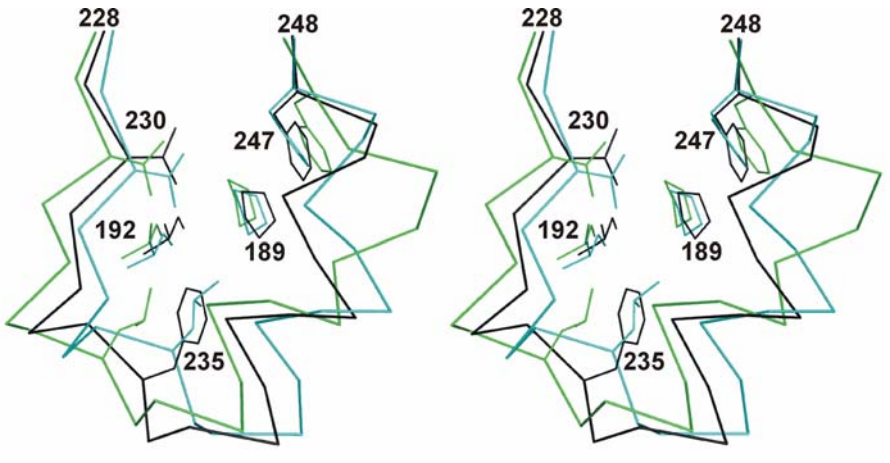


Figure 8

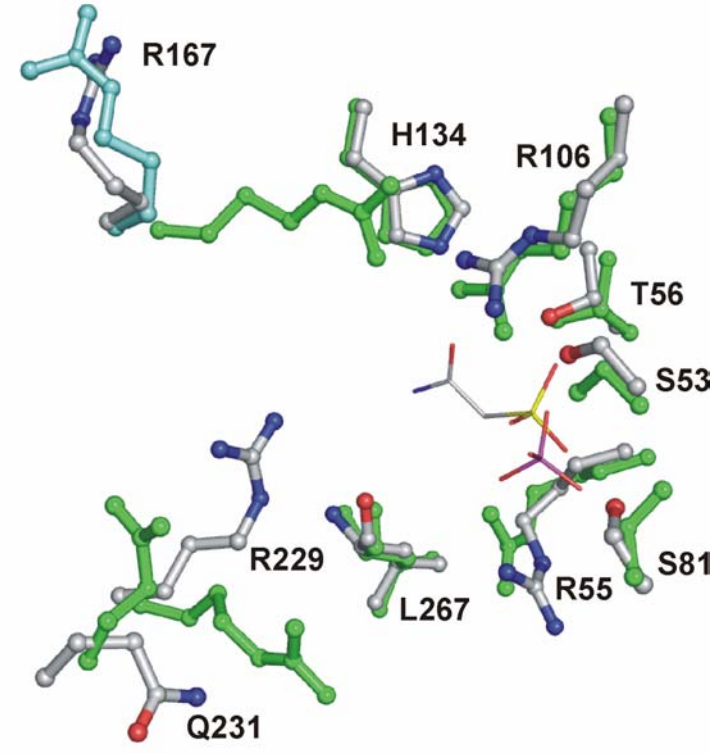


Figure 9

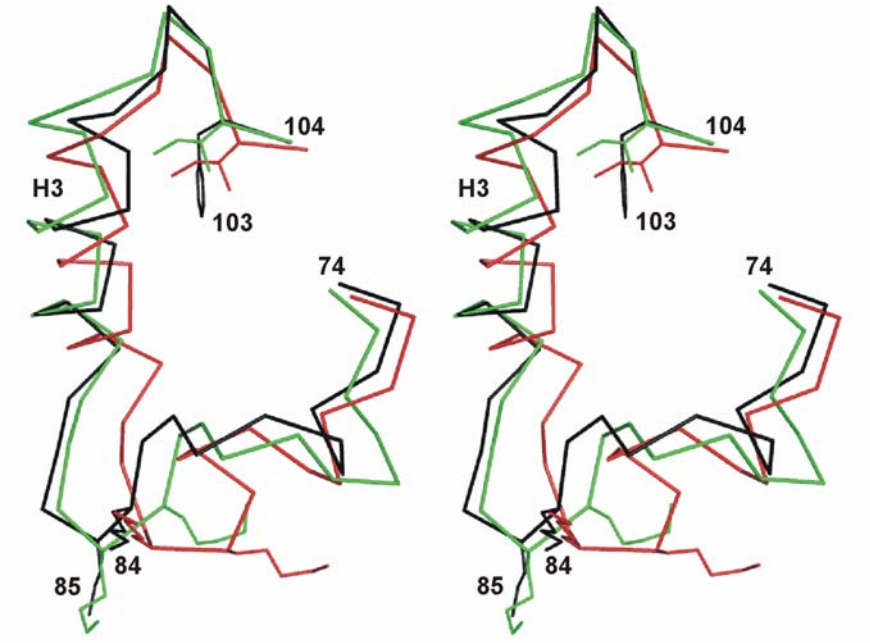


Figure 10

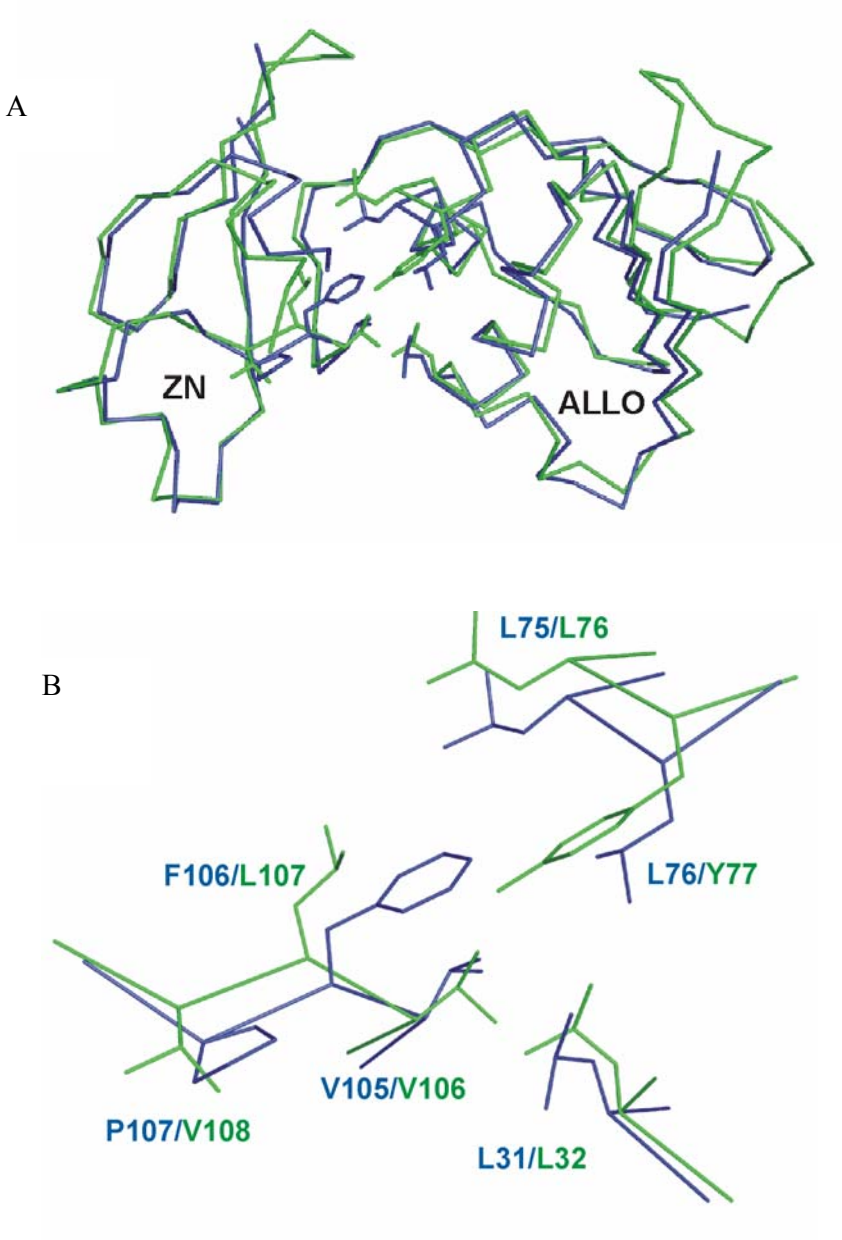
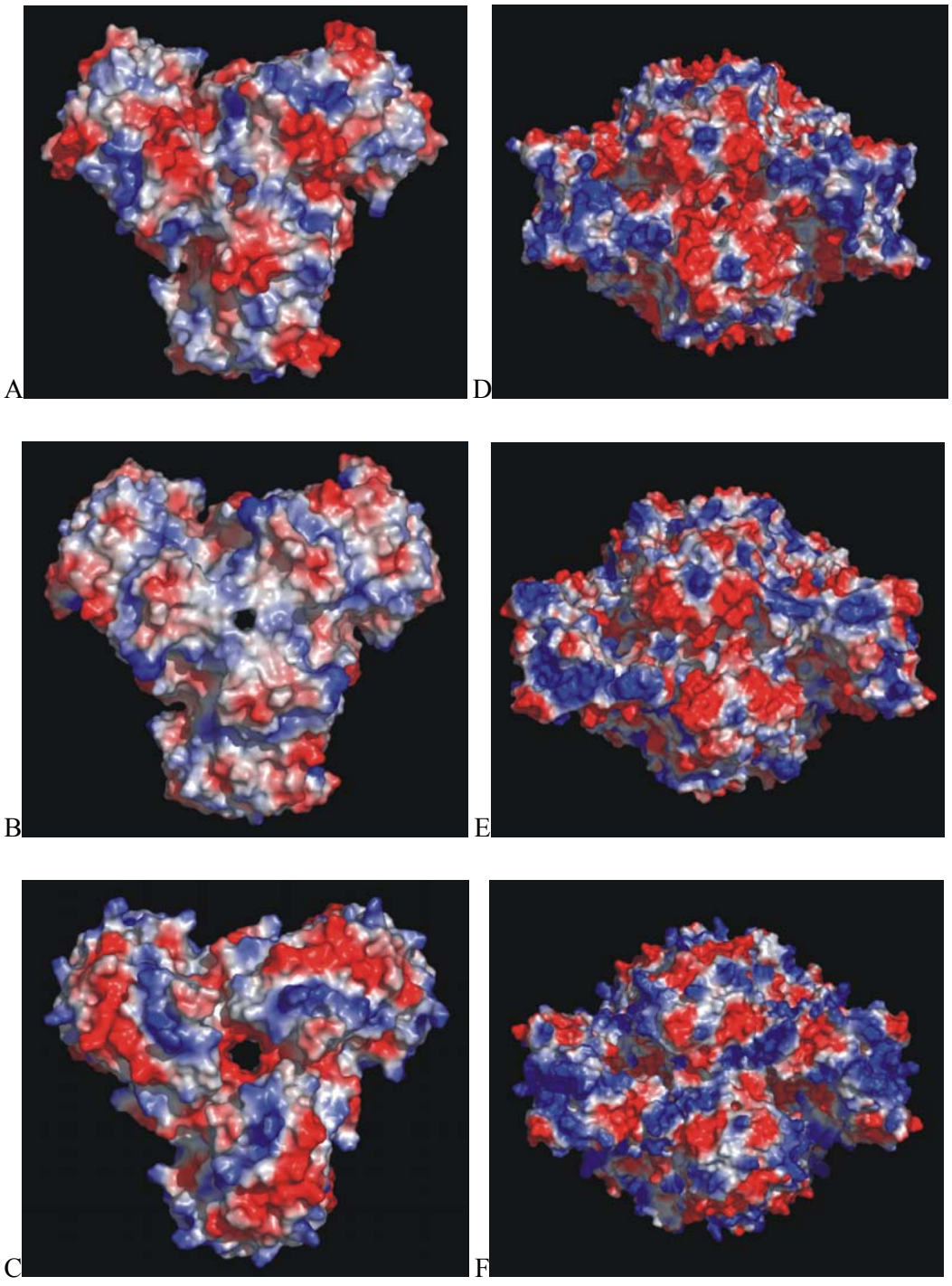

A
ZN
ALLO
B
L75/L76
F106/L107
L76/Y77
V105/V106
P107/V108
L31/L32

Figure 11



7 Supplementary experiments on the allosteric regulation of *S. acidocaldarius* and *M. profunda* aspartate carbamoyltransferase

Supplementary experiments on the allosteric regulation of *S. acidocaldarius* and *M. profunda* aspartate carbamoyltransferase

Crystal structures of effector complexes

S. acidocaldarius and *M. profunda* ATCase both exhibit interesting differences in regulatory behaviour compared to the *E. coli* model system. In contrast to *E. coli* ATCase the *S. acidocaldarius* enzymes is activated by CTP (Durbecq *et al.*, 1999). In fact, activation by CTP has previously been reported for other B-class ATCases (e.g. from *Serratia marcescens* and *Proteus vulgaris*), yet in these cases the amplitude of this effect was smaller than the ATP activating effect. Therefore, CTP can still be considered as an inhibitor of ATCase activity in the presence of both effectors (Wild *et al.*, 1988). The stronger activating effect of CTP (compared to ATP) on the *S. acidocaldarius* enzyme seems, however, to violate the metabolic logic. This also appears to be the case for the *M. profunda* enzyme at high ATP concentrations, which are shown to inhibit the enzyme more than CTP. A fairly obvious explanation for the *M. profunda* enzyme is that cellular levels of ATP are lower than 3 mM, i.e. at concentrations where activation has been demonstrated by our experiments (cf. previous chapter). Applying a similar reasoning to *S. acidocaldarius* ATCase, one has to assume that cellular CTP levels are at least 2-fold lower than ATP levels. Moreover, to offset the strong CTP activation, CTP levels below 0.1 mM are required. Whereas the first assumption appears to be valid in *Enterobacteriaceae*, CTP levels were found to be around 1 mM in this bacterial family (Wild *et al.*, 1988). To investigate whether these aberrant nucleotide responses are reflected in structural differences, the respective effector complexes were determined: ATP in complex with *M. profunda* ATCase, and CTP in complex with *S. acidocaldarius* ATCase.

Inspection of the electron density maps of the CTP complex revealed well-defined contours at the expected allosteric binding site (corresponding to a high ligand occupancy) (Figure 7.1), this at different pH values (cf. Materials and methods). However, the diffraction data of the crystals obtained at pH 9.0 reached a resolution of maximum 3.5 Å (data statistics not shown), the crystals at pH 4 reached 2.6 Å. Notably, there were no significant structural differences between refined structures at pH 4 and pH 9. A sulphate ion with clear interaction partners was also modelled at the effector binding site. Superpositions were performed of the regulatory chains of the low-pH complex with deposited nucleotide complexes of the *E. coli* ATCase. The best fit for the nucleoside part of the effector molecule was found with that of the R6 regulatory chain of the *E. coli* T state in complex with CTP, the best fit for the triphosphate moiety was found with that of the R6 chain of the *E. coli* R state in complex with CTP (Figure 7.2). A superposition of the complete dodecameric complex of unliganded and CTP-liganded *S. acidocaldarius* ATCase indicates no significant differences in quaternary structure. Despite the high overall tertiary structural similarity of the unliganded and CTP-liganded form (expressed by an rmsd of 0.33 Å for the catalytic-regulatory chain dimer), a number of local differences were observed, some of which were at sites previously implicated in the allosteric behaviour of the *E. coli* enzyme. E.g., CTP binding induces a conformational shift of helix H1' of the regulatory chain and small changes at the R1-C4 interface are observed. A more detailed structural analysis is in progress.

Due to poor electron density (at a resolution of 3.1 Å) an ATP molecule was modelled only at the effector binding site of chain E of *M. profunda* ATCase soaked with ATP. Unlike the situation in $Ec\mathrm{ATC}_T$, ATP binding does not induce an expansion of the catalytic trimers along the three-fold axis. The electron density is weak for the purine base, but relatively well-defined for the ribose and triphosphate moiety of ATP. This was already noted in low resolution studies of nucleotide binding to the *E. coli* ATCase (Honzatko *et al.*, 1979).

Superpositioning gave the best fit with the ATP bound to the R1 chain of the R state *E. coli* ATCase (Figure 7.3).

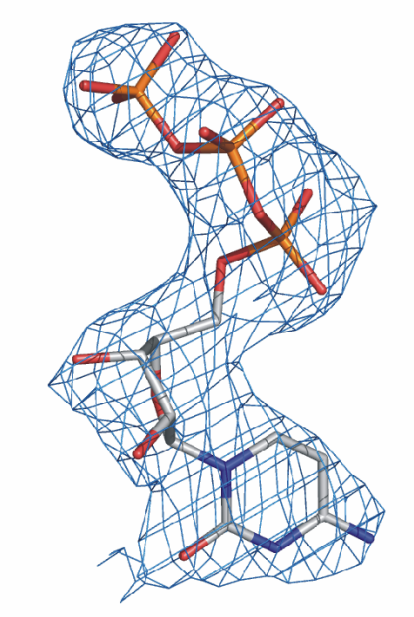


Figure 7.1 The $2F_o$-F_c electron density map for CTP bound at the regulatory chain of *Sa*ATC$_T$ (colour-coded according to atom type), contoured at 1.0σ.

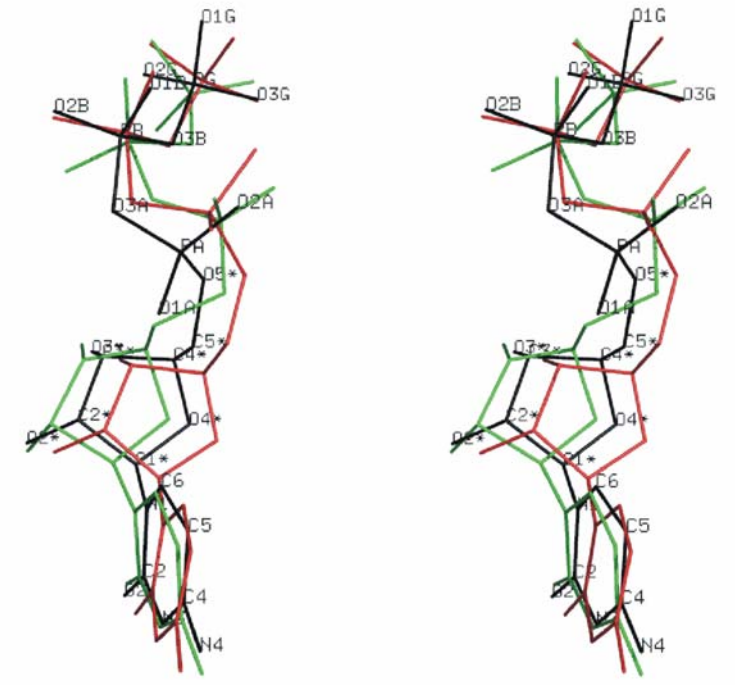


Figure 7.2 Stereoview of CTP molecules (atom labels included) bound at the effector-binding sites of *Sa*ATC$_T$ (coloured black), of the *Ec*ATC$_T$ R6 chain (coloured green) and of the *Ec*ATC$_R$ R6 chain (coloured red), obtained after superposing the respective regulatory chains.

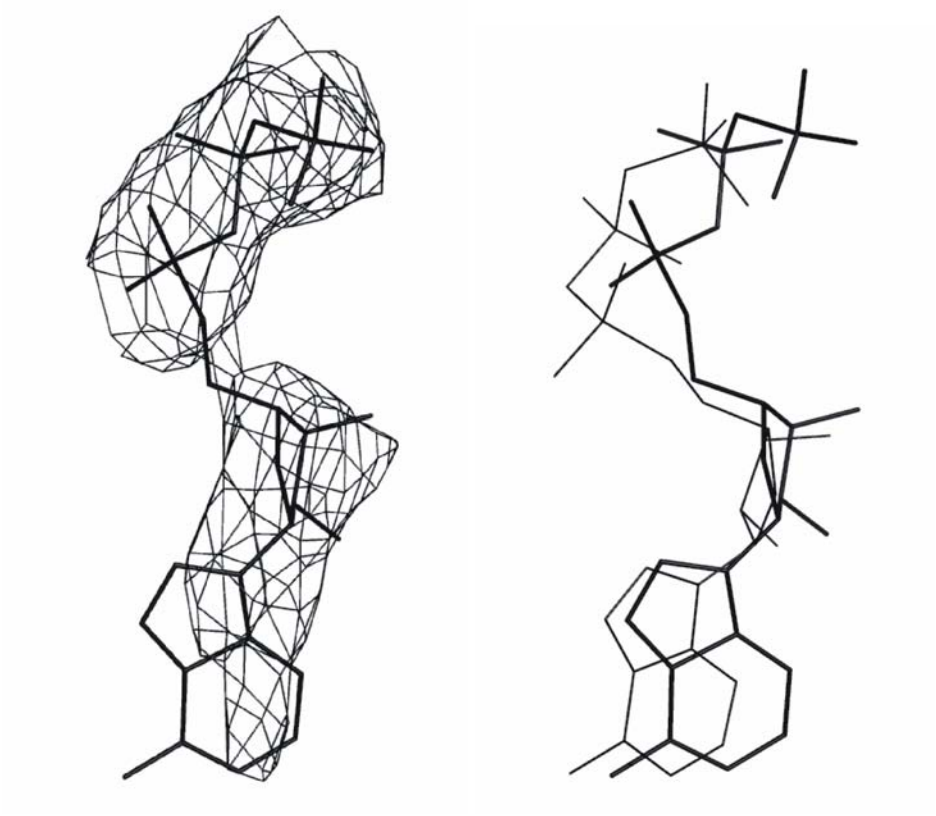


Figure 7.3 (a) $2F_o$-F_c electron density map for ATP bound at regulatory chain "E" of *Mp*ATC$_T$, contoured at 0.8σ. (b) View of ATP bound at the effector binding sites of MpATCT (thick lines) and the R1 chain of EcATCR (thin lines) obtained after superposition of their regulatory chains.

Materials and methods

Crystals of the CTP liganded *S. acidocaldarius* ATCase and of the ATP liganded *M. profunda* ATCase were obtained by transferring crystals, grown as described (De Vos *et al.*, 2004; De Vos *et al.*, 2005), into equilibrated droplets of mother liquor with gradually increasing concentrations of CTP and ATP and over periods of 72 and 48 hours, respectively. For the *S. acidocaldarius* enzyme, crystals obtained at a different pH (Bicine buffer at pH 9.0 instead of acetate at pH 4.0) were also used for soaking. For both complexes the final soaking concentration was 10 mM (at higher concentrations the crystals developed cracks). Diffraction data were collected from single crystals on a MARCCD detector (MarResearch) using 0.80 Å synchrotron radiation at the X13 beamline at the EMBL/DESY in Hamburg. All data were collected at 100 K using 25 % glycerol as the cryo-protectant. Intensity data were integrated, scaled and reduced to structure factor amplitudes with the HKL suite of programs (Otwinowski and Minor, 1997). Data collection statistics are shown in Table 7.1.

The structures of both complexes were isomorphous with the respective unliganded structures and were refined, starting from this structure (pdb codes 1PG5 and 2BE7), using rigid body refinement as the first step, followed by the same procedures as for the unliganded *M. profunda* ATCase, except for TLS refinement. The structure of the ATP-liganded MpATC has some additionally missing residues compared to the starting model (residue N79 from chain A and B, residues 13,14 and 49 from chain D). Atomic coordinates and related structure factors have been deposited in the RCSB Protein Data Bank, Rutgers University, New

Brunswick, NJ (http://www.rcsb.org/) with identification codes 2BE9 for the CTP complex and 2BE8 for the ATP complex.

Table 7.1 Diffraction data processing and refinement statistics for the CTP-liganded *S. acidocaldarius* ATCase and the ATP-liganded *M. profunda* ATCase.

	***Sa*ATC:CTP**	***Mp*ATC:ATP**
Diffraction data		
Resolution (Å)	20-2.6 (2.69-2.60)	20-3.1 (3.21-3.10)
Space group	$P6_322$	$P3_221$
Unit cell parameters	a = 131.9	a = 130.0
	b = 131.9	b = 130.0
	c = 139.1	c = 207.7
No. Reflections		
Total	120,254	119,332
Unique	22,175	36,850
R_{merge}[a] *(%)*	7.9 (51.3)	7.0 (40.9)
I/σ (I)	17.2 (2.2)	14.3 (2.8)
Completeness (%)	98.4 (97.2)	99.3 (98.1)
Refinement		
R_{work}/R_{free} (%)	22.3/27.1	20.8/24.2
Rms bond length deviations (Å)	0.011	0.010
Rms bond angle deviations (°)	1.44	1.21
No. atoms		
Protein (catalytic/regulatory)	2351/1144	7109/2687
Solvent	77	31
Sulphate	5	15
CTP/ATP	29	31
Average B-factor (Å^2)		
Protein (catalytic/regulatory)	44.4/50.7	76.1/77.1
Solvent	43.4	46.6
Sulphate	52.0	93.3
CTP/ATP	73.2	80.0
Ramachandran plot		
% most favoured regions	88.4	85.5
% additional allowed	11.6	14.3
% disallowed	0.0	0.2[b]

Values between parentheses reflect the data in the highest resolution shell.

[a] $R_{merge} = \Sigma_h\Sigma_i|I(h,i)-\langle I(h)\rangle|/\ \Sigma_h\Sigma_i I(h,i)$, where I(h,i) is the intensity of the ith measurement of reflection h and <I(h)> is the average value from multiple measurements.

[b] The peptide link between L267 and P268 is in a *cis* conformation.

Ultracentrifugation studies

As crystal structures of the respective R states of the *S. acidocaldarius* and *M. profunda* ATCases could not be determined, the question remains whether T and R states deduced from kinetic studies correspond to different quaternary states and, furthermore, whether a global quaternary transition with an amplitude comparable with that observed for *E. coli* ATCase does effectively occur.

In order to study the difference in shape between the T and R state of both ATCases, sedimentation velocity experiments were carried out. In this type of ultracentrifugation experiment a high angular speed is used to cause a rapid sedimentation of solute towards the cell bottom. The rate of sedimentation, or sedimentation coefficient s (expressed in svedberg, S) for molecules with the same molecular mass is inversely proportional to the frictional coefficient, which is related to the hydrodynamic volume and shape of the particle. A comparison of the S values in the absence and presence of saturating concentrations of the bisubstrate analogue PALA (*N*-phosphonacetyl-L-aspartate), known to induce the T-R quaternary transition in *E. coli* ATCase, would indicate whether the fractional decrease in sedimentation velocity ($\Delta s/s_0$) is similar to that reported for *E. coli* ATCase.

The results (Table 7.2) reveal similar changes as for *E. coli* ATCase, thus confirming that a swelling, comparable in amplitude, is taking place in both enzymes upon T to R transition at 20 °C. For *S. acidocaldarius* ATCase, the concentration dependence of the PALA effect was not known in advance, and thus it could not be shown that the swelling effect was 'saturated' at the highest value. At the 2.5 μM PALA concentration, the swelling of *S. acidocaldarius* ATCase appears to be intermediate. However, PALA, at this concentration, is not able to induce a swelling for the *M. profunda* enzyme in accordance with our kinetics data, which indicate that a PALA concentration of minimum 10 μM is required for activation. The inherent instability of the *M. profunda* ATCase resulted in less accurate data due to the formation of aggregates.

Table 7.2 Effect of PALA on the sedimentation coefficients of *S. acidocaldarius*, *M. profunda* and *E. coli* ATCase.s_0 is the sedimentation coefficient of the unliganded enzyme. Δs is the difference between the sedimentation coefficients of unliganded and PALA-liganded enzyme.

PALA (μM)	**s, sedimentation coefficient (S)**	**$\Delta s/s_0$, fractional change in sedimentation velocity (%)**
S. acidocaldarius ATCase		
0	10.9	
2.5	10.7	-1.8
1000	10.5	-3.7
M. profunda ATCase		
0	10.9	
2.5	11	0.9
50	10.6	-2.8
1000	10.6	-2.8
E. coli ATCase		
0	11.2	
300	10.9	-2.7

Materials and methods

Sedimentation velocity experiments were performed in a Beckman Optima XL-A analytical ultracentrifuge, equipped with an AN-60 Ti analytical rotor (Analis, NV, Belgium). The protein was at a concentration of 2 $mg.ml^{-1}$ in 50 mM Tris-HCl, pH 7.5, 300 mm NaCl for *M. profunda* ATCase and 150 mM for *S. acidocaldarius* ATCase. The concentration along the cell was measured with an optical absorption detection system at 280 nm. The experiments were carried out at an angular speed of 10,000 r.p.m. at 20 °C. Sedimentation coefficients were determined with Beckman software based on a nonlinear least squares fit, using the MixedFit program for Windows (Stafford, 1992). The calculated sedimentation coefficients were not corrected for density, viscosity of the solvent, and the presence of PALA. The experiments were performed in duplicate.

Part III
pXyl

8 Introduction to pXyl

8.1 Xylanases and their substrate

Xylanases (EC 3.2.1.8; endo-1,4-*β*-xylan xylanohydrolases) are *O*-glycoside hydrolases which catalyze the endohydrolysis of 1,4-*β*-D-xylosidic linkages in xylan. They are a widespread group of enzymes involved in the production of xylose, a carbon source for cell metabolism and plant cell infection by plant pathogens, and are produced by a plethora of organisms like bacteria, algae, fungi, protozoa, gastropods, and arthropods. Xylan is a major structural polysaccharide in plant cells and is the second most abundant polysaccharide in nature, accounting for approximately one-third of all renewable organic carbon on earth (Kulkarni *et al.*, 1999). In contrast to cellulose, a homopolymer of D-glucose that has the same structure in all plants, xylan has a more complex structure, varying between different plant species. *β*-1,4-xylans are heteropolysaccharides, composed of a homopolymeric backbone chain of 1,4-linked *β*-D-xylopyranose units (with a variable degree of polymerization). The backbone can be substituted with *O*-acetyl, *α*-L-arabinofuranosyl, *α*-1,2-linked glucuronic acid or 4-O-methylglucuronic acids. In marine algae, linear xylans composed of a mixture of *β*-1,3 and *β*-1,4 linkages also occur, and in some red algae, *β*-1,3-xylan has been found (Nunn *et al.*, 1973; McDowell, 1967).

Due to its heterogeneity and complexity, the complete hydrolysis of xylan requires a large variety of cooperatively acting enzymes (Puls *et al.*, 1987; Biely, 1985; Subramaniyan and Prema, 2002) (Figure 8.1). Endo-1,4-*β*-D-xylanases (E.C. 3.2.1.8) randomly cleave the xylan backbone, *β*-D-xylosidases (E.C. 3.2.1.37) cleave xylose monomers from the non-reducing end of xylo-oligosaccharides and xylobiose, while removal of the side groups is catalyzed by *α*-L-arabinofuranosidases (E.C. 3.2.1.55), *α*-D-glucuronidases (E.C. 3.2.1.139) acetylxylan esterases (E.C. 3.1.1.72), ferulic acid esterases (E.C. 3.1.1.73) and p-coumaric acid esterases (E.C. 3.1.1.-). Complete xylanolytic enzyme systems, including all of the activities listed, have been found to be quite widespread among fungi and bacteria (Sunna and Antranikian, 1997; Belancic *et al.*, 1994; Elegir *et al.*, 1994). The ecological niches of these organisms are diverse and widespread and typically include environments where plant material accumulates and deteriorates, as well as in the rumen of ruminants (Collins *et al.*, 2004).

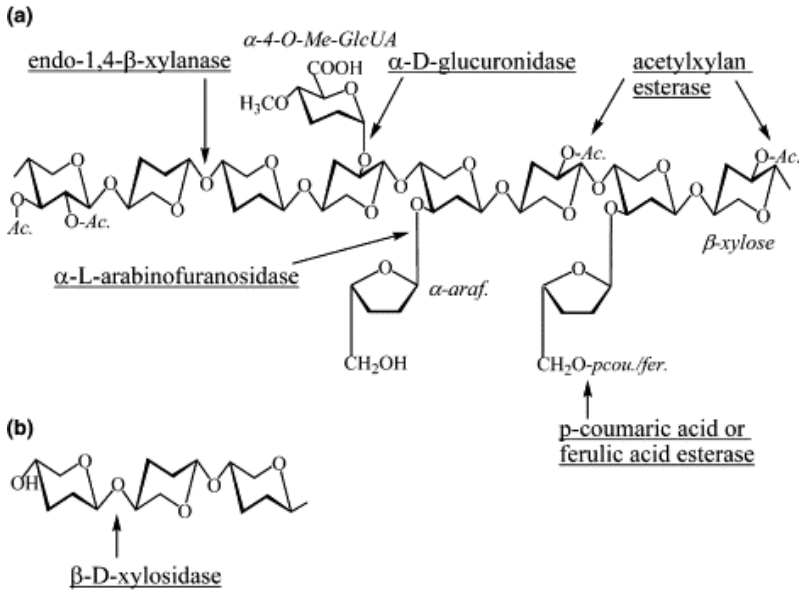


Figure 8.1 (a) Structure of xylan and the sites of its attack by xylanolytic enzymes. The backbone of the substrate is composed of 1,4-*β*-linked xylose residues. Ac., Acetyl group; *α*-araf., *α*-arabinofuranose; *α*-4-O-Me-GlcUA, *α*-4-O-methylglucuronic acid; pcou., p-coumaric acid; fer., ferulic acid. (b) Hydrolysis of xylo-oligosaccaride by β-xylosidase. Adapted from Sunna and Antranikian (1997).

Within the standard (primary structure based) classification system of glycoside hydrolases (CAZY, the carbohydrate-active server at http://afmb.cnrs-mrs.fr/~cazy/CAZY/; Coutinho and Henrissat, 1999), xylanases are normally reported as being confined to families 10 (formerly F) and 11 (formerly G). However, enzymes with xylanase activity have been reported for families 5, 7, 8 and 43 (Collins *et al.*, 2004). Families 5, 7, 10, 11 contain enzymes which catalyze hydrolysis with retention of anomeric configuration, with two glutamate residues (one as a general acid-base catalyst, one as the nucleophile) being implicated in the catalytic mechanism in all cases (CAZY). This indicates a double-displacement mechanism, in which a covalent glycosyl-enzyme intermediate is formed and subsequently hydrolyzed via oxocarbenium-ion-like transition states (Rye and Withers, 2000; Zechel and Withers, 2000; McCarther and Withers, 1994). In contrast, enzymes in families 8 and 43 typically operate with inversion of the anomeric centre and a glutamate and aspartate are believed to be the catalytic residues (CAZY). Inverting enzymes function by a single displacement reaction in which one carboxylate provides for a general acid-catalyzed leaving group departure and the second functions as the general base, activating a nucleophilic water molecule to attack the anomeric carbon, thereby cleaving the glycosidic bond and leading to an inversion of the configuration at the anomeric carbon (Rye and Withers, 2000; Zechel and Withers, 2000; McCarther and Withers, 1994). Table 8.1 gives an overview of the structural fold and catalytic mechanism of the different xylanase containing glycosidase families.

Table 8.1 Properties of glycosidase hydrolase families containing enzymes with a demonstrated activity on xylan. Adapted from Collins *et al*. (2004)

Glycoside hydrolase family	Fold	Catalytic mechanism
5	$(\beta/\alpha)_8$	Retaining
7	β-jelly roll	Retaining
8	$(\alpha/\alpha)_6$	Inverting
10	$(\beta/\alpha)_8$	Retaining
11	β-jelly roll	Retaining
43	5-blade β-propeller	Inverting

The main applications of xylanases are in the paper and pulp industries (Kulkarni *et al.*, 1999). The brown colour in paper can be due to the presence of residual lignin and lignin-carbohydrate molecules which are typically removed with the aid of a chlorine bleaching process. However, environmental regulations have put a restriction on the usage of chlorine in the bleaching process. It has been found that treatment of the pulp with xylanase makes the pulp more permeable for the subsequent chemical extraction of this lignin, thereby decreasing the amounts of chemicals needed for the bleaching. Another application lies in the baking industry, where xylanases act on the gluten fraction of the dough and help in the even redistribution of the water content of the bread. Dietary hemicelluloses have little nutritional significance for non-ruminant organisms because these lack the appropriate digestive enzymes. The use of xylanases along with other hemicellulases can help to decrease the viscosity of the food in the gut, as well as to increase the nutritive value.

8.2 pXyl

A description of the biochemical properties of the recombinant *P. haloplanktis* xylanase (pXyl) has been given by Collins *et al*. (2002), a description of the crystal structure of the unliganded WT enzyme has been given by Van Petegem *et al*. (2003). In order to obtain a better understanding of the work presented on the subject, a summary of the main findings is given here.

The mature xylanase gene has a signal sequence coding for 21 amino acids, and a predicted molecular mass of 45,982 Da. After cleavage of the signal sequence, a protein of 405 residues remains. The amino acid sequence shows highest identity to *Cytophaga hutchinsonii* endoglucanase Y (34%), and contains the glycosyl hydrolase family 8 fingerprint. The thermophilic endoglucanase CelA from *Clostridium thermocellum* (23 % sequence identity) was chosen as the primary target for comparative structural analyses as this is one of the most thoroughly studied family 8 enzymes, and atomic resolution crystal structures are available for this enzyme both in the unliganded and substrate/product bound forms (Beguin *et al.*, 1985; Souchon *et al.*, 1996; Alzari *et al.*, 1996; Guerin *et al.*, 2002). A number of residues are strictly conserved in the family 8 enzymes: Glu 78, Trp 124, Ala 142, Asp 144, Ala 156 and Arg 284 (pXyl numbering). Two further residues are strictly conserved in all enzymes except for the xylanases: Glu 146 (Asp in other family 8 enzymes), and Val 121 or Tyr 121 in xylanases (Leu in other family 8 enzymes).

pXyl has a pI of approximately 9.5 and a pH optimum of 6.5. The high pI is typical for family 11 enzymes, while the relatively high molecular mass is commonly found in family 10 members. The apparent K_m of the enzyme for birchwood xylan is relatively high (~28 mg/ml), compared to typical K_m values between 0.5 and 5 mg/ml (Beg *et al.*, 2001). The enzyme is specific for xylan, is most active on long-chain xylo-oligosaccharides and, in contrast to most

other xylanases studied to date, is not active on aryl-*β*-glycosides of xylobiose or xylotriose. In order to check for parameters associated with adaptation to cold, the xylanase was compared with a family 11 xylanase from *Streptomyces sp*. S38. The cold-adapted enzyme shows a shift in apparent optimal activity of approximately 25 °C towards low temperatures. At 5 °C, the activity is 60 % of the maximum, compared to the mesophilic xylanase for which activity is less than 5 % of the maximum. The mesophilic xylanase has a 10 °C lower melting temperature (53 °C compared to 63 °C), and a reduced half life at 55 °C (1.9 minutes versus 23 minutes).
The crystal structure (at a resolution of 1.3 Å) of the unliganded WT enzymes has shown pXyl to exhibit an $(\alpha/\alpha)_6$-barrel fold (Figure 8.2). An examination of various parameters revealed a reduced number of salt bridges and an increased exposure of hydrophobic residues as potential factors determining the cold-adaptive character of the enzyme. Furthermore, it was suggested that a decreased acidity of the substrate binding cleft and an increased flexibility of aromatic residues lining the subsites may enhance the rate at which substrate is bound. A subsequent biophysical investigation (using differential scanning calorimetry, guanidine unfolding and fluorescence quenching) further demonstrated pXyl to have a reduced stability and increased flexibility in comparison with its thermophilic homologue CelA (Collins *et al*., 2003).

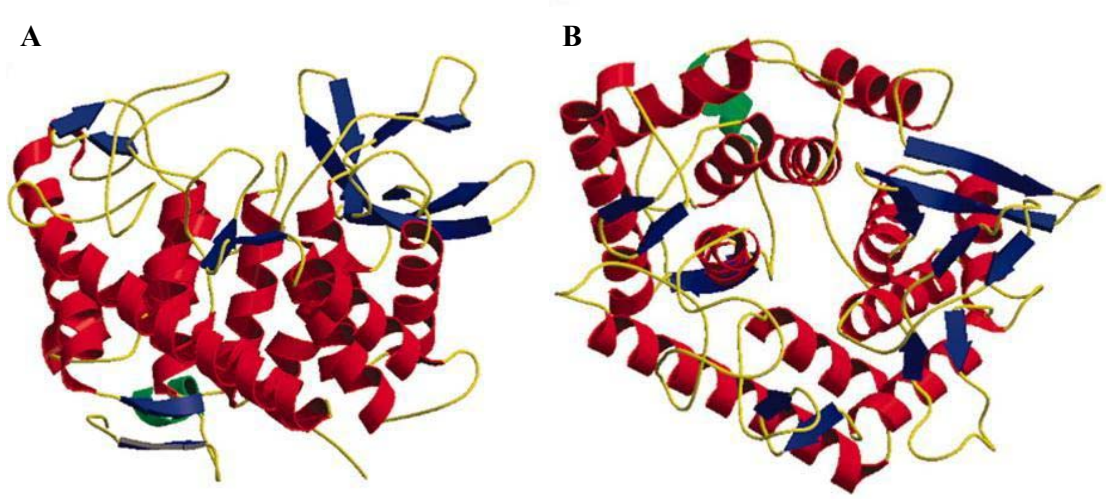


Figure 8.2 Overview of the structures of the cold-adapted xylanase. α-Helices are in red and β-strands in blue (an extra α-helix in the psychrophilic xylanase, compared to CelA, is shown in green). (A) Side view of the $(\alpha/\alpha)_6$ barrel of the native psychrophilic xylanase. (B) Top view of the barrel in A. Adapted from Van Petegem *et al.* (2003).

9 Study of the active site residues of a glycoside hydrolase family 8 xylanase

doi:10.1016/j.jmb.2005.09.064 J. Mol. Biol. (2005) 354, 425–435

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Study of the Active Site Residues of a Glycoside Hydrolase Family 8 Xylanase

T. Collins[1]*†, D. De Vos[2]†, A. Hoyoux[1], S. N. Savvides[2], C. Gerday[1] J. Van Beeumen[2] and G. Feller[1]

[1]*Laboratory of Biochemistry Institute of Chemistry B6 University of Liège, B-4000 Liège, Belgium*

[2]*Laboratorium voor Eiwitbiochemie en Eiwitengineering, Ghent University, Ledeganckstraat 35 B-9000 Gent, Belgium*

**Corresponding author*

Site-directed mutagenesis and a comparative characterisation of the kinetic parameters, pH dependency of activity and thermal stability of mutant and wild-type enzymes have been used in association with crystallographic analysis to delineate the functions of several active site residues in a novel glycoside hydrolase family 8 xylanase. Each of the residues investigated plays an essential role in this enzyme: E78 as the general acid, D281 as the general base and in orientating the nucleophilic water molecule, Y203 in maintaining the position of the nucleophilic water molecule and in structural integrity and D144 in sugar ring distortion and transition state stabilization. Interestingly, although crystal structure analyses and the pH-activity profiles clearly identify the functions of E78 and D281, substitution of these residues with their amide derivatives results in only a 250-fold and 700-fold reduction in their apparent k_{cat} values, respectively. This, in addition to the observation that the proposed general base is not conserved in all glycoside hydrolase family 8 enzymes, indicates that the mechanistic architecture in this family of inverting enzymes is more complex than is conventionally believed and points to a diversity in the identity of the mechanistically important residues as well as in the arrangement of the intricate microenvironment of the active site among members of this family.

Keywords: catalytic site residues; glycoside hydrolase family 8; xylanase

Introduction

Glycoside hydrolases (EC 3.2.1.*x*) catalyse the hydrolysis of the glycosidic bond between two or more carbohydrates or between a carbohydrate and a non-carbohydrate moiety. They are a very diverse group of enzymes and have been classified into 100 different glycoside hydrolase families on the basis of amino acid sequence homology (carbohydrate-active enzyme server (CAZY)‡).[1] Within this classification scheme, members of a particular family have been found to display a similar three-dimensional fold and to catalyse hydrolysis with a similar stereochemical outcome.[2,3]

The cold-adapted xylanase (pXyl) from the Antarctic bacterium *Pseudoalteromonas haloplanktis* TAH3a[4–7] belongs to glycoside hydrolase family 8, a family that displays an $(\alpha/\alpha)_6$ fold (clan GH-M) and that presently groups cellulases, lichenases, chitosanases and a number of other xylanases.[8] This anti-β-inverting enzyme catalyses the random hydrolysis of β-1,4-D-xylosidic linkages in xylan *via* a single displacement mechanism, giving rise to products of α configuration.[4,9] In such inverting glycoside hydrolases, two functional groups are believed to play key roles in the process: typically a carboxyl group (acting as the general acid) and a carboxylate group (acting as the general base), which are usually conserved within a particular family. It is believed that leaving group departure, assisted by partial proton donation from the general acid, is concomitant with nucleophilic attack by water at the anomeric centre, assisted by partial proton abstraction by the general base (Figure 1). This reaction proceeds *via* a high-energy oxocarbenium ion-like transition state that is stabilised by non-covalent interactions with specific, often highly conserved, residues in the enzyme.[10–12]

† T.C. & D.D.V. contributed equally to this work.

Abbreviations used: DSC, differential scanning calorimetry; WT, wild-type.

E-mail address of the corresponding author: tcollins@ulg.ac.be

‡ http://afmb.cnrs-mrs.fr/~cazy/CAZY/

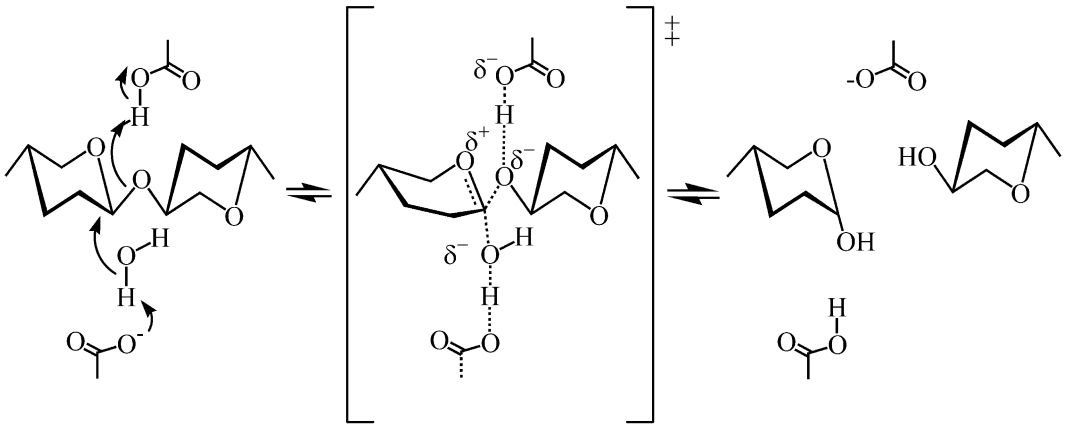


Figure 1. General mechanism for inverting glycosidases. The Figure was prepared with CS Chemdraw Ultra version 6.0.

Previous studies of the catalytic mechanism and catalytic residues of a number of glycoside hydrolase family 8 enzymes have indicated that the general base residue may vary among members of this family. An early study based on site-directed mutagenesis of conserved family 8 residues in endoglucanase K from *Bacillus* sp. KSM-330 identified a glutamic acid residue (E78 in pXyl) and an aspartic acid residue (D144 in pXyl) as the catalytic residues.[13] Subsequent structural analysis of endoglucanase CelA from *Clostridium themocellum*,[14,15] Rex from *Bacillus halodurans* C-125,[16] pXyl,[6] and the chitosanase ChoK from *Bacillus* sp. K17[17] indicated, however, that while E78 (pXyl numbering) is indeed the general acid in all cases, the general base is either D281 (pXyl numbering), Y203 (pXyl numbering) or, in the case of ChoK, a glutamic acid residue located in an additional long inserted loop. Structural analysis indicates that this inserted loop found in ChoK is absent from pXyl, while residues D281 and Y203 are located within hydrogen bonding distance from the putative nucleophilic water molecule[6] and hence either of these two latter residues could function as the general base. On alignment of all 79 family 8 members, however, it is found that, of the proposed general base residues, only D144 is strictly conserved; D281 is substituted by asparagine in 34 of the sequences and by serine in two others, while Y203 is replaced by aspartic acid in 22 sequences.

In the present study, we have used site-directed mutagenesis of the proposed catalytic residues (E78, D281, Y203 and D144) combined with mutant characterisation and crystallographic studies to further unravel the catalytic mechanism and to understand the contributions made by specific residues in determining the activity of pXyl, and hence of family 8 enzymes in general.

Table 1. Kinetic parameters (apparent k_{cat}, apparent K_m and k_{cat}/K_m) and thermal stabilities (T_m) of wild type pXyl and mutant enzymes

Enzyme	Apparent k_{cat} (s^{-1})	Apparent K_m (mg ml^{-1})	k_{cat}/K_m (s^{-1} mg^{-1} ml)	Apparent T_m (°C)
pXyl	1247	28	44.5	53.5
E78Q	5.0	19	0.3	55.6
D281N	1.8	19	0.09	52.9
Y203F	0.32	17	0.02	46.7
Y203A	0.4	15	0.03	46.7
D144N	8.0	36	0.2	52.7
D144A	0.01	20	0.0005	52.7

Average errors in apparent k_{cat} are ±5–20% and in apparent K_m are ±15–20%. Errors of ±0.1 K for T_m have been reported for DSC measurements.[47]

Results

Production and purification

Mutants E78Q, D281N, Y203F, Y203A and D144A were produced and purified using the protocol developed for the wild-type (WT) enzyme, and SDS-PAGE indicated these to be pure and of a molecular mass similar to that of the WT pXyl.

Kinetic parameters

All mutations resulted in large decreases in activity and catalytic efficiency (Table 1) thereby confirming the crucial roles of the mutated amino acids in catalysis. Substitutions D144A (0.0008% residual activity), Y203F (0.03%) and Y203A (0.03%) showed the largest losses, while residual activities of 0.4%, 0.14% and 0.64% were found for the isosteric substitutions E78Q, D281N and D144N, respectively.

Despite the large errors in K_m (Table 1) inherent to the macromolecular nature of the substrate (see Materials and Methods), the results obtained indicate that only the apparent K_m of mutant D144N remains unchanged, or very slightly increased, while all other mutants, in particular Y203F and Y203A, appear to have a slightly reduced K_m.

Thermal stability

Thermal denaturation of all mutants was monitored by differential scanning calorimetry (DSC) and an absence of heat absorption effects on rescanning of denatured samples indicates that thermal unfolding is irreversible in all cases. As can be seen from Table 1, the apparent melting temperature (T_m) values of mutants Y203A and Y203F are reduced by 6.8 deg. C as compared to WT pXyl, indicative of a reduced stability and hence suggesting structural modifications upon substitution of Y203. Very slight reductions in the apparent T_m of both D144 mutants (−0.8 deg. C) and of the D281N mutant (−0.6 deg. C) were observed, whereas the apparent T_m of E78Q is increased (+2.1 deg. C). Such modifications probably reflect the alterations and reorganization of residues and interactions within the active site of these mutants, as discussed below.

pH-dependence of activity

The pH-activity profiles of mutants E78Q and D281N, Y203F and Y203A, and D144N and D144A, are compared with that of WT pXyl in Figure 2(a)–(c), respectively. It can be seen that both the acidic and basic limbs of the pH profiles of Y203F and Y203A closely resemble those of WT pXyl, while slight modifications of the acidic limb, indicative of a slight increase of the pK_a of the general base, occur for both D144N and D144A. In contrast, both E78Q and D281N show major changes in the basic and acidic limbs, respectively, their pH profiles approximating more closely that for a single ionisation.

Structural analysis of mutants

Crystal structures were obtained for E78Q, D281N and D144A at 1.76 Å, 1.3 Å and 3.2 Å resolution, respectively (Table 2). The crystal structure of D144N at 1.5 Å has been described,[6] and is included here for comparative purposes. Efforts to crystallize D144A in the same space group ($P2_12_12_1$) as the other mutants failed, hence alternative crystallisation conditions were used, resulting in data to lower resolution (space group $P3_121$). Despite this, the electron density maps for this mutant were of good quality and all important active site residues were clearly defined (Figure 3). Crystallisation trials were performed under a variety of conditions for Y203F and Y203A but these have not been successful.

The three-dimensional structures of E78Q, D281N, D144N and D144A are very similar to that of WT pXyl, as is evident from overall rmsds for main-chain atoms of 0.14 Å, 0.22 Å, 0.19 Å and 0.41 Å, respectively. Some small conformational differences are indeed observed in the surface loop region comprising residues 69–71 (maximum C^α–C^α distances from the WT of 0.57 Å, 0.64 Å, 1.22 Å and 1.18 Å, respectively). However, these minor modifications do not affect the overall structural integrity of these mutants as compared to the WT enzyme.

Substitution of E78 with Gln results in minor alterations within the active site (Figure 4(a); Table 3), while more important changes are observed for D281N (Figure 4(b); Table 3), D144A (Figure 4(d); Table 3) and, in particular, for D144N (Figure 4(c); Table 3). In WT pXyl, E78 interacts with D144 (2.41 Å), with the proposed nucleophilic water molecule 258 (2.87 Å), as well as with water molecules 122, 330 and 412, and these interactions are retained in mutant E78Q (Figure 4(a)). However, a slight movement of the nucleophilic water towards residue 78 (0.25 Å) and away from residue 281 (0.26 Å), as well as a slight reorientation of this latter around its χ-1 torsion angle are observed in this latter mutant. Furthermore, the fact that residue 78 retains its hydrogen bond (2.50 Å) with D144 and in light of the reorientation of the side-chains of E78 and N144 observed in the D144N substitution (Figure 4(c)),[6] it may be proposed that E78 is the proton donor in the interaction between these two residues. Its side-chain is protonated, in accordance with its proposed function as the general acid in this enzyme.

Mutation of D281 to Asn (Figure 4(b)) appears to result in an increased flexibility in the side-chain of this residue and alternative conformations are indeed required to explain the observed electron density, no other residues were refined with alternative conformations in the active site of this mutant. The mutation results in a movement of the side-chain of residue N281 away from D200, from 2.5 Å to, depending on the alternative conformation, between 2.74 Å and 4.11 Å. This suggests that D281 is non-protonated in the WT enzyme and that it possibly interacts as the H bond acceptor for D200 (donor). A distance of 2.5 Å is not to be expected if D281 and D200 are both negatively charged and the neutral form of D200 is favoured by its location in a highly hydrophobic pocket formed by residues F191, Y203 and F254. Furthermore, if D281 was the H bond donor, one would expect little structural changes upon mutating it to asparagine. In addition, this substitution results in a reorientation of residue 281 such that the hydrogen bonding distance to the nucleophilic water is increased (from 2.8 Å to, depending on the alternative conformation, 3.11 Å and 3.22 Å, respectively) and the distance to R284 is changed (from 3.46 Å to 3.97 Å and 3.29 Å, respectively). Interestingly, while residue 281 is repositioned and displays alternative conformations upon mutation to Asn, the nucleophilic water molecule remains relatively invariant, as only a 0.5 Å displacement from its position in the WT is detected. Indeed this suggests that the positioning of this water molecule is governed by some other residue i.e. Y203 or/and E78.

As described,[6] replacement of D144 with Asn results in a large number of modifications in the active site (Figure 4(c)). In addition to the new interactions of E78 with R284 and Y380, residue R284 now also forms a hydrogen bond with S202 as

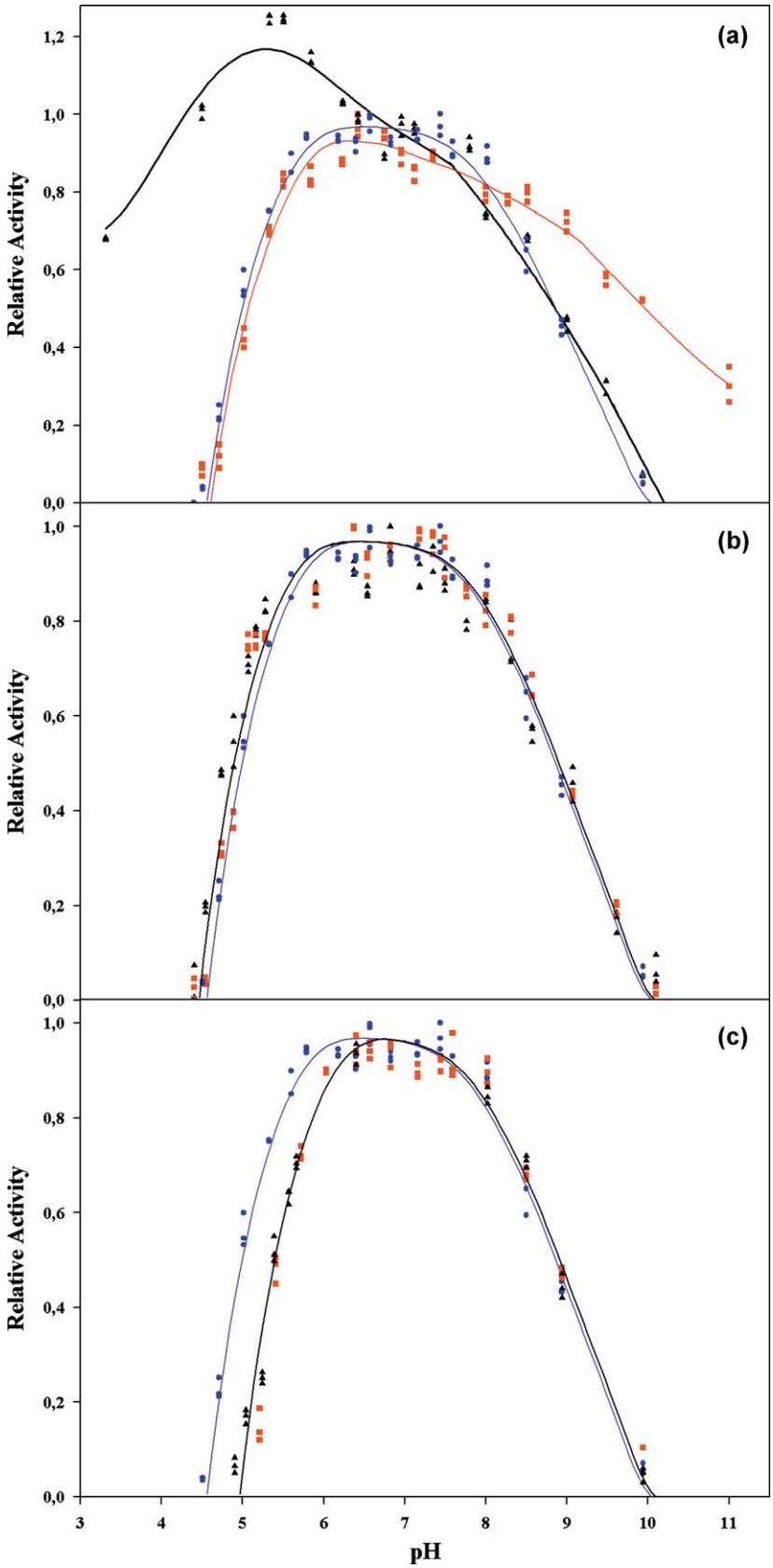


Figure 2. Comparison of pH-activity profiles for pXyl (blue circles, blue lines) with (a) D281N (black triangles, black line) and E78Q (red squares, red line), (b) Y203F (red squares, red line) and Y203A (black triangles, black line) and (c) D144N (red squares, red line) and D144A (black triangles, black line). Activities are expressed relative to the activity at pH 6.5.

the interaction with Y381 is lost. Furthermore, the nucleophilic water molecule shifts by 1.48 Å towards D281 which, along with Y203, also undergoes a slight displacement. In contrast, mutation of D144 to alanine (Figure 4(d)) results in only minor modifications. In particular, E78 occupies a position similar to that of the WT enzyme, with only a slight reorientation of its side-chain away from D144

Table 2. Diffraction data processing and refinement statistics for mutants E78Q, D281N and D144A

	E78Q mutant	D281N mutant	D144A mutant
A. *Diffraction data*			
Resolution (Å)	50–1.76 (1.79–1.76)	50–1.30 (1.34–1.30)	12–3.2 (3.27–3.20)
Space group	$P2_12_12_1$	$P2_12_12_1$	$P3_121$
Unit cell parameters			
a (Å)	51.2	51.6	77.8
b (Å)	91.1	91.6	77.8
c (Å)	97.8	98.0	149.6
No. reflections			
Total	178,916	396,814	33,218
Unique	46,033	104,895	8337
R_{merge}[a] (%)	3.3 (13.7)	7.2 (38.3)	11.9 (43.8)
$I/\sigma(I)$	35.5 (5.7)	18.4 (2.5)	11.1 (3.5)
Completeness (%)	99.4 (95.1)	91.3 (84.3)	93.3 (94.9)
B. *Refinement*			
R_{work}/R_{free} (%)	16.0/17.8	12.3/14.3	21.6/27.6
rms deviations			
Bond lengths (Å)	0.007	0.012	0.008
Bond angles (deg.)	0.96	1.36	1.08
No. atoms			
Protein	3258	3313	3203
Solvent	514	616	57
Average *B*-factor			
Main-chain (Å^2)	11.2	11.3	26.8
Side-chain (Å^2)	12.1	12.9	27.0
Solvent (Å^2)	26.4	29.5	30.7
Ramachandran plot			
Most favoured (%)	91.3	92.1	88.2
Additionally allowed (%)	8.7	7.9	11.8

Values between parentheses reflect the data in the highest resolution shell.

[a] $R_{merge} = \sum_h \sum_i |I(h,i) - \langle I(h)\rangle| / \sum_h \sum_i I(h,i)$, where $I(h,i)$ is the intensity of the ith measurement of reflection h and $\langle I(h)\rangle$ is the average value from multiple measurements.

being observed. This suggests that the repositioning of E78 observed in mutant D144N is due to the loss of its hydrogen bonding interaction with D144 and to repulsion by the introduced asparagine residue. In common with mutant D144N, mutant D144A displays a slight displacement in the side-chains of Y203 and Y381.

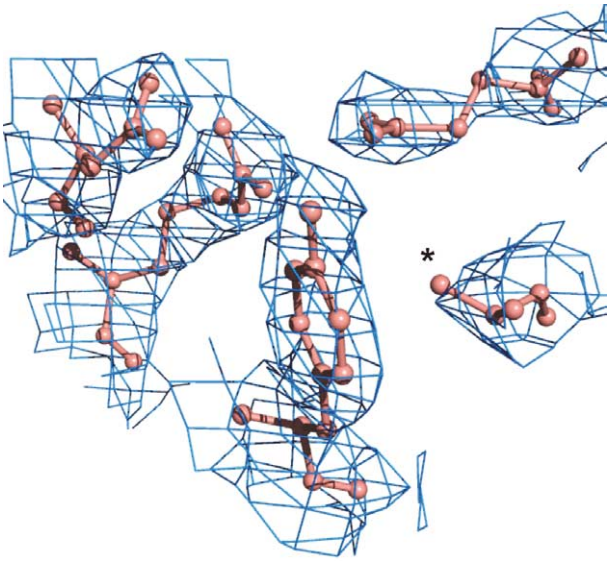


Figure 3. The $2F_o - F_c$ electron density map of the active site of pXyl mutant D144A contoured at 1.0σ. The mutated residue is indicated with an asterisk (*).

Superposing the principal active site residues of pXyl upon those of the CelA–cellopentaose complex (Figure 5) reveals a number of important points: residue 144 in pXyl may form a bifurcated hydrogen bond with the OH-2 and OH-3 groups of a sugar molecule in the −1 subsite; both D281 (2.8 Å) and E78 (2.87 Å) are within hydrogen bonding distance from the proposed nucleophilic water molecule; R76 (conserved strongly in family 8 xylanases) is found to be positioned axially above the sugar ring oxygen atom (O5) and is located near (2.39 Å) the C6 OH of the glucose unit in subsite −1; finally, a conserved Trp (W124) positioned at the −2 subsite stretches into the −1 subsite at the B-face of the sugar unit in this subsite and close to the ring C5 methylene function. Interestingly, the

Table 3. Comparison of mutated residue displacements with coordinate errors for each of the crystallised mutants

Mutant	Mutated residue displacement[a] (Å)	Estimated overall coordinate error[b] (Å)
E78Q	0.20	0.078
D281N	1.88	0.022
D144N	2.05	0.054
D144A	1.05	0.41

[a] Maximum atomic displacement.
[b] Based on maximum likelihood, calculated with REFMAC.[40]

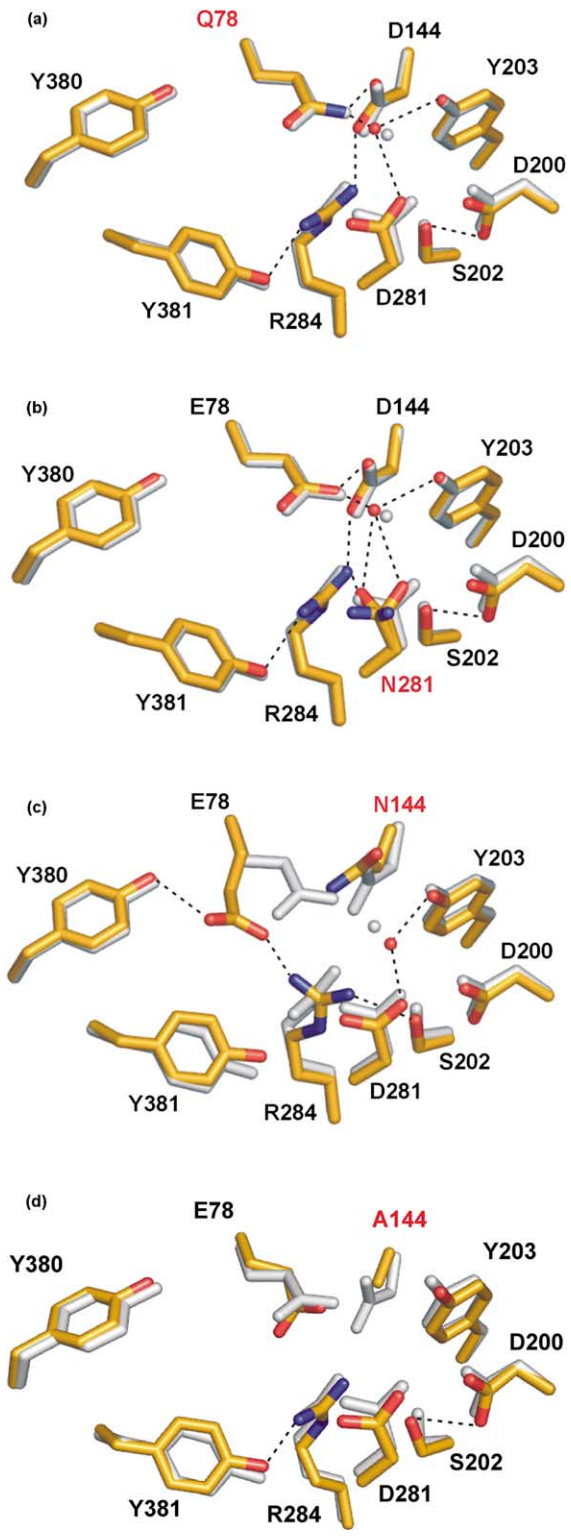


Figure 4. Superposition of the principal active site residues (sticks) and the proposed nucleophilic water (spheres) of wild-type pXyl (grey) with those of mutants (color-coded according to atom type): (a) E78Q; (b) D281N; (c) D144N; and (d) D144A. The hydrogen bonds and nucleophilic water in the mutants are indicated by broken lines and a red sphere, respectively, while the nucleophilic water in the wild-type pXyl is indicated by a grey sphere.

side-chains of E78 and D144 adopt an orientation different from that of their equivalent residues in CelA,[15] ChoK[17] and Rex[16] and indeed, according to the superposition, such a positioning of E78 would lead to a steric clash with the sugar residue at subsite −1. This suggests that some structural modifications, either of the substrate or/and of the protein itself, occur upon substrate binding and thus the structure of a pXyl–substrate complex is required to clarify this point.

Discussion

A combination of kinetic, biophysical and structural studies has allowed us to more clearly define the roles of several active-site residues in the psychrophilic family 8 enzyme pXyl. In agreement with previous studies,[6,15] the function of E78 as the proton donor is confirmed in this work: its side-chain is protonated and the pH-activity profile for the substitution of this residue with Glu is altered drastically in the basic limb as compared to the WT enzyme, this being in accordance with the absence of a catalytic group with high pK_a in this mutant, i.e. lack of the general acid. The decrease in activity observed is similar to that reported for the mutation of the proposed general acid to Gln in the homologous family 8 endoglucanase K from *Bacillus* sp. KSM-330 (460-fold reduction in k_{cat} on CM-cellulose),[13] but is much lower than that observed in ChoK (Glu to Gln mutation) and Rex (Glu to Ala mutation), where approximately 10^4-fold reductions in the specific activity have been reported.[17,18]

The interaction of D281 with the nucleophilic water molecule, the 700-fold reduction in k_{cat} and the substantial alteration of the acidic limb observed in the D281N mutant are consistent with this residue acting as the general base. However, Y203 is also within hydrogen bonding distance from the nucleophilic water molecule and mutation of this residue to Phe or Ala results in a larger drop in activity than for the D281N mutation. On the other hand, the pH-dependences of the Y203 mutants are very similar to that of WT pXyl and only a very slight increase in the pK_a of the general base is observed. This, in addition to the high pK_a of tyrosine, argues against this residue acting as the general base in pXyl. Y203 probably functions with D281 in positioning the nucleophilic water molecule in a catalytically productive orientation and thus its mutation could lead to the observed reduction in activity. Indeed, Y80 in the family 11 xylanase from *Bacillus circulans* has been proposed to carry out a similar function and substitution of this residue by Phe has resulted in a comparable decrease in activity.[19] Furthermore, structural analysis indicates that, in addition to its interaction with the nucleophilic water molecule, this residue is involved in a water-mediated interaction with the main-chain of F191, a residue conserved strongly in family 8 xylanases, and hence disruption of these

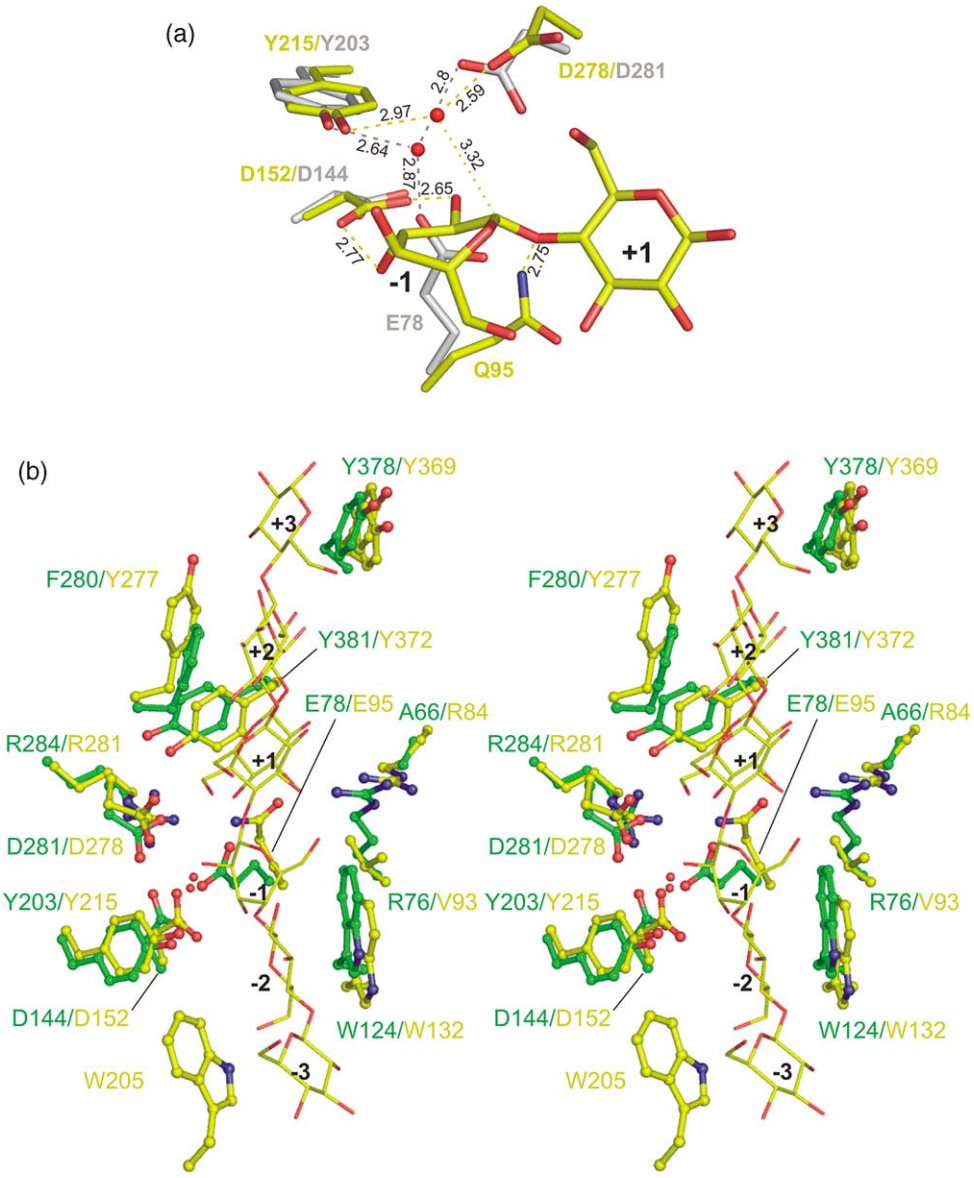


Figure 5. Structural comparison of endoglucanase CelA and endoxylanase pXyl. (a) Detailed view of the catalytic centre of CelA mutant E95Q with glucose residues at subsites −1 and +1 and the most important catalytic residues superimposed on the equivalent residues of WT pXyl. Carbon atoms of CelA and pXyl are coloured yellow and grey, respectively. The proposed nucleophilic water molecules of CelA and pXyl are indicated by spheres. The distance between the nucleophilic water of CelA and the anomeric carbon C1 at subsite −1 is indicated by a dotted line. Hydrogen bonds are indicated by broken lines and distances are in Å. (b) Stereo-drawing of the catalytic site of CelA (E95Q) in complex with cellopentaose (subsites −3 to +2) and cellotriose (subsites +1 to +3), and superposition on the equivalent residues of WT pXyl. Oxygen atoms are coloured red, nitrogen atoms are coloured blue and carbon atoms are coloured yellow (CelA) or green (pXyl).

interactions could lead to the observed destabilization of the protein.

Similar to that observed for the proposed general acid residue, mutation of the proposed general base leads to a smaller decrease in activity than might be expected. However, large variations in the activity decreases are observed upon mutation of the proton acceptor residue in inverting enzymes, pointing to a disparity in the importance of the mechanistic functions in these enzymes, and suggesting that the mechanism of inverting enzymes may be more complex than is currently believed. Mutation of the general base in ChoK (Glu to Gln mutation) and Rex (Asp to Ala mutation) leads to reductions in the specific activity of approximately 10^4-fold and 4500-fold, respectively, while 10^3-fold to 10^4-fold reductions have been observed with a number of other inverting enzymes.[20–22] In contrast, mutation of the proposed general base residue in a family 15 glucoamylase resulted in only a 35-fold to 60-fold loss of activity,[23] while the absence of a general base residue has been reported for the related (clan GH-M) family 48 cellobiohydrolase CelS,[24] as well as for a number of family 6 enzymes.[25,26]

Nevertheless, the small decreases in activity observed in the present study indicate that general acid and general base catalytic assistance may not be the principal factors leading to the hydrolytic rate enhancement in pXyl. In fact, it has been suggested that, although the general acid and base residues are essential for hydrolysis by inverting enzymes, much of the rate improvement is probably due to protein–substrate interactions further removed from the glycosidic scissile bond which stabilise the transition state and improve the "effective leaving group ability".[27] Indeed, an interesting feature of family 8 enzymes is their poor activity on small substrates, as at least five sugar units appear to be required for efficient activity.[4,15,29,30] Here, it has been suggested that part of the energy obtained on substrate binding is utilized for creating a local microenvironment that allows the sugar unit in the −1 subsite to distort into an ALPH-compliant[28] skew-boat conformation in which the scissile glycosidic bond is orientated axially and, as such, is ready for in-line axial attack by the water nucleophile. Furthermore, it has been shown that a high degree of charge accumulates on the leaving phenolate oxygen atom at the transition state in some inverting enzymes, hence indicating a high degree of bond cleavage and little proton donation.[27] Nevertheless, the high pK_a of the leaving group in xylan purports to protonation being absolutely required during departure of the leaving group in this substrate. Another possibility is that family 8 enzymes are highly effective in recruiting other amino acids to act as the catalytic residues in the absence of E78 or D281. Indeed, the proposed general base residue (D281) is not conserved in family 8 enzymes and Adachi *et al*. have even suggested the subdivision of family 8 enzymes into at least three subfamilies, based on the identity of the general base.[17] A further possibility for the small decreases in activity observed is that, as has been seen for the inverting family 15 glucoamylase from *Aspergillus awamori*, the rate-limiting step for xylan degradation by pXyl is something other than hydrolysis,[31,32] but upon mutation of the general acid or base residue the hydrolysis step is slowed sufficiently for it to become rate limiting, leading to the decrease in activity and pH profiles observed. Thus, while our study confirms that E78 and D281 are indeed the general acid and general base residues, further studies are required to investigate the actual importance of these residues in this family of enzymes.

Investigation of the D144 mutants indicates that this residue is extremely important for enzyme activity as a 1.6×10^2-fold and 1.2×10^5-fold reduction is obtained upon mutation to Asn and Ala, respectively. Analysis of the CelA cellopentaose complex,[15] as well as of the superposition of pXyl on CelA-cellopentaose (Figure 5), indicates that this residue (D152 in CelA) is involved in a bidentate hydrogen bonding interaction with the OH-2 and OH-3 hydroxyl groups of the substrate unit in the −1 site. Indeed, interactions with these groups have been shown to be critical in sugar ring distortion and in enhancing the electrophilicity of the anomeric carbon atom.[15,33,34] Therefore, whereas introduction of Asn in place of D144 may maintain the hydrogen bonding interactions, the relatively weaker interaction between the neutral partners (N144 to sugar hydroxyl groups) will result in a reduced activity. Similarly, substitution with Ala will lead to a complete loss of all interactions with the substrate and hence to the large reduction in activity discerned in this study.

From Figure 5 it can be seen that in addition to D144, R76 and W124 may also be proposed as being important for transition state stabilisation in this enzyme. R76 is strongly conserved in family 8 xylanases (replaced by a lysine in xylanase Y from *Bacillus* sp. KK-1) and the free electron pair of its NH1 atom, which is probably in its basic form due to the presence of Asp74 (also highly conserved in family 8 xylanases) adjacent to NH2, may stabilize the carbohydrate ring oxygen atom at the transition state in a manner similar to that of the C6 hydroxyl oxygen atom of the substrate itself in family 8 endoglucanases.[10] The introduction of this strategically located arginine residue in family 8 xylanases may allow for the electrostatic stabilization of the transition state, even in the absence of the C6 hydroxyl oxygen atom in the natural substrate of the xylanases. In relation to W124, this is found to be strictly conserved in family 8 enzymes and is believed to function as a subsite −2 stacking residue;[6] however, as described by Nerinckx *et al*. for CelA,[35] this residue may also function in stabilising the carbohydrate out-of-plane hydrophobic patch created in the transition state.

Materials and Methods

Site-directed mutagenesis

The gene for pXyl was cloned into the NdeI and XhoI sites of the cloning vector pSP73® (Novagen) and the desired mutations were introduced by PCR with Vent DNA polymerase. The D281N and D281A mutations were introduced using the sense primers 5′-GAATTTAAT GCATGGCGCGTAATTATG-3′ or 5′-GAATTTGCTGC ATGGCGCG-3′, containing the Asp to Asn or Ala mutation (underlined), respectively, and the antisense primer 5′-ATACCATTGACCTGGGTTTGCTTGG-3′. The Y203F and Y203A mutations were introduced using the sense primer 5′-CCTAACAGACCCTTCTTTCCATAT ACCTGCG-3′ or 5′-CCTAACAGACCCTTCTGCCCATAT ACCTG-3′, containing the Tyr to Phe or Ala mutation (underlined), respectively, and the antisense primer 5′-TTATCAATGTAAGGTGAAAAGCGGATTATTTGG-3′. The E78Q and E78A mutations were introduced using the sense primer 5′-TATAAGAACACAAGGGCAATCGTG GGGA-3′ or 5′-TATAAGAACAGCAGGGCAATCGTGG GG-3′, containing the Glu to Gln or Ala mutation (underlined), respectively, and the antisense primer 5′-TCGTCGCCTTCGTCTGGGTTAATTGC-3′. The D144A mutation was introduced using the sense primer 5′-GCCCCGCTCCCGCTGGCGAA G-3′, containing the Asp

to Ala mutation (underlined) and the antisense primer 5′-CCTCATCCACTTTATAAACAAAGCCGTTTTGA-3′.

The PCR products were purified and circularised, transformed to *Escherichia coli* DINO RR1® and double-strand sequenced with an ALF DNA sequencer (Amersham Biosciences). The mutated genes were then excised from the cloning vector by restriction digestion with NdeI and XhoI, and ligated into the expression vector pET22b(+) (Novagen). The resulting recombinant plasmid was transformed to the expression host *E. coli* BL21 (DE3) (Stratagene).

Production and purification

Production and purification of all mutants was carried out as described for the WT recombinant enzyme.[4] Briefly, cultures were produced in Terrific broth (12 g/l of Bacto tryptone (Difco), 24 g/l of yeast extract (Difco), 4 ml/l of glycerol, 12.54 g/l of K_2HPO_4, 2.31 g/l of KH_2PO_4, 200 μg/ml of ampicillin) at 18 °C, enzyme expression was induced with 1 mM isopropyl-1-thio-β,D-galacto-pyranoside and purification was carried out by a combination of anion-exchange chromatography (QFF-Sepharose, Amersham Biosciences), cation-exchange chromatography (SFF-Sepharose, Amersham Biosciences) and gel-filtration chromatography (Sephacryl S-100, Amersham Biosciences).

Characterisation of mutants

Protein purity and molecular size was determined by SDS-PAGE run essentially as described by the instrument supplier (Hoefer-Scientific).

Xylanase activity was measured by a modification of the dinitrosalicylic acid method[36] at pH 6.5 in 100 mM citrate-acetate buffer with 3% (w/v) soluble birchwood xylan.[37]

Kinetic parameters were determined at 25 °C by non-linear regression, using the Michaelis–Menten equation, of initial rates between 0 mg/ml and 30 mg/ml of soluble birchwood xylan. The inability to measure activity at concentrations higher due to substrate interference and high viscosity meant that substrate saturation was not achieved and, thus, extrapolation was used in the determination of the kinetic values, giving rise to apparent values.

The pH-dependence of activity was determined with a 2 h xylanase activity assay at 25 °C between pH 4.5 and pH 12, with 3% soluble birchwood xylan as substrate, using the following buffer mix: 40 mM sodium acetate, 20 mM Mes, 20 mM Mops, 20 mM Taps, 20 mM Ches and 20 mM Caps.

The T_m for irreversible thermal unfolding was determined by DSC with a MicroCal MCS-DSC apparatus (Microcal) using the MCS Observer software package for data acquisition, analysis and deconvolution. Analysis was carried out with 2 mg/ml of protein in 20 mM Mops (pH 7.5), 100 mM NaCl, 500 mM 3-(1-pyridinio)-1-propanesulfonate, at a scan rate of 1 deg. C/min and under 1 atm (=101,325 Pa) nitrogen pressure.

Structure determination and refinement

Crystals of the mutant enzymes E78Q and D281N were grown as described for the WT enzyme[5] except that sodium phosphate buffer at pH7 was used. Crystals of mutant D144A were obtained by using the hanging-drop, vapour-diffusion technique at 277 K. A portion (1 μl) of protein solution (10 mg/ml in 20 mM Mops (pH 7.5), 50 mM NaCl, 2% (w/v) trehalose) was mixed with 1 μl of reservoir solution (0.1 M sodium cacodylate (pH 5.6), 27% (w/v) PEG8000, 0.2 M ammonium acetate). Diffraction data were collected from single E78Q crystals on a DIP2030 imaging plate detector with Cu Kα radiation from a Nonius FR591 rotating anode generator with Montel multilayer optics. Diffraction data were collected from a single D281N crystal on a MarResearch Imaging Plate scanner using 0.81 Å synchrotron radiation at the X11 beamline of EMBL/DESY in Hamburg. Diffraction data were collected from a single D144A crystal on a MAR CCD (MarResearch) detector system using 0.80 Å radiation at the X13 beamline of EMBL/DESY in Hamburg. All data were collected at 100 K using paratone as the cryo-protectant. Intensity data were integrated, scaled and reduced to structure factor amplitudes with the HKL suite of programs[38] and data collection statistics are shown in Table 2A. Molecular replacement was carried out with AmoRe[39] using the native xylanase structure (PDB code 1H13) as a search model. Prior to refinement, 5% of the data were randomly flagged for cross-validation (R_{free}) and restrained refinement was performed using the program REFMAC,[40] by monitoring the convergence of R_{work} and R_{free} (Table 2). For the D281N mutant, the final cycles of refinement were carried out assuming anisotropic temperature factors. Fourier maps with coefficients of $2mF_{obs}-DF_{calc}$ and $mF_{obs}-DF_{calc}$ were calculated, where m is the figure of merit and D is the error distribution derived from the σ_A function.[41] Inspection of electron density maps and model refinement were carried out with TURBO-FRODO[42] and side-chains with missing electron density were not modelled. Simulated annealing omit maps were calculated using CNS to confirm the positioning of the major active site residues.[43] The stereochemistry of the model was analysed using PROCHECK,[44] while hydrogen bonds were calculated with the program HBPLUS.[45] Structural superpositions were performed using LSQKAB[46] and Figures were drawn with the program Pymol†.

Protein Data Bank accession codes

Atomic coordinates and related structure factors have been deposited in the RCSB Protein Data Bank, Rutgers University, New Brunswick, NJ‡ with identification codes 1XW2 for E78Q, 1XWT for D281N and 2A8Z for D144A.

Acknowledgements

The authors thank N. Gerardin and R. Marchand for their skilful technical assistance and the French Institute for Polar Research for generously accommodating our research fellows at the French Antarctic station in J.S. Dumont d'Urville, Terre Adelie. This work was supported by the Region Wallonne (contract BIOVAL 981/3860 and FIRST EUROPE R0202/215266), the European Union (network contract CT970131), the Fonds National de la Recherche Scientifique (contracts 2.4515.00 and 2.4536.04) and the Fonds voor Wetenschappelijk

† http://www.pymol.org
‡ http://www.rcsb.org/

Onderzoek-Vlaanderen (contract G.0330.03). We are grateful for access to beamline X11 at the EMBL/DESY, Hamburg Outstation and to Wim Nerincx for helpful discussions. D.D.V. is a Research Fellow of the Fund for Scientific Research-Flanders.

References

1. Coutinho, P. M. & Henrissat, B. (1999). Carbohydrate-active enzymes: an integrated database approach. In, (Gilbert H. J., Henrissat, G. D. & Svensson, B., eds), pp. 3–12, The Royal Society of Chemistry, Cambridge, UK.
2. Gebler, J., Gilkes, N. R., Claeyssens, M., Wilson, D. B., Beguin, P., Wakarchuk, W. W. *et al*. (1992). Stereoselective hydrolysis catalyzed by related beta-1,4-glucanases and beta-1,4-xylanases. *J. Biol. Chem.* **267**, 12559–12561.
3. Henrissat, B. & Coutinho, P. M. (2001). Classification of glycoside hydrolases and glycosyltransferases from hyperthermophiles. *Methods Enzymol.* **330**, 183–201.
4. Collins, T., Meuwis, M. A., Stals, I., Claeyssens, M., Feller, G. & Gerday, C. (2002). A novel family 8 Xylanase: functional and physico-chemical characterization. *J. Biol. Chem.* **277**, 35133–35139.
5. Van Petegem, F., Collins, T., Meuwis, M. A., Gerday, C., Feller, G. & Van Beeumen, J. (2002). Crystallization and preliminary X-ray analysis of a xylanase from the psychrophile *Pseudoalteromonas haloplanktis*. *Acta Crystallog. sect. D*, **58**, 1494–1496.
6. Van Petegem, F., Collins, T., Meuwis, M. A., Gerday, C., Feller, G. & Van Beeumen, J. (2003). The structure of a cold-adapted family 8 xylanase at 1.3 Å resolution. Structural adaptations to cold and investigation of the active site. *J. Biol. Chem.* **278**, 7531–7539.
7. Collins, T., Meuwis, M. A., Gerday, C. & Feller, G. (2003). Activity, stability and flexibility in glycosidases adapted to extreme thermal environments. *J. Mol. Biol.* **328**, 419–428.
8. Collins, T., Gerday, C. & Feller, G. (2005). Xylanases, xylanase families and extremophilic xylanases. *FEMS Microbiol. Rev.* **29**, 3–23.
9. Fierobe, H. P., Bagnara-Tardif, C., Gaudin, C., Guerlesquin, F., Sauve, P., Belaich, A. & Belaich, J. P. (1993). Purification and characterization of endoglucanase C from *Clostridium cellulolyticum*. Catalytic comparison with endoglucanase A. Eur. J. Biochem. **217**, 557–565.
10. Nerinckx, W., Desmet, T., Piens, K. & Claeyssens, M. (2005). An elaboration on the syn-anti proton donor concept of glycoside hydrolases: electrostatic stabilisation of the transition state as a general strategy. *FEBS Letters*, **579**, 302–312.
11. Ly, H. D. & Withers, S. G. (1999). Mutagenesis of glycosidases. *Annu. Rev. Biochem.* **68**, 487–522.
12. Davies, G. J., Ducros, V. M., Varrot, A. & Zechel, D. L. (2003). Mapping the conformational itinerary of beta-glycosidases by X-ray crystallography. *Biochem. Soc. Trans.* **31**, 523–527.
13. Ozaki, K., Sumitomo, N., Hayashi, Y., Kawai, S. & Ito, S. (1994). Site-directed mutagenesis of the putative active site of endoglucanase K from *Bacillus* sp. KSM-330. *Biochim. Biophys. Acta*, **1207**, 159–164.
14. Alzari, P. M., Souchon, H. & Dominguez, R. (1996). The crystal structure of endoglucanase CelA, a family 8 glycosyl hydrolase from *Clostridium thermocellum*. *Structure*, **4**, 265–275.
15. Guerin, D. M., Lascombe, M. B., Costabel, M., Souchon, H., Lamzin, V., Beguin, P. & Alzari, P. M. (2002). Atomic (0.94 Å) resolution structure of an inverting glycosidase in complex with substrate. *J. Mol. Biol.* **316**, 1061–1069.
16. Fushinobu, S., Hidaka, M., Honda, Y., Wakagi, T., Shoun, H. & Kitaoka, M. (2005). Structural basis for the specificity of the reducing end xylose-releasing exo-oligoxylanase from *Bacillus halodurans* C-125. *J. Biol. Chem.* **280**, 17180–17186.
17. Adachi, W., Sakihama, Y., Shimizu, S., Sunami, T., Fukazawa, T., Suzuki, M. *et al*. (2004). Crystal structure of family GH-8 chitosanase with subclass II specificity from *Bacillus* sp. K17. *J. Mol. Biol.* **343**, 785–795.
18. Honda, Y. & Kitaoka, M. (2004). A family 8 glycoside hydrolase from *Bacillus halodurans* C-125 (BH2105) is a reducing end xylose-releasing exo-oligoxylanase. *J. Biol. Chem.* **279**, 55097–55103.
19. Wakarchuk, W. W., Campbell, R. L., Sung, W. L., Davoodi, J. & Yaguchi, M. (1994). Mutational and crystallographic analyses of the active site residues of the *Bacillus circulans* xylanase. *Protein Sci.* **3**, 467–475.
20. Ohnishi, H., Matsumoto, H., Sakai, H. & Ohta, T. (1994). Functional roles of Trp337 and Glu632 in *Clostridium glucoamylase*, as determined by chemical modification, mutagenesis, and the stopped-flow method. *J. Biol. Chem.* **269**, 3503–3510.
21. Chauvaux, S., Beguin, P. & Aubert, J. P. (1992). Site-directed mutagenesis of essential carboxylic residues in *Clostridium thermocellum* endoglucanase CelD. *J. Biol. Chem.* **267**, 4472–4478.
22. Damude, H. G., Withers, S. G., Kilburn, D. G., Miller, R. C., Jr & Warren, R. A. (1995). Site-directed mutation of the putative catalytic residues of endoglucanase CenA from *Cellulomonas fimi*. *Biochemistry*, **34**, 2220–2224.
23. Frandsen, T. P., Dupont, C., Lehmbeck, J., Stoffer, B., Sierks, M. R., Honzatko, R. B. & Svensson, B. (1994). Site-directed mutagenesis of the catalytic base glutamic acid 400 in glucoamylase from *Aspergillus niger* and of tyrosine 48 and glutamine 401, both hydrogen-bonded to the gamma-carboxylate group of glutamic acid 400. *Biochemistry*, **33**, 13808–13816.
24. Guimaraes, B. G., Souchon, H., Lytle, B. L., David Wu, J. H. & Alzari, P. M. (2002). The crystal structure and catalytic mechanism of cellobiohydrolase CelS, the major enzymatic component of the *Clostridium thermocellum* cellulosome. *J. Mol. Biol.* **320**, 587–596.
25. Koivula, A., Ruohonen, L., Wohlfahrt, G., Reinikainen, T., Teeri, T. T., Piens, K. *et al*. (2002). The active site of cellobiohydrolase Cel6A from *Trichoderma reesei*: the roles of aspartic acids D221 and D175. *J. Am. Chem. Soc.* **124**, 10015–10024.
26. Wolfgang, D. E. & Wilson, D. B. (1999). Mechanistic studies of active site mutants of *Thermomonospora fusca* endocellulase E2. *Biochemistry*, **38**, 9746–9751.
27. Damude, H. G., Ferro, V., Withers, S. G. & Warren, R. A. (1996). Substrate specificity of endoglucanase A from *Cellulomonas fimi*: fundamental differences between endoglucanases and exoglucanases from family 6. *Biochem. J.* **315**, 467–472.
28. Deslongchamps, P. (1993). Intramolecular strategies and stereoelectronic effects. Glycosides hydrolysis revisited. *Pure Appl. Chem.* **65**, 1161–1178.

29. Brennan, Y., Callen, W. N., Christoffersen, L., Dupree, P., Goubet, F., Healey, S. *et al*. (2004). Unusual microbial xylanases from insect guts. *Appl. Environ. Microbiol.* **70**, 3609–3617.
30. Petre, J., Longin, R. & Millet, J. (1981). Purification and properties of an endo-beta-1,4-glucanase from *Clostridium thermocellum*. *Biochimie*, **63**, 629–639.
31. Sierks, M. R., Sico, C. & Zaw, M. (1997). Solvent and viscosity effects on the rate-limiting product release step of glucoamylase during maltose hydrolysis. *Biotechnol. Prog.* **13**, 601–608.
32. Natarajan, S. K. & Sierks, M. R. (1996). Identification of enzyme-substrate and enzyme–product complexes in the catalytic mechanism of glucoamylase from *Aspergillus awamori*. *Biochemistry*, **35**, 15269–15279.
33. Uitdehaag, J. C., Mosi, R., Kalk, K. H., van der Veen, B. A., Dijkhuizen, L., Withers, S. G. & Dijkstra, B. W. (1999). X-ray structures along the reaction pathway of cyclodextrin glycosyltransferase elucidate catalysis in the alpha-amylase family. *Nature Struct. Biol.* **6**, 432–436.
34. McCarter, J. D., Adam, M. J. & Withers, S. G. (1992). Binding energy and catalysis. Fluorinated and deoxygenated glycosides as mechanistic probes of *Escherichia coli* (lacZ) beta-galactosidase. *Biochem. J.* **286**, 721–727.
35. Nerinckx, W., Desmet, T. & Claeyssens, M. (2003). A hydrophobic platform as a mechanistically relevant transition state stabilising factor appears to be present in the active centre of all glycoside hydrolases. *FEBS Letters*, **538**, 1–7.
36. Petrescu, I., Lamotte-Brasseur, J., Chessa, J. P., Ntarima, P., Claeyssens, M., Devreese, B. *et al*. (2000). Xylanase from the psychrophilic yeast *Cryptococcus adeliae*. *Extremophiles*, **4**, 137–144.
37. Matte, A. & Forsberg, C. W. (1992). Purification, characterization, and mode of action of endoxylanases 1 and 2 from *Fibrobacter succinogenes* S85. *Appl. Environ. Microbiol.* **58**, 157–168.
38. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
39. Navaza, J. (1994). AmoRe:an automated package for molecular replacement. *Acta Crystallog. sect. D*, **50**, 157–163.
40. Murshudov, G. N. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D*, **53**, 240–255.
41. Read, R. J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallog. sect. A*, **42**, 144–149.
42. Roussel, A. & Cambillau, C. (1992). *TURBO-FRODO Biographics*, AFMB, Marseille, France.
43. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al*. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
44. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemistry quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
45. McDonald, I. K. & Thornton, J. M. (1994). Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238**, 777–793.
46. Collaborative Computing Project Number 4. (1994). CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
47. Feller, G., d'Amico, D. & Gerday, C. (1999). Thermodynamic stability of a cold-active α-amylase from the Antarctic bacterium *Alteromonas haloplanctis*. *Biochemistry*, **38**, 4613–4619.

Edited by M. Guss

(Received 19 July 2005; received in revised form 19 September 2005; accepted 19 September 2005)
Available online 10 October 2005

10 Oligosaccharide binding in family 8 glycosidases: crystal structures of active site mutants of the *β*-1,4-xylanase pXyl from Pseudoaltermonas haloplanktis in complex with substrate and product

Oligosaccharide binding in family 8 glycosidases: crystal structures of active site mutants of the *β*-1,4-xylanase pXyl from *Pseudoaltermonas haloplanktis* in complex with substrate and product

De Vos, D.[‡,#], Collins, T.[§,#], Nerinckx, W.[‖], Savvides, S. N.[‡], Claeyssens, M.[‖], Gerday, C.[§], Feller, G.[§], and Van Beeumen, J.[‡,*]

† This work was supported by the Fonds voor Wetenschappelijk Onderzoek-Vlaanderen (FWO contract G.0330.03) and the 'Region Wallonne' (FIRST EUROPE R0202/215266). D.D.V. is a Research Fellow of the FWO.

‡ Laboratorium voor Eiwitbiochemie en Eiwitengineering, Ghent University, K. L. Ledeganckstraat 35, B-9000 Gent, Belgium.
§ Laboratory of Biochemistry, Institute of Chemistry B6, University of Liège, B-4000 Liège, Belgium.
‖ Laboratory for Glycobiology, Department of Biochemistry, Physiology and Microbiology, Ghent University, K. L. Ledeganckstraat 35, B-9000 Gent, Belgium.

* To whom correspondence should be addressed: Jozef Van Beeumen, Laboratorium voor Eiwitbiochemie en Eiwitengineering, Ghent University, K. L. Ledeganckstraat 35, B-9000 Gent, Belgium. Tel. +32 (0)9 264 51 09; Fax. +32 9 264 53 37; Email: jozef.vanbeeumen@ugent.be

# Both authors contributed equally to this work.

Running Title: Substrate and product complexes of pXyl

Atomic coordinates and related structure factors have been deposited in the RCSB Protein Data Bank, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/) with identification codes 1XWQ for E78Q-X3 and 2B4F for D144A-X5.

Abbreviations: MOPS, 3-(N-morpholino)propanesulphonic acid; rmsd: root-mean-square-deviation.

ABSTRACT: The structures of inactive mutants (D144A and E78Q) of the glycoside hydrolase family 8 (GH-8) endo-*β*-1,4-D-xylanase (pXyl) from the Antarctic bacterium *Pseudoalteromonas haloplanktis* in complex with its substrate xylopentaose (at 1.95 Å resolution) and product xylotriose (at 1.9 Å resolution) have been determined by X-ray crystallography. A detailed comparative analysis of these with the apo-enzyme and with other GH-8 structures indicates an induced fit mechanism upon ligand binding, whereby a number of conformational changes and in particular a repositioning of the proton donor into a more catalytically competent position occur. This has also allowed for the description of protein-ligand interactions in this enzyme and for the demarcation of subsites –3 to +3. An in-depth analysis of each of these subsites gives an insight into the structure – function relationship of this enzyme and the basis of xylose/glucose discrimination in family 8 glycoside hydrolases. Furthermore, the structure of the –1/+1 subsite spanning complex reveals that the substrate is distorted from its ground state conformation. Indeed, structural analysis and *in silico* docking studies indicate that substrate hydrolysis in GH-8 members is preceded by a conformational change, away from the substrate ground state chair conformation, to a pre-transition state local minimum $^{2}S_{O}$ conformation.

Xylans are the most abundant hemicelluloses in plants, accounting for as much as 35 % of the dry weight of higher plants. They are heteropolysaccharides, being typically composed of a backbone of ~150-200 *β*-1,4-linked xylopyranose units substituted, in particular at the OH2 and OH3 positions, to varying degrees with diverse side-groups. Depending on the botanical origin, substituents including arabinose, D-glucuronic acid, 4-*O*-methyl-D-glucuronic acid and acetic acid may be found covalently bound to the xylopyranose units. Indeed, this heterogeneity results in the need for several hemicellulolytic enzymes for complete degradation of xylan. The enzymatic hydrolysis of xylan has raised significant interest because of its biotechnological applications, for example, in biobleaching and in the food and feed industries (*1–4*).

Endo-xylanases (EC 3.2.1.8) are glycoside hydrolases that catalyze the hydrolysis of the internal 1,4-*β*-D-bonds of the xylan backbone. The majority of xylanases cluster into families 10 and 11 in the sequence-based classification system of glycoside hydrolases (afmb.cnrs-mrs.fr/CAZY) (*5*). Within this classification scheme, members of a particular family have been found to display a similar three dimensional fold and to catalyse hydrolysis with a similar stereochemical outcome (*6, 7*). The cold-adapted xylanase (pXyl) from the Antarctic bacterium *Pseudoalteromonas haloplanktis* TAH3a (*8–11*) belongs to GH-8, a family which exhibits an $(\alpha/\alpha)_6$ fold (clan GH-M) and which presently groups cellulases, lichenases, chitosanases and a number of other xylanases (*12*). Family 8 members investigated, including pXyl, have been shown to be inverting enzymes. Such a reaction is believed to involve a direct displacement of the leaving group by water, via an oxocarbenium-ion-like transition state, which is aided by both general acid and general base catalysis (*13, 14*). Typically, one carboxylate group provides for a general acid-catalyzed leaving group departure and a second functions as a general base, activating a nucleophilic water molecule to attack the anomeric carbon. This leads to cleavage of the glycosidic bond and to an inversion of the configuration at the anomeric carbon. In a recent study it has been shown that E78 and D281 in pXyl correspond to the catalytic proton donor and acceptor respectively (*15*). A third catalytically crucial residue is D144, which has been implicated in stabilizing the transition state by a bidentate interaction with the OH2 and OH3 groups of the sugar at subsite –1 (the subsites that bind the glycon and aglycon regions of the substrate are prefixed by – and +, respectively, and their number is related to the site of bond cleavage) (*16*).

Known crystal structures of GH-8 members include pXyl, the *β*–1,4-endoglucanase CelA, the reducing end exo-oligoxylanase Rex and the chitosanase ChoK (*17–20*). CelA is structurally the best characterized GH-8 member, crystal structures being available of its unliganded, substrate- and product-bound forms. Furthermore, crystal structures are available for the unliganded and product-bound forms of Rex. Here we report the high resolution crystal structures of the substrate and product complexes of pXyl which, in combination with *in silico* docking studies, provide insight into the substrate specificity and catalytic mechanism of this enzyme, as well as into the catalytic mechanism of family 8 glycoside hydrolases in general. Furthermore, we clarify aberrant observations previously reported on the structure of the unliganded pXyl, i.e. the unfavourably positioned catalytic proton donor E78 and the presence of an extra +4 subsite (*10*).

Materials and Methods

Crystallization and data collection

To obtain the structure of pXyl in complex with substrate and product, we have made use of active site mutants D144A and E78Q obtained by site-directed mutagenesis as previously described (*15*). Purified protein was prepared as described by Collins *et al.* (8). Crystallization was performed using the hanging-drop vapour-diffusion technique, mixing 1 µl of protein solution (10 mg/ml, 20 mM MOPS pH 7.5, 50 mM NaCl, 2 % trehalose) with 1 µl of reservoir solution as described by Van Petegem *et al.* (*9*) but using sodium phosphate buffer at pH 7.0

instead of sodium acetate buffer at pH 5.0. Since direct soaking of the resulting crystals with xylo-oligosaccharides of varying lengths was unsuccessful, the ligand was included into the crystallization droplet (at 2 mM). Co-crystallization of E78Q with xylotetraose took 2-3 months and resulted in a complex with xylotriose (X3). To obtain the structure of pXyl in complex with substrate, the more highly inactivated mutant D144A (*15*) was used. Here, crystallization was carried out with 1 µl of protein solution (identical to above, but with 10 mM xylohexaose included) mixed with 1 µl of well solution containing 21 % PEG8000, 0.1 M sodium cacodylate and 0.2 M ammonium acetate, pH 5.6. The resulting crystals grew in the same space group as for the E78Q-X3 complex but were smaller and appeared within 2-3 days. Crystals grown in the presence of xylohexaose (X6) led to a fully occupied binding cleft. However, the electron density for the most distal D-xylosyl moiety at the reducing end was not defined clearly enough to allow modelling of the complete xylohexaose molecule.

Diffraction data were collected from a single E78Q-X3 crystal on a DIP2030 imaging plate detector using CuK_α radiation from a Nonius FR591 rotating anode generator with Montel multilayer optics. Diffraction data were collected from a single D144A-X6 crystal on a MARCCD detector (MarResearch) using 0.94 Å synchrotron radiation at the BW7A beamline at the EMBL/DESY in Hamburg. All data were collected at 100 K using 25 % glycerol as the cryo-protectant. Intensity data were integrated, scaled and reduced to structure factor amplitudes with the *HKL* suite of programs (*21*). Data collection statistics are shown in Table 1.

Structure refinement and analysis

The structures were isomorphous with the native structure described previously (pdb code 1H13) and were refined, starting from this structure, using the program REFMAC (*22*) of the *CCP*4 suite (*23*) by monitoring the convergence of R_{work} and R_{free} (Table 1). Inspection of electron density maps and model refinement were carried out with TURBO-FRODO (*24*) and side chains with missing electron density were not modelled. After refinement of the protein, ligand and water molecules were added to the model based on residual difference density in $F_o - F_c$ and $2F_o - F_c$ electron density maps, followed by further refinement and editing of the water molecules. The stereochemistry of the protein model was analyzed using PROCHECK (*25*), while the program PLATON (*26*) was used for evaluation of the carbohydrate conformations. Hydrogen bonds were calculated with the programs HBPLUS (*27*) and LIGPLOT (*28*). Structural superpositions were performed using LSQKAB (*23*) and figures were drawn with the program PYMOL (*29*).

Atomic coordinates and related structure factors have been deposited in the RCSB Protein Data Bank, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/) with identification codes 1XWQ for E78Q-X3 and 2B4F for D144A-X5.

Docking studies

The program AUTODOCK version 3.0.5 (*30*) was used for the automated docking of ligands to the active site of the E78Q-X3 structure (pdb entry 1XWQ). Xylotriose ligands, with the central D-xylosyl moiety in a $^1S_3$, $^1S_5$, ^2S_O and $^4C_1$ conformation, were drawn using the program HYPERCHEM® 7.5 (Hypercube Inc., Gainesville, USA). The same program was subsequently used for energy minimisation with the MM+ molecular mechanics force field (*31, 32*). The protein was prepared for docking by removing ligand and water molecules as well as side chains corresponding to alternative conformations (none being implicated in the binding of sugars at subsites –2, –1 or +1). Following this, missing side chains (none being implicated in sugar binding) were added with the program DEEPVIEW/SWISS-PDBVIEWER 3.7 (*33*) and polar hydrogens were added with the AUTODOCK TOOLS graphical user interface (ADT). ADT was then used to add atomic partial charges to the ligand and protein using the respective Gasteiger and Kollman methods (*34, 35*) and atomic solvation parameters

were added to the protein file (*36*). AUTODOCK uses a grid-based method for energy-evaluations in which the grid points contain precalculated affinities for the different atom types of the ligand. The number of grid points was set at 40×44×44, with a point spacing of 0.297 Å, yielding a grid box that spans in excess of the –2 to +1 subsites. For each ligand, 50 docking runs were performed using a Lamarckian Genetic Algorithm (using the pseudo-Solis and Wets local search method) (*37*) with a run-termination of 2,500 generations (at a maximum of 25×10^6 energy evaluations), thereby allowing free rotation around the 12 exocyclic bonds between the heavy atoms (C/O) of the ligands. After docking, all structures generated for a single compound were assigned to clusters based on a tolerance of 2 Å for all-atom rmsds from the lowest-energy structure.

Results and Discussion

Crystallographic data and model quality

Inactive pXyl-D144A was crystallized in complex with the substrate xylohexaose (X6) and the structure was determined to 1.95 Å resolution using synchrotron radiation (Table 1). The final model includes 404 amino acid residues, 448 solvent sites, five enzyme-bound D-xyloside residues at the active site cleft and five enzyme-bound D-xyloside residues at a remote binding site. pXyl-mutant E78Q was crystallized in complex with the product xylotriose and led to the structure of the complex at 1.9 Å resolution based on data from an in-house source (Table 1). The final model includes 404 amino acid residues, 431 solvent sites and three enzyme-bound D-xyloside residues at the active site cleft. The stereochemical quality of both models is comparable to that of other crystal structures refined at a similar resolution. All non-glycine residues display main-chain dihedral angles that are localised within allowed or additional allowed regions of the Ramachandran plot, as defined by the program PROCHECK (*25*). Representative electron density for the ligands is depicted in Figure 1A**.**

Structure of the pXyl substrate and product complexes

The overall structure of pXyl in the substrate and product complexes is very similar to that of the unliganded wild type enzyme (rmsds for C_α atoms of resp. 0.27 Å and 0.2 Å). Superimpositions reveal, however, that small concerted backbone movements (maximal C_α distances of 1.0 Å) leading to an opening of the binding cleft take place upon binding of substrate or product. The regions showing the largest shifts are the elongated loops lining the binding cleft, as well as a segment of the $(\alpha/\alpha)_6$ barrel comprising helices H3, H4 and H5 (residues 77-160) which interestingly leads to a repositioning of residues E78 and D144. In particular, and as a result of this, the side-chain of D144 is repositioned to be closer to the OH2 and OH3 groups of the –1 subsite sugar

At first sight, the crystal contacts do not appear to impede active site cleft accessibility. However, crystal soaking in relatively high concentrations of xylohexaose results in the occupation of the remote +2 and +3 subsites only (unpublished results). A possible explanation is that the crystal packing precludes a more pronounced hinge type movement needed to accommodate the substrate. In this respect, it is worth noting that the omega loop region 263-275, which lines the narrowest part of the binding cleft (at subsite –2) is involved in lattice contacts (via Y266, F268 and N269).

In contrast to pXyl, no significant large scale movements are observed upon binding of cellopentaose/cellotriose to CelA. While this may arise from differences in crystal packing, the higher thermostability (pXyl is cold-adapted), and hence possibly also rigidity, of this thermophilic enzyme is also a possible explanation. On the other hand, in Rex (a mesophilic enzyme) large-scale movements (C_α shifts up to 1.8 Å) of loops and helices are also observed upon binding of D-xylose and xylobiose (in particular residues 232-358). Contrary to pXyl, however, these movements result in a closing of the binding cleft. Indeed, since Rex displays

an exo-mode of substrate-cleavage and contains fewer binding sites, comparison to pXyl with respect to ligand-induced conformational changes is not straight-forward.

In addition to the structural changes described above, a number of local changes within the binding cleft of the enzyme are also observed. Firstly, substrate stacking interactions at subsites +3 and +1 are optimized. At subsite +3, a small movement of the Y378 side-chain around its χ-2 dihedral angle is observed, while at subsite +1 the side chain of Y381 moves towards the ligand by 0.7 Å and 1.1 Å in the product and substrate complexes, respectively (Figure 2). Remarkably, this movement of Y381 is counterbalanced by W82 which has its indole ring lying approximately in the same plane at a minimal distance of 3.5 Å (the corresponding movement is resp. 0.6 Å and 1.1 Å in the opposite direction). The importance of these movements may reside in the fact that residues Y381 and W82 interact directly with the strictly conserved residues R284 and E78 that undergo the largest conformational shifts of all the active site residues upon ligand binding.

R284 moves considerably only in the substrate-complex, losing its H bond interactions with D144 and Y381 in favour of H bonds with E78 and S202 (Figure 2A). As no similar movement is observed in the unliganded D144A mutant (*15*) it can be concluded that substrate binding, and not the mutation itself, causes this rearrangement. It can also be seen in this complex that the proton donor side chain is fixed in a similar position to that in the unliganded wild-type enzyme, mainly by H bonding to the OH3 group of the D-xylose at subsite –1. Interestingly however, modelling of D144 into this structure (results not shown) indicates that it would make a steric clash with the side chain of E78 and, hence, the position observed for this latter would probably not occur in the wild-type enzyme. It is possible that the D144A mutation allows for an enhanced conformational freedom of the E78 side-chain which can now adapt a position other than that which can occur upon ligand binding in the wild-type enzyme.

In the product complex (Figure 2B), it can be seen that the side chain of Q78 moves considerably upon ligand binding, losing its H bond interaction with D144 in favour of a H bond with Y380. This repositioning is not observed in the structure of the same E78Q mutant in the absence of ligand (*15*), and hence confirms that the rearrangement is induced by ligand binding. Indeed, this conformational displacement clearly results in a considerably more catalytically competent position for E78 which can now form a 2.8 Å H bond with the oxygen of the scissile glycosidic bond. Interestingly, assuming R284 takes on the position observed in the D144A-X5 complex, the proton donor may now also interact with R284, thus allowing for stabilization of a negative charge on E78 at the transition state. Moreover, a similar position for the side chains of E78 and R284 is observed in all complexes of family 8 enzymes studied to date (*18, 20*) as well as in the unliganded pXyl mutant D144N (*10, 15*) which is the only apo-enzyme for which an opening of the substrate binding cleft has been noted. Thus, it is possible that substrate binding induces this repositioning of side-chains within the active-site into more catalytically competent positions, with residue D144 playing a critical role in the process. In the case of the D144A-X5 complex (Figure 2A) it is possible that mutation of D144 to alanine affects this 'induced-fit' mechanism, resulting in a mispositioning of the proton donor and leading to the observed large activity decrease (10^5-fold decrease in its apparent k_{cat} as compared to the wild-type enzyme) (*15*). Also noteworthy is that the observed repositioning of E78, together with the small readjustments of the region containing the catalytic proton acceptor D281, results in an increase of the donor-acceptor distance from 4.5 to 5.1 Å in the product complex, though this is still considerably shorter than the archetypical distance of ~9.5 Å for inverters (*38*).

Analysis of the available family 8 structures (*17-20*) indicates that in all cases the proton donor residue is located in a solvent-accessible cavity containing one or more water molecules (Figure 3) and this conserved cavity is probably necessary to allow for the 'induced fit' movements described above. Moreover, primary sequence alignments indicate that the critical

residues lining this pocket are either strictly conserved (A142, pXyl numbering) or are replaced by residues with short side-chains (G145 in pXyl is replaced by A or S only). However, it is important to note that A142 may also have a role in substrate binding at subsites –1 and –2. Although the conserved hydrophilic cavity may lower the p*Ka* value of the catalytic proton donor, this effect is likely to be negligible considering the close proximity of D144 (expected to be negatively charged around the optimum pH (*10*)) to E78 in the 'inactive' conformation. Indeed, increasing the p*Ka* of the catalytic proton donor may be one of the potential roles of the 'inactive' conformation. Additionally, in this 'inactive' position, E78 is bridged to the proposed catalytic base D281 via the nucleophilic water molecule. Hence, one can even envisage this conformation being important for proton exchange (via the water bridge) after product release, and thus allowing for completion of the catalytic cycle. Interestingly, the 'inactive' conformation of the proton donor has been observed in one other GH-8 (atomic resolution) structure, that of unliganded CelA (*39*), while all other structures to date display the 'active' conformation of the proton donor in the absence of substrate.

Carbohydrate conformation

Analysis of the puckering parameters of the bound sugar (Table 2) as defined by Cremer and Pople (*40*) reveals significant distortion from the most stable ground state $^4C_1$ conformation for the D-xylosyl moieties at subsites –1 and –2. The Θ angle deviations are indeed not extremely large at these subsites but they are significant as they occur at two consecutive subsites. Although the distortion is modest compared to the $^{2,5}B$ like conformation observed at subsite –1 in the CelA-cellopentaose complex (*18*), it indicates that pXyl is not optimized for stabilizing the ground state conformation of its substrate. Examination of the puckering parameters of X3 in the product complex and of X5 at the non-catalytic binding site reveals no significant deviations from the $^4C_1$ form, except for the D-xylosyl moiety at the reducing end (subsite 5) of the latter. However, the aberrant puckering colatitude (Θ), in this case, is unreliable due to a weak electron density.

With the exception of the –3/–2 linkage, the φ and ψ torsion angles of the glycosidic linkages deviate significantly from the values reported for the CelA complexes. In CelA, the V-shaped form of the binding cleft results in deviations from the normal torsion angles mainly at the –1/+1 linkage (*18*). In contrast, these torsion angles (Table 2) are anomalous at the –2/–1 linkage in pXyl, whereas the subsequent linkages –1/+1 and +1/+2 approach the values reported for the approximate left-handed 3-fold helical structure of crystalline xylan (*41*) (maximum deviation is 22°). Finally, surprisingly, the +2/+3 linkage is more in accordance with the flat, extended, ribbon-like structure of crystalline cellulose than that of xylan and an intramolecular ($O5_i$-$O3_{i+1}$) hydrogen bond (2.7 Å) characteristic of cellulose is even found between these sugars. Interestingly, CelA exhibits the reverse situation: the +1/+2 linkage approximates the cellulose conformation (including a 2.8 Å $O5_i$-$O3_{i+1}$ H bond) (*42*) while the +2/+3 linkage approaches the xylan conformation. A further important observation is the 2.6 Å hydrogen bond between the OH6 of D-glucose at subsite –1 and of OH3 at subsite +1 of CelA, a linkage that evidently is absent in pXyl.

Finally, the linkages between the D-xylosyl moieties at the aspecific binding site presents smaller deviations from the extended forms of cellulose and xylan, yet do not clearly conform to the 3n-helical structure of xylan. Illustrative is the relative orientation of the second and third D-xylosyl moieties which is more cellulose-like (a 2.7 Å intramolecular hydrogen bond is observed between $O5_i$ and $OH3_{i+1}$).

Substrate binding sites

In accordance with the kinetic characterization of pXyl (*8*), we have found that the substrate binding cleft contains six subsites, three on each side of the cleavage site. Their location and binding mode (Figures 1B and 4) are quite similar to those of CelA, with aromatic side chains

lining the different subsites and water-mediated and direct hydrogen bonds being involved. Furthermore, the subsites previously described as possible +3 and +4 binding sites on the basis of wild type pXyl in complex with xylobiose (*10*) are more likely alternative +2 and +3 sites. The aberrant location of the D-xylosides at these alternative subsites (about 3 Å more distal from the active site) can be attributed to the fact that the xylobiose complex was obtained by direct soaking of crystals which, as discussed above, prevents full access to the binding cleft. At the same time, it illustrates that the enzyme presents a rather smooth binding profile (gliding surface) (*43–45*) around subsites +2 and +3. This could enable directing of the leaving product as previously suggested (*10*).

Inspection of $F_o - F_c$ maps of the substrate complex reveals the presence of an extra xylo-oligosaccharide binding site in a low-homology region remote from the catalytic site, as evidenced by clear electron density for a linear chain of five D-xylosyl moieties (Figure 1). Although hydrogen bonds are formed with the D-xylosyl moieties at positions 1, 3 and 5 (starting from the non-reducing end of the chain), binding appears to be dominated by the stacking of two aromatic residues (W249 and Y315) with the first two sugars. This sandwich-type of hydrophobic platform (*46, 47*) forms a short cleft having the potential to accommodate a range of small xylo- or gluco-configured ligands. Nevertheless, as the D-xylose at position 5 is involved in a lattice contact with Q119 and as non-catalytic binding sites are traditionally found in separate carbohydrate-binding domains (*47*), further investigation is needed to assess the significance of this carbohydrate-binding site in pXyl.

Even though the catalytic site of pXyl is found to be highly similar to that of CelA, a more detailed analysis of each of the individual subsites, as well as a comparison with those of Rex, ChoK and CelA as described below allows for a better understanding of the substrate binding and specificity of pXyl.

The plus subsites

Of subsites +1, +2 and +3, the latter appears to be of least importance, both in terms of strength of binding and substrate specificity. In effect, the D-xylosyl moiety at this subsite could not be determined in the D144A-X5 complex because of poor electron density. Furthermore, even though a ~3 Å shift in the relative position is observed, the D-xylosyl moiety in pXyl (E78Q-X3) and D-glucosyl in CelA have a similar orientation (Figure 1B) and, in both cases, a conserved tyrosine residue is involved whereas the polar interactions differ. A further interesting observation at this subsite is that the α-configuration of the D-xylosyl moiety is stabilized by a hydrogen bonding interaction with glutamine 327 (Figure 4B), this being in contrast to xylotriose in solution where the β-configuration dominates (*48*). In fact, Q327 was also refined with two alternate conformations, in both unliganded and liganded pXyl, and both of these give rise to hydrogen bonds (3.0 and 2.7 Å) with the α-anomeric form of the xylopyranosyl residue. Presently, the reason for the stabilization of the α-configuration in this subsite is unknown.

The presence of well-defined electron density at subsite +2 in the substrate complex indicates stronger substrate binding at this subsite as compared to that at subsite +3. Furthermore, while the relative positions of the D-xylosyl and D-glucosyl moieties of pXyl and CelA are closer (~1 Å), their orientation differs by ~35° (Figure 1B). No direct hydrogen bonds are observed (neither in pXyl nor CelA) and, hence, this position of the sugar unit is governed by aromatic stacking interactions with F280 in pXyl and Y277 in CelA (Figures 4 and 5A). In addition, the positions of these aromatic side chains themselves are governed, via steric effects, by residues W283 and F336, respectively (Figure 5A). In fact, the tryptophan residue (W283 in pXyl) is found to be conserved in GH-8 xylanases, whereas more compact residues are found in all GH-8 cellulases, chitosanases and lichenases. Interestingly, a serine is present at this position in the GH-8 enzyme BglBC2 from *Bacillus circulans* ATCC 21367, annotated as a cellulolytic xylanase (*49*) while Rex, which has been shown to have an obstructed +2 site, does not have a

conserved stacking aromatic residue. In addition, ChoK, whose substrate glucosamine is structurally similar to glucose, has an aromatic stacking system similar to that of CelA.

The altered orientation of the aromatic stacking system observed in pXyl and in GH-8 xylanases may play a role in substrate specificity (*50, 51*), firstly, via an optimized topography for xylan binding, as it imposes a helical turn upon the +1/+2 glycosidic link similar to that observed in crystalline xylan, and secondly, via the narrowing effect on the binding cleft which may cause steric repulsion against the extra hydroxymethyl group of D-glucose (by means of Y381). Finally, the orientation of the D-xyloside observed at subsite +2 may be important for the binding of decorated xylan substrates due to the absence of steric hindrance to substituents at OH2 and OH3 and the presence of R337 which is favorably located to interact with negatively charged (4-*O*-methyl-) D-glucuronic acid substituents.

In contrast to subsite +2, subsite +1 is characterized by a combination of aromatic stacking (Y381) and conserved polar interactions (Figure 4) and, not surprisingly, the positions of the D-xylosyl and D-glucosyl moieties at this subsite are very similar in pXyl and CelA (Figure 1B). The dominant interaction at this subsite appears to be that with Y381 which, with one exception (the endo–1,4-glucanase fragment from *Cellulomonas uda*) (*49*) is found to be highly conserved, or replaced by a phenylalanine, in GH-8 enzymes. Nevertheless, a slight rotation of the D-xylosyl moiety in the pXyl-substrate complex as compared to D-glucose in CelA is observed due to the presence of a H bond between the OH6-group and the proton acceptor in the CelA complex. Interestingly, this rotation becomes even more pronounced in the pXyl product complex, due to the formation of a H bond between the D-xylosyl OH4 group and the proton acceptor residue. Furthermore, and as a result of this, the product is shifted towards subsite –1 compared to the substrate. This observation contrasts with the situation observed in CelA where the product moves in the opposite direction, resulting in the disruption of the OH6-D278 interaction. This phenomenon, which was suggested to play a role in product release in CelA (*18*), is thus not operative in pXyl. Finally, because the OH2 and OH3 groups of the D-xylosyl in this subsite point towards the protein, a decorated sugar unit (*52*) is unlikely to bind here in pXyl.

Subsites –3, –2

Subsite –3 of pXyl shows only superficial resemblance to that of CelA because of local differences in backbone conformation and consequently, the amino acids involved in carbohydrate recognition differ. Furthermore, the number and strength of these interactions is decreased as compared to CelA (Figures 1B and 4A). The aromatic stacking interaction (with W205) and hydrogen bonds (three) with the D-glucose OH2- and OH3-groups in CelA are replaced by weaker hydrophobic interactions (with F191, I195 and F268) and only one direct hydrogen bond (D138 - OH2). Furthermore, despite the extended loop region (residues 270-274) which narrows the binding cleft at this subsite, D-glucose binding and decorations at OH3 still appears possible. On, the other hand, substituents at OH2 or a 1,3-glycosidic linkage at the –3/–2 junction would cause severe steric hindrance.

The well-defined electron density obtained at subsite –2 (Figure 1A) indicates that this is a strong binding site. Together with subsite +1, it is essential in bringing the scissile glycosidic bond at subsite –1 into a catalytically competent position. Indeed, our observation of residual electron density at these subsites in maps from crystals without added ligand further indicates that subsites –2 and +1 are high affinity sites. W124, which seems to be the dominant aromatic stacking partner at subsite –2, is strictly conserved among family 8 members, whereas the direct hydrogen bonds observed in CelA (between the D-glucose-specific OH6-group and W205-N_ε and A149-O) are lost in pXyl (Figure 4A) due to variations in the backbone conformation in the 140s region and the replacement of W205 by S191. The former observation is especially important as the ~1.7 Å shift of the P141 (equivalent to A149 of CelA) backbone carbonyl into the cleft would cause a steric clash with a D-glucose-

hydroxymethyl group, both in pXyl and Rex, and thus appears to be a determining factor for D-xylose specificity (Figure 5B). Importantly, this relatively small difference will only have its full effect in combination with the strong stacking interaction with W124 (or W112 in Rex) and the close-fitting binding cleft at subsite –2. In addition, the presence of F191 and F187 in pXyl and Rex, respectively, seems to restrain the ligand at subsite –2 even more compared to that in CelA and ChoK. Primary sequence comparisons indicate, however, that the presence of a bulky aromatic residue at this position is not limited to xylanases. As a final remark on the substrate specificity of subsite –2, the extended 270s loop region does not appear to confer specificity towards D-xylose, but rather may play a role in binding cleft accessibility. Moreover, it seems to prevent subsite –2 from binding substituted xylan due to steric constraints.

Because of the pseudo-symmetry of xylan, it may be able to bind to the cleft in the wrong direction and thereby lead to the formation of a dead-end complex. The question thus arises as to how family 8 xylanases differentiate between the glycon and aglycon regions of their substrate. Strikingly, there are no direct hydrogen bonds with the endocyclic O5 in pXyl and Rex (nor in CelA), but one indirect interaction that may allow for correct orientation of the substrate is observed in all GH-8 complexes: a water bridge between Y203 (pXyl numbering) and O5 at subsite –2 (Figure 4A). Interestingly, this water molecule corresponds to a strong electron density peak and it is one of only four conserved water molecules found in the binding cleft of all GH-8 structures known to date.

Subsite –1

Subsite –1 of pXyl exhibits some notable differences with that of CelA. Firstly, the average D-xylose ring plane is tilted relative to the average D-glucose ring plane and the OH2 and OH3 groups engage in H bonds with residues D281 and Y203 (Figure 5C) probably as a result of the loss of the bidentate interaction with residue 144 as well as steric repulsion by the 'mispositioned' proton donor in the D144A mutant. Secondly, the xylopyranose ring in subsite –1 reveals a distorted $^{4}C_{1}$ conformation instead of a distorted boat ($^{2,5}B$) form as occurs in CelA. Although it cannot be excluded to result from the absence of the D144 side chain in the mutant, a number of arguments can be put forward to support the proposition that the presence of a $^{2,5}B$ conformation is unlikely to occur in the [ES] complex of family 8 xylanases. Given that the stabilizing influence of D144 on the transition state (TS) can primarily be ascribed to electrostatic interactions with the transient redistribution of charges during the transition state (*53, 54*), it does not follow that D144 preferentially stabilizes the neutral $^{2,5}B$ form, as present in the CelA complex. Moreover, the influence of a $^{4}C_{1} \rightarrow {}^{2,5}B$ conformational change on the relative positions of the OH2 and OH3 groups is negligible, atoms C5 and O5 being involved in the major rearrangements. Closer investigation of the complexes of pXyl and CelA at subsite –1 shows that the preference for a $^{2,5}B$ like conformation in the ground-state rather results from other factors. Guérin *et al.* (*18*) have proposed that binding of substrate in the V-shaped binding cleft, with a sufficient number of binding sites being occupied, induces a chain bending ('kink') at the scissile glycosidic linkage, thus imposing a $^{2,5}B$ conformation on the sugar ring at the –1 subsite. Indeed, while we have indications of 'strain' at subsite –1 in the pXyl substrate complex, as discussed above, it appears that chain bending is insufficient to induce a boat conformation. In fact, a number of interactions that specifically stabilize the observed boat conformation appear to be present in CelA. Firstly, the sugars at subsites –1 and at +1 are optimally positioned to allow for the formation of a 2.6 Å hydrogen bond between OH6 at –1 and OH3 at +1, secondly, a H bond, which can not occur with a $^{4}C_{1}$ conformation, is found between OH6 at –1 and R84, and finally, a $^{2,5}B \rightarrow {}^{4}C_{1}$ transition would lead to steric repulsion between the C6OH group and W132. Since these interactions are D-glucose-specific, these forms of 'local strain' are inexistent in pXyl, and this is further supported by the presence of a regular $^{4}C_{1}$ chair form at subsite –1 in the non-(cleavage-site)-spanning

xylobiose complex of Rex (*20*). Importantly, R84 is substituted by A66 and D61 in pXyl and Rex, resp. However, at a similar position, but ~1.5 Å closer to the potential position of a D-glucose-hydroxymethyl group, residues R76 and R68 are found in pXyl and Rex. Together with resp. W124 and W112 (equivalent to W132 in CelA) these residues would seriously impede access of D-glucose to subsite –1. In effect, inspection of the GH-8 sequences indicates that the presence of R or K at this position is xylanase-specific (*15*).

Unfavourable 'bowsprit' interactions and the requirement for strong compensatory protein-carbohydrate interactions are major barriers to the formation of a ground state D-glucoside $^{2,5}B$ conformation and may be one of the reasons behind the observation of a $^4C_1$ form in family 12 cellulases whereas the closely related family 11 xylanases adapt a $^{2,5}B$ conformation (*55–57*). Indeed, the $^{2,5}B$ form of the CelA Michaelis complex is distorted towards the ^2S_O form, which more likely represents a local conformational minimum (*58*) and, furthermore, is located next to $^{2,5}B$ in the pseudo-rotational itinerary (PSI; (*59*)). In the next section we present elements in support of the proposal that family 8 cellulases and xylanases are committed to a conformational change towards a ^2S_O like conformation.

A mechanistic itinerary for substrate hydrolysis by family 8 enzymes

It has been generally accepted that, as a consequence of the antiperiplanar lone pair hypothesis (ALPH), the hydrolysis of β-glycosides should be preceded by a conformational change, away from the ground state chair to a high-energy skew-boat conformation in which the leaving-group has an axial position and thus enables an essential set-up with an antiperiplanar orbital of the ring oxygen lone pair (*60–63*). According to these stereoelectronic considerations, the $^1S_3$, $^1S_5$ and ^2S_O forms are the most likely candidates for the pre-TS local minimum. Given that GH-8 consists of *β*-glycoside hydrolases, at least one of these conformations should be stabilized by its members. This finding agrees well with the presence of a ^2S_O like conformation in the Michaelis complex of CelA. To determine the conformational preference of pXyl in relation to these three potential pre-TS conformations, we have performed docking studies with the program AUTODOCK. Xylotriose ligands differing only in the puckering conformation of the central D-xylosyl moiety were docked into the crystal structure of the E78Q-X3 complex because, as discussed above, this is proposed to best resemble the conformation of the active site of the wild type enzyme in complex with substrate (especially in relation to the proton donor E78). The $^1S_3$, $^1S_5$ and ^2S_O conformations were used for the central D-xylosyl moiety, as a control-ligand, the all-β $^4C_1$ xylotriose was also docked.

The results of the dockings show a clear preference for $^4C_1$ and ^2S_O over $^1S_5$ and $^1S_3$, with the first two conformations giving the best fit to the crystallographic complex (Table 3). Considering the fact that the energetic penalty for a $^4C_1 \rightarrow S$ change is not paid for by AUTODOCK, a preference for ^2S_O instead of $^4C_1$ might even be expected based on the release of strain imposed by the binding cleft on an all-$^4C_1$ form. Docking runs performed with larger substrates to confirm this potential effect, however, were unsuccesful. Despite of this, the docking experiments with xylotriose clearly demonstrate that the $^1S_5$ and $^1S_3$ forms are not complementary to the shape of the binding cleft. In addition, a $^4C_1 \rightarrow 1S_{3/5}$ conformational change from the ground state would be disruptive to hydrogen bonds involving the OH2 and OH3 groups at subsite –1, leading to a destabilizing effect, whereas a $^4C_1 \rightarrow {}^2S_O$ conformational reorganization would leave most of the ground state interactions unaltered. Consequently, taking the strongly conserved binding cleft of GH-8 members into account, the conformational route β-$^4C_1 \rightarrow {}^2S_O \rightarrow$ TS $\rightarrow \alpha$-$^4C_1$ can be put forth as a reasonable, yet incomplete, mechanistic itinerary for the substrate hydrolysis catalyzed by this family. This raises the question if the observation of a distorted $^{2,5}B$ conformation in the CelA-Michaelis complex should be explained as resulting from the stabilization of an ALPH-compliant pre-TS local minimum ^2S_O conformation, rather than a high-energy $^{2,5}B$ like TS as proposed by Guérin *et al*. (*18*) and if the E_3 and $^4H_3$ forms should not be considered as TS candidates. Indeed their

C5, O5, C1 and C2 atoms are coplanar (as in the oxocarbenium-ion-like TS) and only a minor repositioning is required when coming from a $^{2}S_{O}$ form. However, in contrast to a $^{2,5}B$ form, they avoid the unfavourable flagpole interaction of the hydroxymethyl group, axial on the C5 bowsprit, with the C2 hydrogen atom, which *inter alia* would be present in a pre-TS $^{2,5}B$ just as well as in a $^{2,5}B$ TS. Moreover, a (borderline) S_N2-type nucleophilic attack of a water molecule on an anomeric center should proceed in-line from the other side of the leaving group, which can occur when the leaving group is axially positioned (as in the ALPH-compliant $^{1}S_{3}$, $^{1}S_{5}$ and $^{2}S_{O}$ conformations), but is problematic on a $^{2,5}B$ pre-TS since here the leaving group occupies an equatorial position for which the region for nucleophilic attack is occupied by the hydrogen atom on C3 and is as such sterically inaccessible. A strict $^{2,5}B$ conformation of a classical β-substituted glycopyranoside is therefore not a productive situation and should be avoided by the enzyme.

Conclusions

Our analysis of the crystal structures of the pXyl mutants D144A and E78Q in complex with substrate and product, respectively, indicates that substrate specificity in GH-8 xylanases is determined by structural features at subsites +2, –1 and –2. At subsite +2, an alternative orientation of the aromatic stacking system provides an optimized topography for xylan binding and, furthermore, would impose steric repulsion on a D-glucosyl moiety. Subtle differences in backbone conformation impede binding of D-glucose at subsite –2 and steric hindrance may also hamper binding of D-glucose at subsite –1.

This work further delineates the function of several active site residues and follows up on a previous study on pXyl in which we have determined the identity of the general acid and general base catalysts and of the nucleophilic water (15). Comparison of the complexes with the wild-type and mutant structures in their unliganded form has allowed for the demonstration of an induced fit mechanism taking place upon ligand binding. The global conformational changes are in agreement with the high conformational flexibility previously reported for this cold-adapted enzyme (11). Among local changes, the repositioning of the proton donor in a more catalytically competent position is the most notable event.

The structure of the –1/+1 subsite spanning complex reveals that the enzyme imposes ‘strain’ on the substrate, which, along the reaction path, may be alleviated by a conformational change of the sugar residue at subsite –1, bringing the scissile glycosidic linkage into a pseudoaxial position. This is consistent with stereoelectronic expectations (as dictated by ALPH), while simultaneously allowing unrestricted nucleophilic attack at the opposite face of the anomeric carbon. According to our study, a $^{2}S_{O}$ form most likely fulfils this role in family 8 enzymes.

Acknowledgement

We gratefully acknowledge access to beamline BW7A at the EMBL/DESY Hamburg Outstation and thank N. Gerardin, A. Dernier and R. Marchand for their skillful technical assistance.

References

1. Biely, P. (1985) Microbial xylanolytic systems, *Trends Biotechnol. 3*, 286–290.
2. Poutanen, K., Rättö, M., Puls, J., and Viikari, L. (1987) Evaluation of different microbial xylanolytic systems, *J. Biotechnol. 6,* 49–60.
3. Gilbert, H. J., and Hazlewood, G. P. (1993) Bacterial cellulases and xylanases, *J. Gen. Microbiol. 139,* 187–194.
4. Kulkarni, N., Shendye, A., and Rao, M. (1999) Molecular and biotechnological aspects of xylanases, *FEMS Microbiol. Rev. 23,* 411–456.
5. Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities, *Biochem. J. 280*, 309–316.
6. Gebler, J., Gilkes, N. R., Claeyssens, M., Wilson, D. B., Beguin, P., Wakarchuk, W. W., Kilburn, D. G., Miller, R. C., Jr., Warren, R. A., and Withers, S. G. (1992) Stereoselective hydrolysis catalyzed by related β–1,4–glucanases and β–1,4–xylanases, *J. Biol. Chem. 267*, 12559–12561.
7. Henrissat, B., and Coutinho, P. M. (2001) Classification of glycoside hydrolases and glycosyltransferases from hyperthermophiles. *Methods Enzymol. 330*, 183–201.
8. Collins, T., Meuwis, M. A., Stals, I., Claeyssens, M., Feller, G., and Gerday, C. (2002) A novel family 8 xylanase: functional and physico-chemical characterization, *J. Biol. Chem. 277*, 35133–35139.
9. Van Petegem, F., Collins, T., Meuwis, M. A., Gerday, C., Feller, G., and Van Beeumen, J. (2002) Crystallization and preliminary X-ray analysis of a xylanase from the psychrophile *Pseudoalteromonas haloplanktis*, *Acta Crystallogr. D Biol. Crystallogr.* 58, 1494–1496.
10. Van Petegem, F., Collins, T., Meuwis, M. A., Gerday, C., Feller, G., and Van Beeumen, J. (2003) The structure of a cold-adapted family 8 xylanase at 1.3 A resolution. Structural adaptations to cold and investigation of the active site, *J. Biol. Chem. 278*, 7531–7539.
11. Collins, T., Meuwis, M. A., Gerday, C., and Feller, G. (2003) Activity, stability and flexibility in glycosidases adapted to extreme thermal environments, *J. Mol. Biol. 328*, 419–428.
12. Collins, T., Gerday, C. & Feller, G. (2005). Xylanases, xylanase families and extremophilic xylanases. *FEMS Microbiol. Rev. 29*, 3–23.
13. Koshland, D. E. (1953) Stereochemistry and the mechanism of enzymatic reactions, *Biol. Rev. 28*, 416–436.
14. Sinnott, M. L. (1990) Catalytic mechanisms of enzymatic glycosyl transfer, *Chem. Rev. 90*, 1171–1202.
15. Collins, T., De Vos, D., Hoyoux, A., Savvides, S. N., Gerday, C., Van Beeumen, J., and Feller, G. (2005) Study of the active site residues of a glycoside hydrolase family 8 xylanase, *J. Mol. Biol., 354,* 425-435.
16. Davies, G. J., Wilson, K. S., and Henrissat, B. (1997) Nomenclature for sugar-binding substrates in glycosyl hydrolases, *Biochem. J. 321*, 557-559.
17. Alzari, P. M., Souchon, H., and Dominguez, R. (1996) The crystal structure of endoglucanase CelA, a family 8 glycosyl hydrolase from *Clostridium thermocellum*, *Structure 4*, 265–275.
18. Guérin, D. M. A., Lascombe, M.-B., Costabel, M., Souchon, H., Lamzin, V., Béguin, P., and Alzari, P. M. (2002) Atomic (0.94 Å) resolution structure of an inverting glycosidase in complex with substrate, *J. Mol. Biol. 316*, 1061–1069.
19. Adachi, W., Sakihama, Y., Shimizu, S., Sunami, T., Fukazawa, T., Suzuki, M., Yatsunami, R., Nakamura, S., and Takénaka, A. (2004) Crystal structure of family GH-8 chitosanase with subclass II specificity from *Bacillus sp.* K17, *J. Mol. Biol. 343*, 785–795.

20. Fushinobu, S., Hidaka, M., Honda, Y., Wakagi, T., Shoun, H., and Kitaoka, M. (2005) Structural basis for the specificity of the reducing end xylose-releasing exo-oligoxylanase from *Bacillus halodurans* C-125, *J. Biol. Chem. 280*, 17180–17186.
21. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307–326.
22. Murshudov, G. N. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr. D Biol. Crystallogr. 53*, 240–255.
23. Collaborative Comutational Project, Number 4 (1994) *Acta Cryst. D50*, 760–763.
24. Roussel, A., and Cambillau, C. (1992) Turbo-Frodo, Biographics, AFMB, Marseille, France.
25. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemistry quality of protein structures, *J. Appl. Crystallog. 26*, 283–291.
26. Speck, A. L. (2001) PLATON, a multipurpose crystallographic tool, Utrecht University, The Netherlands.
27. McDonald, I. K. & Thornton, J. M. (1994) Satisfying hydrogen bonding potential in proteins, *J. Mol. Biol. 238*, 777–793.
28. 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.
29. DeLano, W. L. (2002) The PyMOL Molecular Graphics System at URL: http://www.pymol.org.
30. Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., and Olson, A. J. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, *J. Computational Chemistry 19*, 1639–1662.
31. Allinger, N. L. (1997) MM2. A hydrocarbon force field utilizing v_1 and v_2 torsional terms, *J. Am. Chem. Soc. 99,* 8127–8134.
32. Burkert, U., and Allinger, N. L. (1982) *Molecular Mechanics*, ACS Monograph 177, American Chemical Society, Washington, DC.
33. Guex, N., and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling, *Electrophoresis 18*, 2714–2723.
34. Gasteiger, J., and Marsili, M. (1980) Iterative partial equalization of orbital electronegativity – a rapid access to atomic charges, *Tetrahedron 36,* 3219–3228.
35. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Jr., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) A second generation force-field for the simulation of proteins, nucleic acids, and organic molecules, *J. Am. Chem. Soc. 117,* 5179–5197.
36. Stouten, P. F. W., Frömmel, C., Nakamura, H., and Sander, C. (1993) An effective solvation term based on atomic occupancies for use in protein simulations, *Molecular Simulations 10*, 97–120.
37. Solis, F. J., and Wets, R. J.–B. (1981) Minimization by random search techniques, *Math. Oper. Res. 6,* 19–30.
38. Rye, C. S., and Withers, S. G. (2000) Glycosidase mechanisms, *Curr. Opin. Chem. Biol. 4*, 573–580.
39. Schmidt, A., Gonzalez, A., Morris, R. J., Costabel, M., Alzari, P.M., and Lamzin, V.S. (2002) Advantages of high-resolution phasing: MAD to atomic resolution, *Acta Crystallogr D Biol Crystallogr. 58*, 1433–1441.
40. Cremer, D., and Pople, J. A. (1975) A general definition of ring puckering coordinates, *J. Am. Chem. Soc. 97*, 1354–1358.

41. Atkins, E. D. T. (1992) Three-dimensional structure, interactions and properties of xylans, in *Xylans and xylanases* (Visser, J., Ed.), pp 349–360, Elsevier Science Publisher, Amsterdam.
42. Kroon-Batenburg, L. M. J., and Kroon, J. (1995) The crystal and molecular structures of cellulose, *Carbohydrates in Europe 12*, 15–19.
43. Dutzler, R., Wang, Y.-F., Rizkallah, P. J., Rosenbusch, J. P., and Schirmer, T. (1996) Crystal structure of various maltooligosaccharides bound to maltoporin reveal a specific translocation pathway, *Structure 4*, 127–134.
44. Meyer, J. E. W., and Schulz, G. E. (1997) Energy profile of maltooligosaccharide permeation through maltoporin as derived from the structure and from a statistical analysis of saccharide-protein interactions, *Protein Sci. 6*, 1084–1091.
45. Parsiegla, G., Juy, M., Reverbel-Leroy, C., Tardif, C., Belaïch, J.-P., Driguez, H., and Haser, R. (1998) The crystal structure of the processive endocellulase CelF of *Clostridium cellulolyticum* in complex with a thiooligosaccharide inhibitor at 2.0Å resolution, *EMBO J. 17*, 5551–5562.
46. Robert, X., Haser, R., Gottschalk, T. E., Ratajczak, F., Driguez, H., Svensson, B., and Aghajari, N. (2003) The structure of barley α-amylase isozyme 1 reveals a novel role of domain C in substrate recognition and binding: a pair of sugar tongs, *Structure 11*, 973–984.
47. Boraston, A. B., Bolam, D. N., Gilbert, H. J. and Davies, G. J. (2004) Carbohydrat-binding modules: fine-tuning polysaccharide recognition, *Biochem. J. 382*, 769–781.
48. Honda, Y., and Kitaoka, M. (2004) A family 8 glycoside hydrolase from *Bacillus halodurans* C-125 (BH2105) is a reducing end xylose-releasing exo-oligoxylanase, *J. Biol. Chem. 279*, 55097–55103.
49. Coutinho, P. M., and Henrissat, B. (1999) Carbohydrate-active enzyme server (CAZY) at URL: http://afmb.cnres-mrs.fr/~cazy/CAZY/.
50. Simpson, P. J., Bolam, D. N., Cooper, A., Ciruela, A., Hazlewood, G. P., Gilbert, H. J., and Williamson, M. P. (1999) A family IIb xylan-binding domain has a similar secondary structure to a homologous family IIa cellulose-binding domain but different ligand specificity, *Structure 7*, 853–864.
51. Szabo, L., Jamal, S., Xie, H., Charnock, S. J., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2001) Structure of a family 15 carbohydrate-binding module in complex with xylopentaose, *J. Biol. Chem. 276*, 49061–49065.
52. Pell, G., Taylor, E. J., Gloster, T. M., Turkenburg, J. P., Fontes, C. M. G. A., Ferreira, L. M. A., Nagy, T., Clark, S. J., Davies, G. J. and Gilbert, H. J. (2004) The mechanisms by which family 10 glycoside hydrolases bind decorated substrates, *J. Biol. Chem. 279*, 9597–9605.
53. Warshel, A. (1991) Computer simulation of chemical reactions in enzymes and solutions, John Wiley & Sons, New York.
54. Warshel, A. (1998) Electrostatic origin of the catalytic power of enzymes and the role of preorganized active sites, *J. Biol. Chem. 273*, 27035–27038.
55. Sulzenbacher, G., Mackenzie, L. F., Wilson, K. S., Withers, S. G., Dupont, C., and Davies, G. J. (1999) The crystal structure of a 2-fluorocellotriosyl complex of the *Streptomyces lividans* endoglucanase CelB2 at 1.2 Å resolution, *Biochemistry 33*, 12546–12552.
56. Sabini, E., Wilson, K. S., Danielsen, S., Schulein, M., and Davies, G. J. (2001) Oligosaccharide binding to family 11 xylanases: both covalent intermediate and mutant product complexes display $^{2,5}B$ conformations at the active centre, *Acta Cryst. D57*, 1344–1347.

57. Davies, G. J., Ducros, V. M.-A., Varrot, A., and Zechel, D. L. (2003) Mapping the conformational itinerary of β-glycosidases by X-ray crystallography, *Biochem. Soc. 31*, 523–527.
58. Dowd, M. K., French, A. D., and Reilly, P. J. (1994) Modeling of aldopyranosyl ring puckering with MM3(92), *Carbohydr. Res. 264,* 1–19.
59. Stoddart, J. F. (1971) *Stereochemistry of Carbohydrates,* Wiley Interscience, Toronto.
60. Deslongchamps, P. (1983) Stereoelectronic effects in organic chemistry, Pergamon Press, Oxford, New York.
61. Kirby, A. J. (1983) The anomeric effect and related stereoelectronic effects at oxygen, Springer-Verslag, Berlin.
62. Deslongchamps, P. (1993) Intramolecular strategies and stereoelectronic effects. Glycosides hydrolysis revisited, *Pure & Appl. Biochem. 65*, 1161–1178.
63. Davies, G. J., Ducros, V. M.–A., Varrot, A., and Zechel, D. L. (2003) Mapping the conformational itinerary of β-glycosidases by X-ray crystallography, *Bioch. Soc. Trans. 31,* 523–527.

Table 1: Diffraction data processing and refinement statistics for the xylopentaose and xylotriose complexes

	D144A-X5	E78Q-X3
A. ***Diffraction data***		
Resolution (Å)	20-1.95 (2.01-1.95)	50-1.88 (1.92-1.88)
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters	a = 50.7	a = 51.0
	b = 91.1	b = 90.8
	c = 99.6	c = 98.1
No. Reflections		
Total	174,379	94,091
Unique	34,436	34,526
R_{merge} [a] (%)	6.0 (28.6)	8.5 (40.0)
I/σ (I)	26.6 (3.5)	10.3 (2.0)
Completeness (%)	100 (100)	90.3 (82.4)
B. ***Refinement***		
R_{work}/R_{free} (%)	14.6/17.9	16.0/18.7
Rms bond length deviations (Å)	0.011	0.008
Rms bond angle deviations (°)	1.29	1.02
No. atoms		
Protein	3263	3256
Solvent	448	431
Ligand	46/46[b]	29
Average B-factor (Å^2)		
Main chain	18.0	13.4
Side chain	18.9	14.1
Solvent	32.8	26.6
Ligand	34.8/36.7[b]	28.4
Ramachandran plot		
% most favoured regions	92.1	92.1
% additional allowed	7.9	7.9

Values between parentheses reflect the data in the highest resolution shell.

[a] $R_{merge} = \Sigma_h\Sigma_i|I(h,i)-\langle I(h)\rangle|/\ \Sigma_h\Sigma_i I(h,i)$, where I(h,i) is the intensity of the ith measurement of reflection h and $\langle I(h)\rangle$ is the average value from multiple measurements.

[b] The second value is for the ligand at the non-catalytic binding site.

Table 2: Sugar conformation parameters for the D144A-X5 and E78Q-X3 complexes

Sugar site	Ring puckering parameters[a]			Form	Dihedral angles[b]	
	Φ_2 (°)	Θ (°)	Q (Å)		φ (°)	ψ (°)
D144A						
Catalytic						
–3	272	4	0.57	$^4C_1$	–75	
–2	320	14	0.50	$^4C_1$	–120	130
–1	2	14	0.55	$^4C_1$	–59	165
+1	283	9	0.56	$^4C_1$	–78	138
+2	347	6	0.56	$^4C_1$		157
Non-catalytic						
1	186	6	0.54	$^4C_1$	–103	
2	96	6	0.52	$^4C_1$	–74	139
3	296	4	0.57	$^4C_1$	–79	98
4	90	4	0.53	$^4C_1$	–98	131
5	71	21	0.55	$^4C_1$		137
E78Q						
+1	326	8	0.54	$^4C_1$	–70	
+2	241	4	0.54	$^4C_1$	–85	153
+3	71	1	0.53	$^4C_1$		101
Xylan[c]					–65	+135
Cellulose[d]					–94	+92

[a] The puckering parameters (Φ_2, Θ, Q) were calculated with PLATON (*26*).
[b] The dihedral angles are defined as: φ ($O5_i$-$C1_i$-$O4_{i+1}$-$C4_{i+1}$) and ψ ($C1_i$-$O4_{i+1}$-$C4_{i+1}$-$C3_{i+1}$).
[c,d] Values for xylan and cellulose are shown for comparison (*41, 42*).

Table 3: Xylotriose-docking statistics

Conformation of the central xylopyranose	Number of members in the lowest-energy cluster	Rmsd[a] from the substrate complex
$^{1}S_{5}$	22/50	4.20
$^{1}S_{3}$	4/50	4.50
$^{2}S_{O}$	17/50	2.17
$^{4}C_{1}$	12/50	1.53

[a] Based on the best-docked member of the lowest-energy cluster.

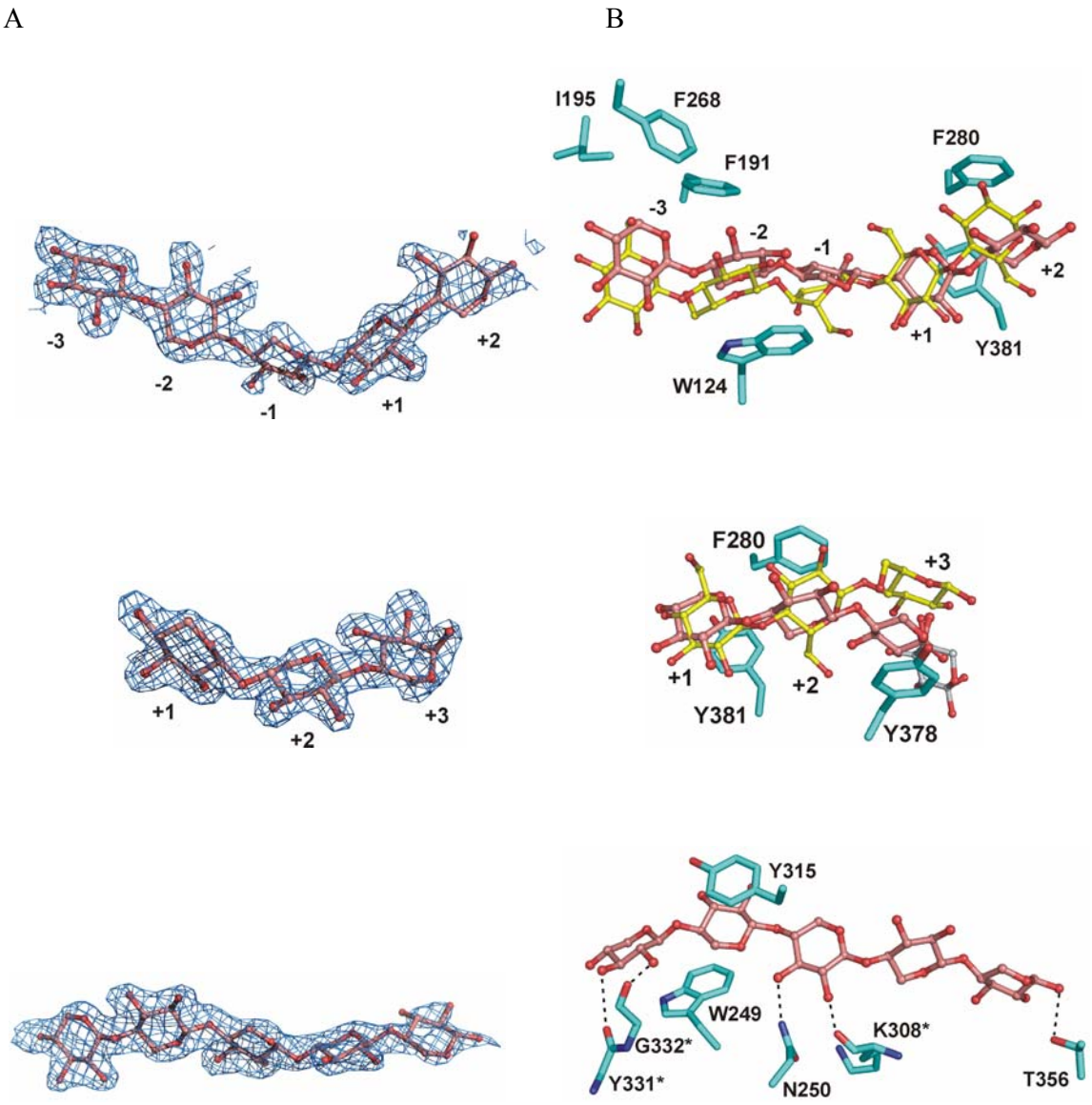
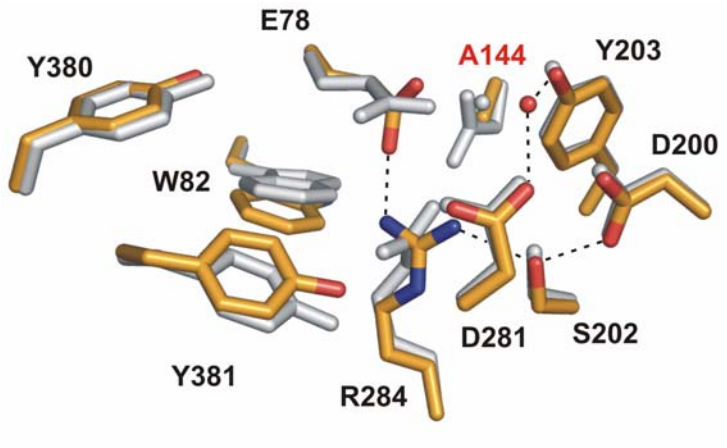


Figure 1. (A, left) $2F_o–F_c$ electron density maps contoured at 1.0σ for the substrate (top), product (middle) and xylopentaose at the non-catalytic binding site (bottom). (B, right) Superposition of the pXyl (in pink) and CelA (in yellow) substrate (top), and product (middle) complexes. For the product complex the xyloside at the alternative subsite +3 is shown in grey and the xylosyls at the reducing end are both shown as a combination of α- and β-configured models. Side chains involved in important hydrophobic interactions are shown for pXyl. At the bottom right is shown a view of the major protein-carbohydrate interactions of the non-catalytic binding site. The reducing end of the xylopentaose chain is located at the right. Main chain interactions are indicated by a '*'.

A



B

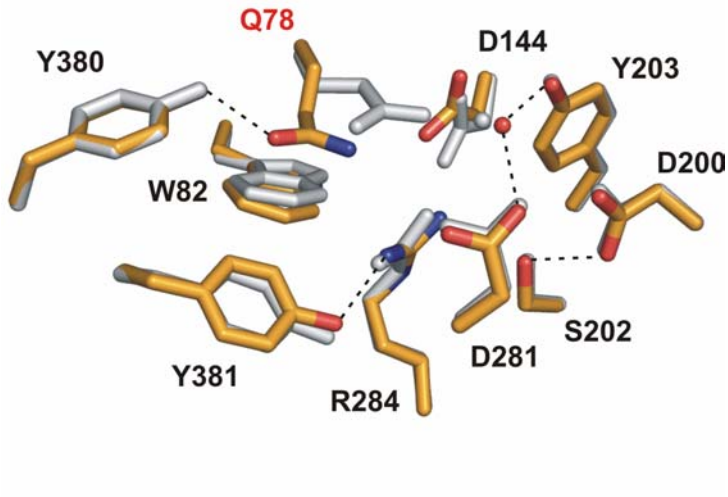
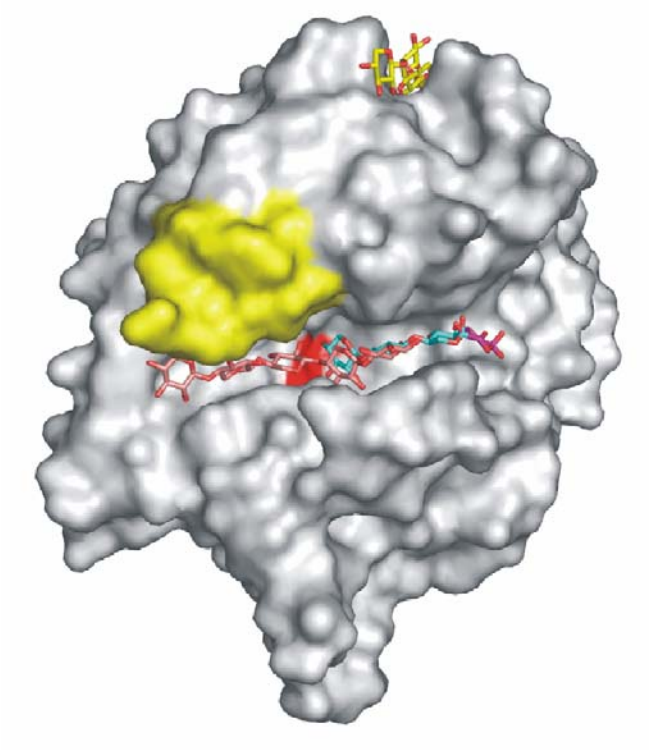


Figure 2. Superposition of the main active site residues (sticks) and the proposed nucleophilic water (spheres) of wild-type pXyl (grey) with those of the D144A-X5 (A), and E78Q-X3 (B) complexes (color-coded according to atom type). The hydrogen bonds and nucleophilic water in the complexes are indicated by dashed lines and a red sphere, respectively, while the nucleophilic water in the wild-type is indicated by a grey sphere.

A



B

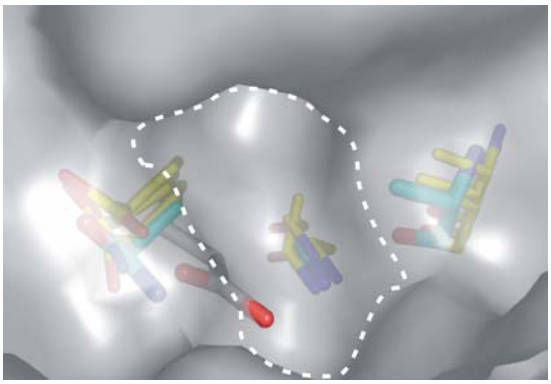


Figure 3. (A) Superposition of different ligands onto the surface of pXyl E78Q-X3. Xylopentaose at the catalytic site is shown in pink, xylotriose in cyan, the xylosyl of the pXyl-xylobiose complex described previously (*10*) is shown in purple and xylopentaose at the non-catalytic binding site is depicted in yellow. The region of the molecular surface involved in lattice contacts (270s region) is in yellow and the approximate location of the hydrophilic cavity is in red. (B) Details of the molecular surface at the hydrophilic cavity conserved in GH-8 members. Residues lining the cavity of pXyl mutant E78Q in complex with xylotriose (in cyan) and the superposed equivalent residues of CelA, ChoK and Rex (in yellow) are depicted as sticks. The cluster of residues at the left correspond to the catalytic proton donor. The side chain (in grey) that perforates the molecular surface is that of the catalytic proton donor of the unliganded wild-type pXyl. The cluster of residues in the middle and on the right correspond to G145 (and equivalent residues) and A142 (and equivalent residues) respectively.

A. B

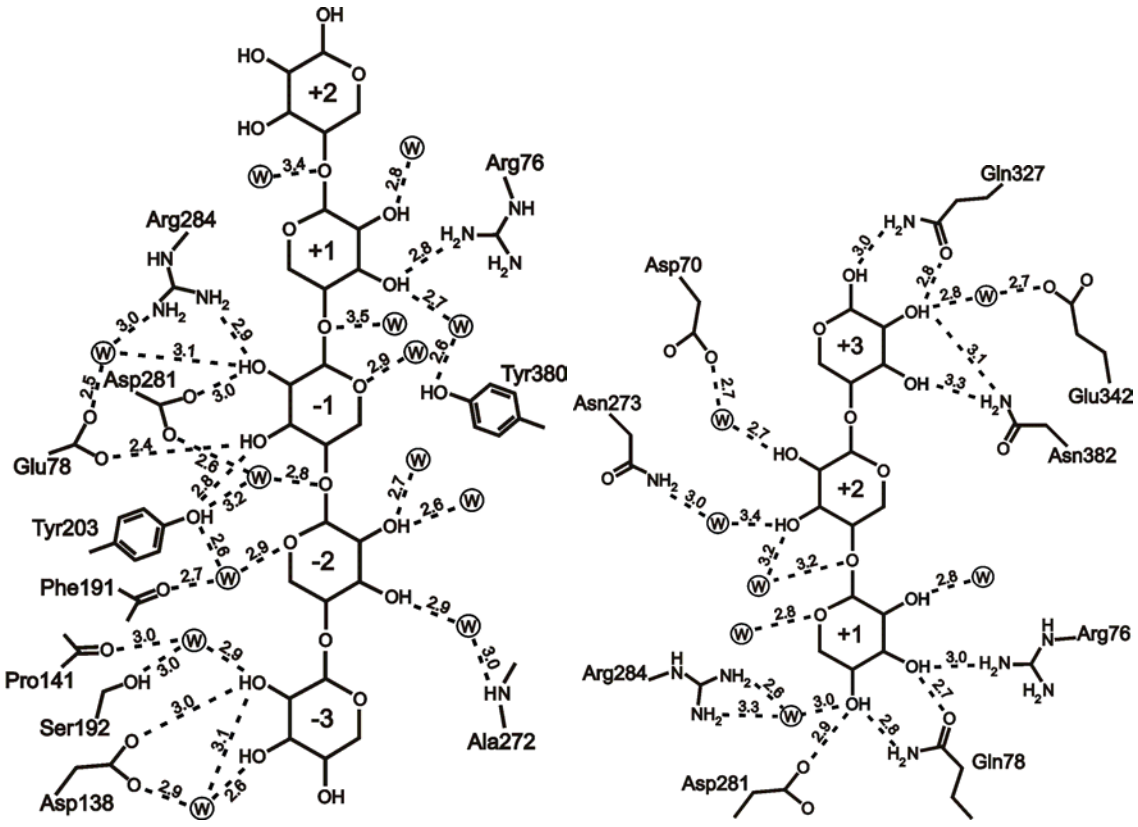


Figure 4. Schematic presentation of the potential water mediated and direct protein-carbohydrate H bonding interactions (in Å) of the D144A-X5 (left) and E78Q-X3 (right) complexes.

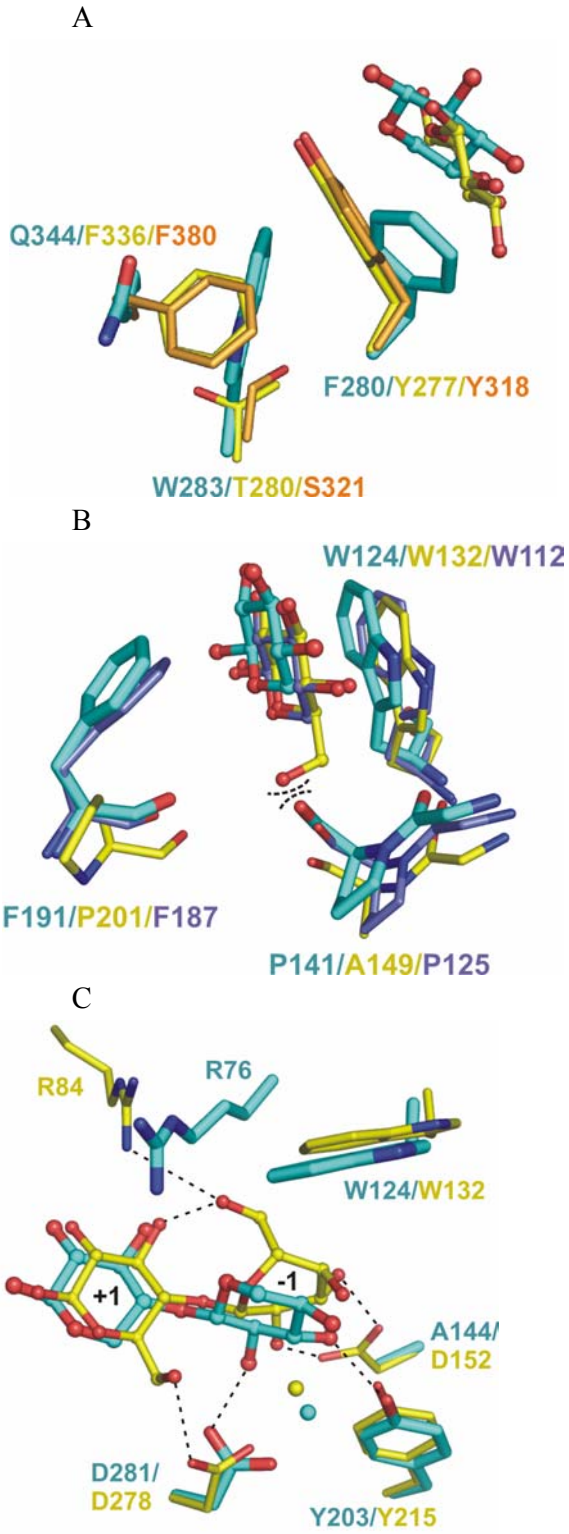


Figure 5. Detailed presentation of the superposed subsites +2 (A), –2 (B) and –1 (C) of GH-8 members. Amino acids are shown as sticks, sugar residues as ball-and-stick and nucleophilic water molecules as spheres. pXyl residues are colored cyan, CelA residues yellow, ChoK residues orange and Rex residues purple. In (B), the repulsive interactions between the glucose of CelA and the pXyl and Rex backbone carbonyl groups are indicated by curved lines. In (C), hydrogen bonds specific to the pXyl D144A- and CelA E95Q-substrate complexes are shown as dashed lines.

Overall conclusions and future prospects

This work represents a contribution to the understanding of enzymes from extremophiles through the study of the aspartate carbamoyltransferases from the hyperthermophilic archaeon *Sulfolobus acidocaldarius* and from the psychrophilic deep-sea bacterium *Moritella profunda*, as well as a glycoside hydrolase family 8 xylanase from the psychrophilic bacterium *Pseudoalteromonas haloplanktis*.

Sulfolobus acidocaldarius ATCase

E. coli aspartate carbamoyltransferase has become a model system for cooperativity and allosteric behaviour of which its dodecameric architecture is an essential component. The crystal structure of the unliganded ATCase of *S. acidocaldarius*, determined by X-ray crystallography to 2.6 Å resolution, represents the first structure of a holoenzyme of the same structural class. Comparison with the *E. coli* ATCase revealed differences in tertiary and quaternary structure. Conformational differences in the 240s loop region of the catalytic chain and in the C-terminal region of the regulatory chain affect intersubunit and interdomain interfaces implicated previously in the *E. coli* ATCase allosteric behaviour. These structural rearrangements may well promote the thermostability of the hyperthermophilic enzyme. The changes at the interfaces may also explain the importance of the regulatory subunits in stabilizing the holoenzyme and furthermore why the catalytic trimers are intrinsically unstable.

Several elements from the crystal structure indicate that the unliganded *S. acidocaldarius* enzyme is in a structural T state, yet one that is shifted towards the *E. coli* R state. Our preliminary analysis of sedimentation velocity experiments with the bisubstrate analogue PALA seem to confirm that the unliganded enzyme is indeed in a structural T state and that a conformational transition takes place, similar in amplitude to that of *E. coli* ATCase. A first analysis of the *S. acidocaldarius* ATCase in complex with CTP indicates very similar modes of binding, which confirms the expectation that the differences in regulatory behaviour are caused by differences in the interfaces that are thought to be involved in effector-signal transmission. Interestingly, a global conformational shift has not been observed, yet small local differences are present.

Future work should be primarily directed towards obtaining crystals of the allosteric R state to obtain a more solid ground for explaining homotropic and heterotropic effects. Several attempts have been made, including direct soaking as well as cocrystallization with different substrates and inhibitors. Direct soakings of crystals obtained in low salt conditions (avoiding weakened substrate affinity) have not yet been performed. However, this should become possible in the near future as we have found a crystallization condition with an uncharged precipitant. If still unsuccesful, site-directed mutagenesis (targeted at potential T state specific interactions) may be used to engineer a version of the enzyme ‘locked’ in the R state. The complementary ATP complex is expected to be more easily obtained and would clarify if activation by CTP and ATP are caused by the same mechanism. Several mutants can be proposed based on our identification of potentially thermostabilizing features, e.g. a mutant carrying a truncated C terminus at the regulatory chain or mutants with disrupted ion pair networks such as the extensive ion pair network involving the novel C1-R2 type interface.

Moritella profunda ATCase

Still relatively few psychrophilic protein structures have been solved by X-ray crystallography and few enzymes are characterized across the whole physiological temperature range. The crystal structure of the *M. profunda* ATCase (at 2.85 Å) thus provides a rare insight into the adaptation to cold of this family of enzymes, even more so because the ATCase is a model for allosteric regulation. The functional characterization (to our knowledge the first one performed on a purified cold-adapted ATCase) reveals similar homotropic behaviour, yet differences in heterotropic regulation. The structure can be described as an extreme T state compared to the *E. coli* T state ATCase. Structural features were identified that may explain the decreased stability and increased cold activity: weakened interfaces, the loss of conformational restraints at the 240s loop region and allosteric-zinc interdomain interface, and an increase in net negative surface charge (contrasting with the hyperthermophilic ATCases).

As for the *S. acidocaldarius* enzyme, analysis of sedimentation velocity experiments with the bisubstrate analogue PALA seems to confirm that the unliganded enzyme is indeed in a structural T state and that a conformational transition takes place similar in amplitude to *E. coli* ATCase. A first analysis of the ATCase in complex with ATP (at 3.1 Å resolution) indicates rather similar modes of binding (resembling the ATP binding mode of *E. coli* R state ATCase) and the absence of a global structural change.

Clearly, the crystal structure of the *M. profunda* ATCase R state would provide a wealth of information on the catalytic and allosteric properties of the enzyme. However, the crystallization condition of the *M. profunda* ATCase also uses ammonium sulphate as a precipitant, which may explain why direct soaking and cocrystallization with PALA at different experimental settings have been unsuccessful to date. No alternative crystal forms in the presence of PALA could be obtained, yet a low-salt condition that yielded small needles was found in the absence of PALA and could potentially be optimized for soaking experiments. A higher quality and higher resolution data set for the ATP complex and a crystal structure of the CTP complex would give more insight into the aberrant heterotropic regulation. More conclusive kinetic studies are required to discriminate between competitive and heterotropic ATP inhibition and to investigate the influence of pressure on activity, stability and regulation. Studies (preferably including calorimetry) on site-specific mutants (e.g. mutating glycine 246 of the 240s loop region back to glutamine), selected on basis of the crystal structure, furthermore would allow confirmation of essential adaptive features.

Pseudoalteromonas haloplanktis xylanase

Xylanases are usually grouped into glycoside hydrolase families 10 and 11. The *P. haloplanktis* family 8 xylanase pXyl is novel among xylanases, displaying a novel mechanism of xylan hydrolysis and a novel three-dimensional structure. The number of studies of the catalytic mechanism and catalytic residues of family 8 enzymes is minimal. By the use of site-directed mutagenesis and a comparative characterization of the kinetic parameters, pH dependency of activity and thermal stability of mutant and wild-type enzymes in association with crystallographic analysis, we have been able to delineate the functions of several active site residues in this enzyme (e.g. the general acid and general base catalyst). Determination of the crystal structures of inactive mutants in complex with substrate and product has allowed for the demarcation of subsites –3 to +3 and has given insight into the basis of xylose/glucose discrimination in family 8 enzymes. Furthermore, we have demonstrated the existence of an induced fit mechanism upon ligand binding whereby a number of conformational changes (possibly linked to the increased flexibility of the cold enzyme), and in particular a

repositioning of the catalytic proton donor occurs. Finally, based on these novel data we have presented an elaboration on the conformational itinerary of family 8 enzymes.

Future work should include characterization of the enzyme mutated at Arg 284, a strictly conserved residue in family 8 with a largely unappreciated role. Our analysis indicates a potential role in transition state stabilization and in mediating structural rearrangements during the catalytic cycle. The relevance of the repositioning of the general acid (Glu 78) during catalysis could be investigated by preventing its movement, e.g. replacing residue Ala 142 (or Gly 145) with a bulkier side chain. It should also be verified whether Asp 144 is able to stabilize the substrate xylose at subsite –1 in a boat conformation, by solving the crystal structure of the substrate complex of pXyl mutated at a different position, e.g. at residue 284. Finally, X-ray crystallography could be used to study the structural basis for binding mixed linkage xylan (seaweed is expected to be the main source of xylan in its natural environment) or decorated substrates.

References

Abe, F., Kato, C. and Horikoshi, K. (1999) Pressure-regulated metabolism in microorganisms. *Trends Microbiol.* 7, 447-453.

Ahonkhai, I., Kamekura, M. and Kushner, D. J. (1989) Effects of salts on the aspartate transcarbamylase of a halophilic eubacterium, *Vibrio costicola. Biochem. Cell Biol.* 67, 666-669.

Albery, W. J. and Knowles, J. R. (1976) Evolution of enzyme function and the development of catalytic effeciency. *Biochemistry,* 15, 5631-5640.

Allewell, N. M. (1989) *Escherichia coli* aspartate carbamoylase: structure, energetics, and catalytic and regulatory mechanisms. *Annu. Rev. Biophys. Biophys. Chem,* 18, 71-92.

Argos, P., Rossmann, M. G., Grau, U. M., Zuber, H., Frank, G. and Tratschin, J. D. (1979) Thermal stability and protein structure. *Biochemistry,* 18, 5698-5703.

Aucoin, J. M., Pishko, E. J., Baker, D. P. and Kantrowitz, E. R. (1996) Engineered complementation in *Escherichia coli* aspartate transcarbamylase. Heterotropic regulation by quaternary structure stabilization. *J. Biol. Chem.* 271, 29865-29869.
Auerbach, G., Ostendorp, R., Prade, L., Korndörfer, I., Dams, T., Huber, R. and Jaenicke, R. (1998) Lactate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*: the crystal structure at 2.1 Å resolution reveals strategies for intrinsic protein stabilization. *Structure*, 6, 769-781.

Baillon, J., Tauc, P. and Hervé, G. (1985) A noncooperative substrate for *Escherichia coli* aspartate transcarbamylase. *Biochemistry,* 24, 7182-7187.

Baker, E. N. and Hubbard, R. E. (1984) Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.* 44, 97-179.

Baldwin, R. L. (1986) Temperature dependence of the hydrophobic interaction in protein folding. *Proc. Natl Acad. Sci. USA,* 83, 8069-8072.

Barlow, D. J. and Thornton, J. M. (1983) Ion-pairs in proteins. *J. Mol. Biol.*168, 867-885.

Bartlett, D. H. (2002) Pressure effects on in vivo microbial processes. *Biochim. Biophys. Acta,* 1595, 367-381.

Bartlett, D. H. and Bidle, K. A. (1999) *Enigmatic microorganisms and life in extreme environments* (ed. Seckbach, J., Kluwer, Dordrecht), 503-512.

Beaucamp, N., Hofmann, A., Kellerer, B. and Jaenicke, R. (1997) Dissection of the gene of the bifunctional PGK-TIM fusion protein from the hyperthermophilic bacterium *Thermotoga maritima:* design and characterization of the separate triosephosphate isomerase. *Protein Sci.*6, 2159-2165.

Bec, N., Villa, A., Tortora, P., Mozhaev, V. V., Balny, C. and Lange, R. (1996) Enhanced stability of carboxypeptidase from Sulfolobus solfataricus at high pressure. *Biotechnol. Lett.* 18, 483-488.

Becktel, W. J. and Schellman, J.A. (1987) Protein stability curves. *Biopol.* 26, 1859-1877.

Beg, Q. K., Kapoor, M., Mahajan, L. and Hoondal, G. S. (2001) Microbial xylanases and their industrial applications: a review. *Appl. Microbiol. Biotechnol.* 56, 326-338.

Belancic, A., Scarpa, J., Peirano, A., Dias, R., Steiner, J. and Ezaguirre, J; (1995) *Penicillium purpurogenum* produces several xylanases: purification of and properties of two of the enzymes. *J. Biotechnol.* 41, 71-79.

Bell, G. S., Russell, R. J., Connaris H., Hough, D. W., Danson, M. J., and Taylor, G. L. (2002) Stepwise adaptations of citrate synthase to survival at life's extremes. From psychrophile to hyperthermophile. *Eur. J. Biochem.* 269, 6250-6260.

Bentahir, M., Geller, G., Aittaleb, M., Lamotte-Brasseur, J., Himri, T., Chessa, J. P. and Gerday, C. (2000) Structural, kinetic, and calorimetric characterization of the cold-active phosphoglycerate kinase from the antarctic *Pseudomonas sp.* TACII18. *J. Biol. Chem.* 275, 11147-11153.

Berezovsky, I. N. and Shakhnovich, E. I. (2005) Physics and evolution of thermophilic adaptation. *Proc. Natl Acad. Sci.USA,* 102, 12742-12747.

Biely, P. (1985) Microbial xylanolytic systems, *Trends Biotechnol. 3*, 286–290.

Blochl, E. et al. (1997) *Pyrolobus fumarii,* gen. and sp. Nov, represents a novel group of archaea, extending the upper temperature limit for life to 113°C. *Extremophiles*, 1, 14-21.

Bönisch, H., Backmann, J., Kath, T., Naumann, D. and Schäfer, G. (1996) Adenylate kinase from *Sulfolobus acidocaldarius:* expression in *Escherichia coli* and characterization by Fourier transform infrared spectroscopy. *Arch. Biochem. Biophys.* 333, 75-84.

Brabson, J. S. and Switzer, R. L. (1975) Purification and properties of *Bacillus subtilis* aspartate carbamoyltransferase. *J. Biol. Chem.* 250, 8664-8669.

Brock, T. D., Brock, K. M., Belly, R. T. and Weiss, R. L. (1972) *Sulfolobus*: a new genus of sulphur-oxidizing bacteria living at low pH and high temperature. *Arch. Microbiol.* 84, 54-68.

Burley, S. K. and Petsko, G. A. (1985) Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science,* 229, 23-28.

Cambillau, C. and Claverie J.-M. (2000) Structural and genomic correlates of hyperthermostability. *J. Biol. Chem.* 275, 32383-32386.

Chan, M. K., Mukund, S., Kletzin, A., Adams, M. W. W. and Rees, D. C. (1995) Structure of a hyperthermophilic tungstopterin enzyme, aldehyde ferredoxin oxidoreductase. *Science*, 267, 1463-1469.

Chen, P., Van Vliet, F., Van de Casteele, M., Legrain, C., Cunin, R. and Glansdorff, N. (1998) Aspartate transcarbamylase from the hyperthermophilic eubacterium *Thermotoga maritima:* fused catalytic and regulatory polypeptides form an allosteric enzyme. *J. Bacteriol*. 180, 6389-6391.

Cherry, J. R., Lamsa, M. H., Schneider, P., Vind, J., Svendsen, A., Jones, A. and Pedersen, A. H. (1999) Directed evolution of a fungal peroxidase. *Nature Biotechnol.* 17, 379-384.

Chi, Y. I., Martinez-Cruz, L. A., Jancarik, J., Swanson, R. V., Robertson, D. E. and Kim, S. H. (1999) Crystal structure of the β-glycosidase from the hyperthermophile *Thermosphaera aggregans*: insights into its activity and thermostability. *FEBS Lett.* 445, 375-383.

Chilukuri, L. N. and Bartlett, D. H. (1997) Isolation and characterization of the gene encoding single-stranded-DNA-binding protein (SSB) from four marine *Shewanella* strains that differ in their temperature and pressure optima for growth. *Microbiology,* 143, 1163-1174.

Chyba, C. F. and Phillips, C. B. (2001) Possible ecosystems and the search for life on Europa. *Proc. Natl Acad. Sci.USA,* 98, 801-804.

Clarke, A. (2003) Evolution on planet Earth: the impact of the physical environment (eds. Rothschild, L. and Lister, A., Academic, London)

Collins, K. D. (2004) Ions from the Hofmeister series and osmolytes: effects on proteins in solution and in the crystallization process. *Methods,* 34, 300-311.

Collins, T., Gerday, C. and Feller, G. (2005) Xylanases, xylanase families and extremophilic xylanases. *FEMS Microbiol. Rev.* 29, 3-23.

Collins, T., Meuwis, M.-A., Gerday, C. and Feller, G. (2003) Activity, stability and flexibility in glycosidases adapted to extreme thermal environments. *J. Mol. Biol.* 328, 419-428.

Collins, T., Meuwis, M. A., Stals, I., Claeyssens, M., Feller, G. & Gerday, C. (2002) A novel family 8 Xylanase: Functional and physico-chemical characterization. *J. Biol. Chem.* 277, 35133-35139.

Cooper, A. (1976) Thermodynamic fluctuations in protein molecules. *Proc. Natl Acad. Sci. USA,* 73, 2740-2741.

Cornish-Bowden, A. (1976) The effect of natural selection on enzymic catalysis. *J. Mol. Biol.* 101, 1-9.

Coutinho, P. M. & Henrissat, B. (1999). Carbohydrate-active enzyme server (CAZY) at URL: http://afmb.cnrs-mrs.fr/~cazy/CAZY/.

Dams, T., Auerbach, G., Bader, G., Jacob, U., Ploom, T., Huber, R. and Jaenicke, R. (2000) The crystal structure of dihydrofolate reductase from *Thermotoga maritima*: molecular features of thermostability. *J. Mol. Biol.* 297, 659-672.

Daniel, R. M., Dunn, R. V., Finney, J. L. and Smith, J. C. (2003) The role of dynamics in enzyme activity. *Annu. Rev. Biophys. B iomol. Struct.* 32, 69-92.

De Staercke, C., Van Vliet, F., Xi, X.-G., Rani, C.X., Ladjimi, M., Jacobs, A., Triniolles, F., Hervé, G. and Cunin, R. (1995) Intramolecular transmission of the ATP regulatory signal in *Escherichia coli* aspartate transcarbamylase: specific involvement of a clustered set of amino

acid interactions at an interface between regulatory and catalytic subunits. *J. Mol. Biol.* 246, 132-143.

De Vos, D., Van Petegem, F., Remaut, H., Legrain, C., Glansdorff, N. and Van Beeumen, J. J. (2004) Crystal structure of T State aspartate carbamoyltransferase of the hyperthermophilic archaeon *Sulfolobus acidocaldarius*. *J. Mol. Biol.* 339, 887–900.

De Vos, D., Hulpiau, P., Vergauwen, B., Savvides, S. N. and Van Beeumen, J. (2005) Expression, purification, crystallization and preliminary X-ray crystallographic studies of a cold-adapted aspartate carbamoyltransferase from *Moritella profunda*. *Acta Crystall*. F61, 279-281.

Di Giulio, M. (2004) A comparison of proteins from *Pyrococcus furiosus* and *Pyrococcus abyssi*: barophily in the physicochemical properties of amino acids and in the genetic code. *Gene*, 346, 1-6.

Dill, K. A. (1990) Dominant forces in protein folding. *Biochemistry* 29, 7133-7155.

Dougherty, D. A. (1996) Cation-π interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. *Science,* 271, 163-168.

Durbecq, V., Thia-Toong, T.-L., Charlier, D., Villeret, V., Roovers, M., Wattiez, R., Legrain, C. and Glansdorff, N. (1999) Aspartate carbamoyltransferase from the thermoacidophilic arschaeon *Sulfolobus acidocaldarius.* Cloning, sequence analysis, enzyme purification and characterization. *Eur. J. Biochem.* 264, 233-241.

Dykhuizen, D. E., Dean, A. M. and Hartl, D. L. (1987) Metabolic flux and fitness. *Genetics*, 115, 25-31.

Edwards, K. J., Bond, P. L., Gihring, T. M. and Banfield, J. F. (2000) An archaeal iron-oxidizing extreme acidophile important in acid mine drainage. *Science,* 287, 1796-1799.

Eijsink, V. G., Vriend, G., Hardy, F., Veltman, O. R., Van der Vinne, B., Van der Burg, B., Dijkstra, B. W., Van der Zee, J. R. and Venema, G. (1993) Structural determinants of the thermostability of thermolysin-like *Bacillus* neutral proteases, p91-99. (in Stability and stabilization of enzymes; eds. W. J. Van den Tweel, A. Harder and R. M. Buitelaar, Elservier Science Publishers BV, Amsterdam, The Netherlands)

Edsall, J. T. (1935) Apparent molal heat capacities of amino acids and other organic compounds. *J. Am. Chem. Soc.* 57, 1506-1507.

Elcock, A. H. (1998) The stability of salt bridges at high temperatures: implications for hyperthermophilic proteins. *J. Mol. Biol.* 284, 489-502.

Elcock, E. H. and McCammon, J. A. (1998) Electrostatic contributions to the stability of halophilic proteins. *J. Mol. Biol.,* 280, 731-748.

Elegir, G., Szakacs, M. and Jeffries, T. W. (1994) Purification, characterization and substrate specificities of multiple xylanases from *Streptomyces* sp. strain B-12-2. *Appl. Environ. Microbiol.* 60, 2609-2615.

England, P. and Hervé, G. (1994) Involvement of the γ-phosphate of UTP in the synergistic inhibition fo *Escherichia coli* aspartate transcarbamoylase by CTP and UTP. *Biochemistry,* 33, 3913-3918.

Enns, C. A. and Chan, W. W.-C. (1978) Stabilization of the relaxed state of aspartate transcarbomoylase by modification with a bifunctional reagent. *J. Biol. Chem.* 253, 2511-2513.

Eriksson, A. E., Baase, W. A., Zhang, X.-J., Heinz, D. W., Blaber, M., Baldwin, E. P. and Matthews, B. W. (1992) Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science,* 255, 178-183.

Facchiano, A. M., Colonna, G. and Ragone, R. (1998) Helix stabiizing factors and stabilization of thermophilic proteins: an X-ray based study. *Protein Eng.* 11, 753-760.

Fell, D. (1997) Understanding the control of metabolism (Portland Press).

Feller, G. (2003) Molecular adaptations to cold in psychrophilic enzymes. *Cell. Mol. Life Sci.* 60, 648-662.

Feller, G. and Gerday, C. (2003) Psychrophilic enzymes: hot topics in cold adaptation. *Nature,* 1, 200-208.

Feller, G., D'Amico, D. and Gerday, C. (1999) Thermodynamic stability of a cold-active α-amylase from the Antarctic bacterium *Alteromonas haloplanctis. Biochemistry,* 38, 4613-4619.

Fetler, L., Tauc, P., Hervé, G., Moody, M. F. and Vachette, P. (1995) X-ray scattering titration of the quaternary structure transition of aspartate transcarbamylase with a bisubstrate analogue: influence of nucleotide effectors. *J. Mol. Biol.* 251, 243-255.

Fields, P. A. and Somero, G. N. (1998) Hot spots in cold adaptation: localized increases in conformational flexibility in lactate dehydrogenase A_4 orthologs of Antarctic notothenioid fishes. *Proc. Natl Acad. Sci. USA,* 95, 11476-11481.

Foote, J., Lauritzen, A. M. and Lipscomb, W. N. (1985) Substrate specificity in aspartate transcarbamylase; interaction of the enzyme with analogs of aspartate and succinate. *J. Biol. Chem.* 260, 9624-9629.

Forst, C. V. and Schulten, K. (1999) Evolution of metabolisms: a new method for the comparison of metabolic pathways using genomics information. *J. Comp. Biol.* 6, 343-360.

Freire, E. (2000) Can allosteric regulation be predicted from structure? *Proc. Natl Acad Sci. USA,* 97, 11680-11682.

Freye, K. J., Perman, C. S. and Royer, C. A. (1996) Testing the correlation between delta A and delta V of protein unfolding using m value mutants of staphylococcal nuclease. *Biochemistry,* 35, 10234-10239.

Gekko, K. and Hadegawa, Y. (1986) Compressibility – structure relationship of globular proteins. *Biochemistry,* 25, 6563-6571.

Gerday, C. Aittaleb, M., Arpigny, J. L., Baise, E., Chessa, J. P., Garsoux, G., Petrescu, I. and Feller, G. (1997) Psychrophilic enzymes: a thermodynamic challenge. *Biochim. Biophys. Acta,* 1342, 119-131.

Gerhart, J. C. and Pardee, A. B. (1962) Enzymology of control by feedback inhibition. *J. Biol. Chem.* 237, 891-896.

Gerhart, J. C. and Schachman, H. K. (1968) Allosteric interactions in aspartate transcarbamylase. II. Evidence for different conformational states of the protein in the presence and absence of specific ligands. *Biochemistry,* 7, 538-552.

Gershenson, A., Schauerte, J. A., Giver, L. and Arnold, F. H. (2000) Tryptophan fluorescence study of enzyme flexibility and unfolding in laboratory-evolved thermostable esterases. *Biochemistry,* 39, 4658-4665.

Giver, L., Gershenson, A., Freskgard, P. O. and Arnold, F. H. (1998) Directed evolution of a thermostable esterase. *Proc. Natl Acad. Sci. USA,* 95, 12809-12813.

Gouaux, J. E. and Lipscomb, W. N. (1990) Crystal structure of phosphonoacetamide ligated T and Phosphonoacetamide and malonate ligated R states of aspartate carbamoyltransferase at 2.8 Å resolution and neutral pH. *Biochemistry,* 29, 389-402.

Gouaux, J. E., Stevens, R. C. and Lipscomb, W. N. (1990) Crystal structures of aspartate carbamoyltransferase ligated with phosphonoacetamide, malonate, and CTP or ATP at 2.8 Å resolution and neutral pH. *Biochemistry,* 29, 7702-7715.

Grabarse, W., Vaupel, M., Vorholt, J. A., Shima, S., Thauer, R. K., Wittershagen, A., Bourenkov, G., Bartunik, H. D. and Ermler, U. (1999) The crystal structure of methenyltetrahydromethanopterin cyclohydrolase from the hyperthermophilic archaeon *Methanopyrus kandleri. Struct. Fold. Des.* 7, 1257-1268.

Graziano, G., Catanzano, F., Riccio, A. and Barone, G. (1997) A reassessment of the molecular origin of cold denaturation. *J. Biochem.* 122, 395-401.

Gross, M. and Jaenicke, R. (1992) High pressure and biotechnology. *Colloq.Inserm.* 224, 83-87.

Gross, M. and Jaenicke, R. (1994) Proteins under pressure. The influence of high hydrostatic pressure on structure, function and assembly of proteins and protein complexes. *Eur. J. Biochem.* 221, 617-630.

Hack, E. S., Voroyova, T., Saksh, J. B., West, J. M., Macol, C. P., Hervé, G., Williams, M. K. and Kantrowitz, E. R. (2000) Characterization of the aspartate transcarbamoylase from *Methanococcus jannaschii. J. Biol. Chem.* 275, 15820-15827.

Harpaz, Y., Gerstein, M. and Chothia, C. (1994) Volume changes on protein folding. *Structure* 2, 641-649.

Hei, D. J. and Clark, D. S. (1994) Pressure stabilization of proteins from extreme thermophiles. *Appl. Environ. Microbiol.* 60, 932-939.

Helmstaedt, K., Krappmann, S. and Braus, G. H. (2001) Allosteric regulation of catalytic activity: *Escherichia coli* aspartate transcarbamoylase versus yeast chorismate mutase. *Microbiol. & Mol. Biol. Rev.* 65, 404-421.

Hendsch, Z. S. and Tidor, B. (1994) Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Protein Sci.* 3, 211-226.

Hervé, G. (1989) Aspartate transcarbamylase from *Escherichia coli*, in Allosteric enzymes (eds. Hervé, G., CRC Press, Boca Raton, FL, pp. 61-79)

Hervé, G., Moody, M. F., Tauc, P., Vachette, P. and Jones, P. T. (1985) Quaternary structure changes in aspartate transcarbamylase studied by X-ray solution scattering. Signal transmission following effector binding. *J. Mol. Biol.* 185, 189-199.

Hess, D., Kruger, K., Knappik, A., Palm, P. and Hensel, R. (1995) Dimeric 3-phosphoglycerate kinases from hyperthermophilic archaea. Cloning, sequencing and expression of the 3-phosphoglycerate kinase gene of *Pyrococcus woesei* in *Escherichia coli* and characterization of the protein. Structural and functional comparison with the 3-phosphoglycerate kinase of *Methanothermus fervidus*. *Eur. J. Biochem.* 233, 227-237.

Hilser, V. J. and Freire, E. (1996) Structure-based calculation of the equilibrium folding pathway of proteins. Correlation with hydrogen exchange protection factors. *J. Mol. Biol.* 263, 756-772.

Hollien, J. and Marqusee, S. (1999) A thermodynamic comparison of mesophilic and thermophilic ribonucleases H. *Biochemistry,* 38, 3831-3836.

Honzatko, R. B., Crawford, J. L., Monaco, H. L., Ladner, J. E., Edwards, B. F. P., Evans, D., Warren, G., Wiley, D. C., Ladner, R. C. and Lipscomb, W. N. (1982) Crystal and molecular structures of native and CTP-liganded aspartate transcarbamylase in the crystalline state. *J. Mol. Biol.* 160, 219-263.

Honzatko, R. B., Monaco, H. L. and Lipscomb, W. N. (1979) A 3.0 Å resolution study of nucleotide complexes with aspartate carbamoyltransferase. *Proc. Natl Acad. Sci. USA,* 76, 5105-5109.

Hoyoux, A., Jennes, I., Dubois, P., Genicot, S., Dubail, F., Francois, J. M., Baise, E., Feller, G. and Gerday, C. (2001) Cold-adapted beta-galactosidase from the Antarctic psychrophile *Pseudoalteromonas haloplanktis*. *Appl. Environ. Micorbiol.* 67, 1529-1535.

Huang, J. and Lipscomb, W. N. (2004) Products in the T-state of aspartate transcarbamylase: crystal structure of the phosphate and N-carbamyl-L-aspartate ligated enzyme. *Biochemistry,* 43, 6422-6426.

Huber, R. (1988) Flexibility and rigidity of proteins and protein-pigment complexes. *Angew. Chem. Int. Ed. Engl.* 27, 79-88.

Hummer, G., Garde, S., Garcia, A. E. and Paulaitis, M. E. (1998) The pressure dependence of hydrophobic interactions is consistent with the observed pressure denaturation of proteins. *Proc. Natl Acad. Sci. USA,* 95, 1552-1555.

Ichikawa, J. K. and Clarke, S. (1998) A highly active protein repair enzyme from an extreme thermophile: the L-isoaspartyl methyltransferase from *Thermotoga maritima. Arch. Biochem. Biophys.* 358, 222-231.

Ishikawa, K., Okumura, M., Katayanagi, K., Kimura, S., Kanaya, S., Nakamura, H. and Morikawa, K. (1993) Crystal structure of ribonuclease H from *Thermus thermophilus* HB8 refined at 2.8 Å resolution. *J. Mol. Biol.* 230, 529-542.

Jaenicke R. (1981) Enzymes under extremes of physical conditions. *Annu. Rev. Biophys. Bioeng.* 10, 1-67.

Jaenicke R. (1987) in Current perspectives in high pressure biology (Jannasch, H. W. et al. Eds) pp. 257-272, Academic Press, London.

Jaenicke R. (1991) Protein stability and molecular adaptation to extreme conditions. *Eur. J. Biochem.* 202, 715-728.

Jaenicke R. (2000) Do ultrastable proteins from hyperthermophiles have high or low conformational rigidity? *Proc. Natl Acad. Sci. USA,* 97, 2962-2964.

Jaenicke, R. (2000) Stability and stabilization of globular proteins in solution. *J. Biotech.,* 79, 193-203.

Jaenicke R. and Böhm, G. (1998) The stability of proteins in extreme environments. *Curr. Opin. Struct. Biol.* 8, 738-748.

Jones, M. E., Spector, L. and Lipmann, F. (1955) Carbamyl phosphate. The carbamyl donor in enzymatic citrulline synthesis. *J. Am. Chem. Soc.* 77, 819-820.

Kantrowitz, E. R. and Lipscomb, W. N. (1988) *Escherichia coli* aspartate transcarbamoylase: the relation between structure and function. *Science,* 241, 669-674.

Kantrowitz, E. R., Reed, H. W., Ferraro, R. A. and Daigneault, J. P. (1981) Analysis of mutant *Escherichia coli* aspartate transcarbamylases isolated from a series of suppressed pyrB nonsense strains. *J. Mol. Biol.* 153, 569-587.

Karplus, M. and Kuriyan, J. (2005) Molecular dynamics and protein function. *Proc. Natl Acad. Sci. USA,* 102, 6679-6685.

Karschikoff, A. and Ladenstein, R. (1998) Proteins from thermophilic and mesophilic organisms essentially do not differ in packing. *Protein Eng.* 11, 867-872.

Kato, C. et al. (1998) Extremely barophilic bacteria isolated from the Mariana trench, Challenger deep, at a depth of 11,000 meters. *Appl. Environ. Microbiol.* 64, 1510-1513.

Kauzmann, W. (1959) Some factors in the interpretation of protein denaturation. *Adv. Protein Chem.* 14, 1-63.

Kauzmann, W. (1987) Thermodynamics of unfolding. *Nature,* 325, 763-764.

Kawamura, S., Kakuta, Y., Tanaka, I., Hikichi, K., Kuhara, S., Yamasaki, N. and Kimura, M. (1996) Glycine-15 in the bend between two α-helices can explain the thermostability of DNA binding protein HU from *Bacillus stearothermophilus. Biochemistry,* 35, 1195-1200.

Ke, H.-M., Honzatko, R. B. and Lipscomb, W. N. (1984) Structure of unligated aspartate carbamoyltransferase of *Escherichia coli* at 2.6 Å resolution. *Proc. Natl Acad. Sci. USA,* 81, 4037-4040.

Ke, H.-M., Lipscomb, W. N., Cho, Y. and Honzatko, R. B. (1988) Complex of *N*-phosphonacetyl-L-aspartate with aspartate carbamoyltransferase: a molecular modeling study. *J. Mol. Biol.* 204, 725-747.

Kim, K. H., Pan, Z., Honzatko, R. B., Ke, H.-M. and Lipscomb, W. N. (1987) Structural asymmetry in the CTP-liganded form of aspartate carbamoyltransferase from *Escherichia coli. J. Mol. Biol.*196, 853-875.

Kimura, S., Kanaya, S. and Nakamura, H. (1992) Thermostabilization of *Escherichia coli* ribonuclease HI by replacing left-handed helical Lys^{95} with Gly or Asn. *J. Biol. Chem.* 267, 22014-22017.

Kirino, H., Aoki, M., Aoshima, M., Hayashi, Y., Ohba, M., Yamagishi, A., Wakagi, T. and Oshima, T. (1994) Hydrophobic interaction at the subunit interface contributes to the thermostability of 3-isopropylmalate dehydrogenase from an extreme thermophile, *Thermus thermophilus. Eur. J. Biochem.* 220, 275-281.

Konisky, J., Michels, P. C. and Clark, D. S. (1995) Pressure stabilization is not a general property of thermophilic enzymes: the adenylate kinases of *Methanococcus voltae, Methanococcus maripludis, Methanococcus thermolithotrophicus,* and *Methanococcus jannaschii. Appl. Environm. Microbiol.* 61, 2762-2764.

Koshland, D. E. Jr., Némethy, G. and Filmer, D. (1966) Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry,* 365-385.

Krause, K. L., Volz, K. W. and Lipscomb, W. N. (1985) Structure at 2.9 Å resolution of aspartate carbamoyltransferase complexed with the bisubstrate analogue *N*-(phoshonacetyl)-L-aspartate. *Proc. Natl Acad. Sci. USA,* 82, 1643-1647.

Krause, K. L., Volz, K. W. and Lipscomb, W. N. (1987) 2.5 Å structure of aspartate carbamoyltransferase complexed with the bisubstrate analog *N*-(phoshonacetyl)-L-aspartate. *J. Mol. Biol.* 193, 527-553.

Kulkarni, N., Shendye, A., and Rao, M. (1999) Molecular and biotechnological aspects of xylanases. *FEMS Microbiol. Rev.* 23, 411–456.

Kumar S. and Nussinov R. (2004) Different roles of electrostatics in heat and in cold: adaptation by citrate synthase. *Chem. Biochem.* 5, 280-290.

Kumar, S., Tsai, C.-J. and Nussinov, R. (2000) Factors enhancing protein thermostability. *Prot. Engin.* 13, 179-191.

Kumar, S., Tsai, C.-J. and Nussinov, R. (2001) Thermodynamic differences among homologous thermophilic and mesophilic proteins. *Biochemistry* 40, 14152-14165.

Kurzynsi, M., Palacz, K. and Chelminiak, P. (1998) Time course of reactions controlled and gated by intramolecular dynamics of proteins: predictions of the model of random walk on fractal lattices. *Proc. Natl Acad. Sci. USA,* 95, 11685-11690.

Labedan, B., Xu, Y., Naumoff, D. G. and Glansdorff, N. (2004) Using quaternary structures to assess the evolutionary history of proteins: the case of the aspartate carbamoyltransferase. *Mol. Biol. Evol.* 21, 364-373.

Ladner, J. E., Kitchell, J. P., Honzatko, R. B., Ke, H.-M., Volz, K. W., Kalb, A. J., Ladner, R. C. and

Lipscomb, W. N. (1982) Gross quaternary changes in aspartate carbamoyltransferase are induced by the binding of *N*-phosphonacetyl)-L-aspartate: a 3.5 Å resolution study. *Proc. Natl Acad. Sci. USA,* 79, 3125-3128.

Lazaridis, T., Lee, I. and Karplus, M. (1997) Dynamics and unfolding pathways of a hyperthermophilic and a mesophilic rubredoxin. *Protein Sci.* 6, 2589-2605.

Lebbink, J. H. G., Knapp, S., van der Oost, J., Rice, D., Ladenstein, R. and deVos, W. M. (1999) Engineering activity and stability of *Thermotoga maritima* glutamate dehydrogenase. II: construction of a 16-residue io*N*-pair network at the subunit interface. *J. Mol. Biol.* 289, 357-369.

Lee, C.-F., Allen, M. D., Bycroft, M. and Wong, K.-B. (2005) Electrostatic interactions contribute to reduced heat capacity change of unfolding in a thermophilic ribosomal protein L30e. *J. Mol. Biol.* 348, 419-431.

Levy, M. and Miller, S. L. (1998) The stability of the RNA bases: implications for the origin of life. *Proc. Natl Acad. Sci. USA,* 95, 7933-7938.

Liebl, V., Kaplan, J. G. and Kushner, D. J. (1969) Regulation of a salt-dependent enzyme: the aspartate transcarbamylase of an extreme halophile. *Canad. J. Biochem.* 47, 1095-1097.

Li, W., Grayling, R. A., Sandman, K., Edmondson, S., Shiver, J. W. and Reeve, J. N. (1998) Thermodynamic stability of archaeal histones. *Biochemistry*, 37, 10563-10572.

LiCata, V. J. and Allewell, N. M. (1998) Solvent perturbation of the allosteric regulation of aspartate transcarbamylase. *Biochim. Biophys. Acta,* 1384, 306-314.

Lippert, K. and Galinski, E. A. (1992) Enzyme stabilization by ectoi*N*-type compatible solute interaction, protection against heating and drying. *Appl. Microbiol. Biotech.* 37, 61-65.

Lipscomb, W. N. (1994) Aspartate transcarbamylase from *Escherichia coli*: activity and regulation. *Adv. Enzymol. Relat. Areas Mol. Biol.* 68, 67-151.

Livingstone, J. R., Spolar, R. S. and Record, M. T. Jr. (1991) Contribution to the thermodynamics of protein folding from the reduction in wateraccessible nonpolar surface area. *Biochemistry,* 30, 4237-4244.

Lonhienne, T., Gerday, C. and Feller, G. (2000) Psychrophilic enzymes: revisiting the thermodynamic parameters of activation may explain local flexibility. *Biochim. Biophys. Acta,* 1543, 1-10.

Lonhienne, T., Zoidakis, J., Vorgias, C. E., Feller, G., Gerday, C. and Bouriotis, V. (2001) Modular structure, local flexibility and cold-activity of a novel chitobiase from a psychrophilic Antarctic bacterium. *J. Mol. Biol.* 310, 291-297.

Low, P. S., Bada, J. L. and Somero, G. N. (1973) Temperature adaptation of enzymes: roles of the free energy, the enthalpy, and the entropy of activation. *Proc. Natl Acad. Sci.USA,* 70,430-432.

Luque, I., Leavitt, S. and Freire, E. (2002) The linkage between protein folding and functional cooperativity. *Annu. Rev. Biophys. Biomol. Struct.,* 31, 235-256.

Manco, G., Giosuè, E., D'Auria, S., Herman, P., Carrea, G. and Rossi, M. (2000) Cloning, overexpression, and properties of a new thermophilic and thermostable esterase with sequence similarity to hormone-sensitive lipase subfamily from the archaeon *Archaeoglobus fulgidus. Arch. Biochem. Biophys.* 373, 182-192.

Matsumura, M., Signor, G. and Matthews, B. W. (1989) Substantial increase of protein stability by multiple disulphide bonds. *Nature*, 342, 291-293.

Makhatadze, G. I. and Privalov, P. L. (1990) Heat capacity fo proteins. I. Partial molar heat capacity of individual amino acid residues in aqueous solution: hydration effect. *J. Mol. Biol.* 213, 375-384.

Makhatadze, G. I. and Privalov, P. L. (1995) Energetics of protein structure. *Adv. Prot. Chem.* 47, 307-425.

Massant, J. (2004) Molecular physiology of hyperthermophiles: metabolic channeling of carbamoyl phosphate, a thermolabile ant potentially toxic intermediate. *PhD thesis*

Matthews, B. W., Nicholson, H. and Becktel, W. J. (1987) Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. *Proc. Natl Acad. Sci.USA,* 84, 6663-6667.

Mc Lean, M. A., Maves, S. A., Weiss, K. E., Krepich, S. and Sligar, S. G. (1998) *Biochem. Biophys. Res. Commun.* 252, 166-172.

McCarter, J. D. and Withers, S. G. (1994) Mechanisms of enzymatic glycoside hydrolysis. *Curr. Opin. Struct. Biol.* 4, 885-892.

McDowell, R. H. (1967) Chemistry and enzymology of marine algal polysaccarides. *Academic Press, London,* 88-96, 134-137.

Melendez-Hevia, E., Waddell, T. G. and Montero, F. (1994) Optimization of metabolism: the evolution of metabolic pathways toward simplicity through the game of the pentose phosphate cycle. *J. theor. Biol.* 166, 201-220.

Merz, A., Knöchel, T., Jansonius, J. N. and Kirschner, K. (1999) The hyperthermostable indoleglycerol phosphate synthase from *Thermotoga maritima* is destabilized by mutational disruption of two solvent-exposed salt bridges. *J. Mol. Biol.* 288, 753-763.

Mevarech, M., Frolow, F. and Gloss, L.M. (2000) Halophilic enzymes: proteins with a grain of salt. *Biophys. Chem.,* 86, 155-164.

Mombelli, E., Shehi, E., Fusi, P. and Tortora, P. (2002) Exploring hyperthermophilic proteins under pressure: theoretical aspects and experimental findings. *Biochim. Biophys. Acta,* 1595, 392-396.

Monaco, H. L., Crawford, J. L. and Lipscomb, W. N. (1978) Three-dimensional structures of aspartate carbamoyltransferase from *Escherichia coli* and its complex with cytidine triphosphate. *Proc. Natl Acad. Sci. USA,* 75, 5276-5280.

Monod, J., Wyman, J. and Changeux, J.-P. (1965) On the nature of allosteric transitions: a plausible model. *J. Mol. Biol.* 12, 88-118.

Motono, C., Oshima, t. and Yamagishi, A. (2001) High thermal stability of 3-isopropylmalate dehydrogenase from *Thermus thermophilus* resulting from low ΔC(p) of unfolding. *Protein Eng.* 14, 961-966.

Mozhaev, V. V., Heremans, K., Frank, J., Masson, P. and Balny, C. (1994) Exploiting the effects of high hydrostatic pressure in biological applications. *Trends Biotechnol.* 12, 493-501.

Mozhaev, V. V., Heremans, K., Frank, J., Masson, P. and Balny, C. (1996) High pressure effects on protein structure and function. *Proteins,* 24, 81-91.

Murphy, K. P. and Freire, E. (1992) Thermodynamics of structural stability and cooperative folding behavior in proteins. *Advan. Protein Chem.* 43, 313-361.

Myers, J. K., Pace, C. N. and Scholtz, J. M. (1995) Denaturant *m* values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* 4, 2138-2148.

Nakai, T., Okada, K., Akutsu, S;, Miyahara, I., Kawaguchi, S., Kato, R., Kuramitsu, S. and Hirotsu, K. (1999) Structure of *Thermus thermophilus* HB8 aspartate aminotransferase and its complex with maleate. *Biochemistry,* 38, 2413-2424.

Newton, C. J. and Kantrowitz, E. R. (1990) The regulatory subunit of *Escherichia coli* aspartate carbamoyltransferase may influence homotropic cooperativity and heterotropic interactions by a direct interaction with the loop containing residues 230-245 of the catalytic chain. *Proc. Natl Acad. Sci. USA,* 87, 2309-2313.

Nojima, H., Ikai, A., Oshima, T. and Noda, H. (1977) Reversible thermal unfolding of thermostable phosphoglycerate kinase. Thermostability associated with mean zero enthalpy change. *J. Mol. Biol.* 116, 429-442.

Norberg, P., Kaplan, J. G. and Kushner, D. J. (1973) Kinetics and regulation of the salt-dependent aspartate transcarbamylase of *Halobacterium cutirubrum. J. Bacteriol.* 113, 680-686.

Nunn, J. R., Parolis, H., Russell, I. (1973) Polysaccarides of the red alga *Chaetangium erinaceum.* Part I. Isolation and characterization of the water-soluble xylan. *Carbohydr. Res.* 26, 169-180.

Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276, 307-326.

Ovadi, J. (1991) Physiological significance of metabolic channelling. *J. theor. Biol.* 152, 1-22.
Pace, C. N., Shirley, B. A., McNutt, M. and Gajiwala, K. (1996) Forces contributing to the conformational stability of proteins. *FASEB J.* 10, 75-83.

Pappenberger, G., Schurig, H. and Jaenicke, R. (1997) Disruption of an ionic network leads to accelerated thermal denaturation of D-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima. J. Mol. Biol.* 274, 676-683.

Privalov, P. L. and Gill, S. J. (1988) Stability of protein structure and hydrophobic interaction. *Adv. Protein Chem.* 39, 191-234.

Privalov, P. L. and Khechinashvili, N. N. (1974) A thermodynamic approach to the problem of stabilization of globular protein structure: a calorimetric study. *J. Mol. Biol.* 86, 665-684.

Privalov, P. L. and Makhatadze, G. I. (1990) Heat capacity of proteins. II. Partial molar heat capacity of the unfolded polyperptide chain of proteins: protein unfolding effects. *J. Mol. Biol.* 213, 385-391.

Privalov, P. L. and Makhatadze, G. I. (1993) Contribution of hydration to protein folding thermodynamics. II. The entropy and Gibbs energy of hydration. *J. Mol. Biol.* 232, 660-679.

Puls, J., Schmidt, O. and Granzow, C. (1987) Glucuronidase in two microbial xylanolytic systems. *Enzyme Microb. Technol.* 9, 83-88.

Purcarea, C. Ahuja, A., Lu, T., Kovari, L., Guy, H. I. And Evans, D. R. (2003) *Aquifex aeolicus* aspartate transcarbamoylase, an enzyme specialized for the efficient utilization of unstable carbamoyl phosphate at elevated temperature. *J. Biol. Chem.* 278, 52924-52934.

Purcarea, C., Erauso, G., Prieur, D. and Hervé, G. (1994) The catalytic and regulatory properties of aspartate transcarbmoylase from *Pyrococcus abyssi,*a new deep-sea hyperthermophilic archaeobacterium. *Microbiol.* 140, 1967-1975.

Rahman, R. N. Z. A., Fujiwara, S., Nakamura, H., Takagi, M. and Imanaka, T. (1998) Ion pairs involved in maintaining a thermostable structure of glutamate dehydrogenase from a hyperthermophilic archaeon. *Biochem. Biophys. Res. Commun.* 248, 920-926.

Record Jr., M. T., Anderson, C. F. and Lohman, T. M. (1978) Thermodynamics analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening and ion effects on water activity. *Q. Rev. Biophys.,* 11, 103-178.

Refaee, M., Tezuka, T., Akasaka, K. and Williamson, M. P. (2003) Pressure-dependent changes in the solution structure of hen egg-white lysozyme. *J. Mol. Biol.* 327, 857-865.

Reichard, P. and Hanshoff, G. (1956) Aspartate carbamyl transferase from *Escherichia coli. Acta Chem. Scand.* 10, 548-560.

Richards, F. M. (1979) Packing defects, cavities, volume flucturations and access to the interior of proteins. *Carlsberg Res. Commun.* 44, 47-63.

Robey, E. A., Schachman, H. K. (1985) Regeneration of active enzyme by formation of hybrids from inactive derivatives: implications for active sites shared between polypeptide chains of aspartate transcarbmoylase. *Proc. Natl Acad. Sci. USA,* 82, 361-365.

Robinson, G. W. and Cho, C. H. (1999) Role of hydration water in protein unfolding. *Biophys. J.* 77, 3311-3318.

Roche, O. and Field, M. J. (1999) Simulations of the T ↔ R conformational transition in aspartate transcarbamylase. *Prot. Eng.* 12, 285-295.

Roovers, M., Sanchez, R., Legrain, C. and Glansdorff, N. (2001) Experimental evolution of enzyme temperature activity profile: selection *in vivo* and characterization of low-temperature-adapted mutants of *Pyrococcus furiosus* ornithine carbamoyltransferase. *J. Bacteriol.* 183, 1101-1105.

Rothschild, L. J. and Mancinelli, R. L. (2001) Life in extreme environments. *Nature*, 409, 1092-1101.

Royer, C. A. (2002) Revisiting volume changes in pressure-induced protein unfolding. *Biochim. Biophys. Acta,* 1595, 201-209.

Royer, C. A., Hinck, A. P., Loh, S. N., Prehoba, K. E., Peng, X., Jonas, J. and Markley, J. L. (1993) Effects of amino acid substitutions on the pressure denaturation of staphylococcal nuclease as monitored by fluorescence and nuclear magnetic resonance spectroscopy. *Biochemistry,* 32, 5222.

Ruelle, P. and Kesselring, U. W. (1998) The hydrophobic effect. I. A consequence of the mobile order in H-bonded liquids. *J. Pharm. Sci.* 87, 987-997.

Russell, N. J. (2000) Toward a molecular understanding of cold activity of enzymes from psychrophiles. *Extremophiles,* 4, 83-90.

Russell, R. J. M., Ferguson, J. M. C., Haugh, D. W., Danson, M. J. and Taylor, G. L. (1997) The crystal structure of citrate synthase from the hyperthermophilic archaeon *Pyrococcus furiosus* at 1.9 Å resolution. *Biochemistry,* 36, 9983-9994.

Rye, C. S., and Withers, S. G. (2000) Glycosidase mechanisms, *Curr. Opin. Chem. Biol.* 4, 573-580.

Saito, R. and Nakayam, A. (2004) Differences in malate dehydrogenases from the obligately piezophilic deep-sea bacterium *Moritella* sp. strain 2D2 and the psychrophilic bacterium *Moritella* sp. strain 5710. *FEMS Microbiol. Lett.* 233, 165-172.

Sali, D., Bycroft, M. and Fersht, A. R. (1991) Surface electrostatic interactions contribute little of stability of barnase. *J. Mol. Biol.* 220, 779-788.

Schachman, H. K. (1988) Can a simple model account for the allosteric transition of aspartate transcarbamoylase. *J. Biol. Chem.* 263, 18583-18586.

Scharnagl, C., Reif, M. and Friedrich, J. (2005) Stability of proteins: temperature, pressure and the role of the solvent. *Biochim. Biophys. Acta,* 1749, 187-213.

Schurr, M. J., Vickrey, J. F., Kumar, A. P., Campbell, A. L., Cunin, R., Benhjiamin, R. C., Shanley, M. S. and O'Donovan, G. A. (1995) Aspartate transcarbamoylase genes of *Pseudomonas putida:* requirement for an inactive dihydrotase for assesmbly into the dodecameric homoenzyme. *J. Bacteriol.* 177, 1751-1759.

Sharp, K. A. and Honig, B. (1990) *Annu. Rev. Biophys. Biophys. Chem.* 19, 301-332.

Sheridan, P. P., Panasik, N., Coombs, J. M. and Brenchley, J. E. (2000) Approaches for deciphering the structural basis of low temperature enzyme activity. *Biochim. Biophys. Acta*, 1543, 417-433.

Shirley, B. A., Stanssens, P., Hanhn, U. and Pace, C. N. (1992) Contribution of hydrogen bonding to the conformational stability of ribonuclease T1. *Biochemistry* 31, 725-732.

Simmer, J. P., Kelly, R. E., Scully, J. L., Grayson, D. R., Ringer, A. G., Bergh, S. T. and Evans, D. R. (1989) Mammalian aspartate transcarbamylase (ATCase): sequence of the ATCase domain and interdomain linker in the CAD multifunctional polypeptide and properties of isolated domain. *Proc. Natl Acad. Sci. USA,* 86, 4382-4386.

Somero, G. N. (1983) Environmental adaptation of protein: strategies for the conservation of critical functional and structural traits. *Comp. Biochem. Physiol.* A 76, 621-633.

Somero, G. N. (1990) Life at low volume change: hydrostatic pressure as a selective factor in the aquatic environment. *Am. Zool.* 30, 123-135.

Somero, G. N. (1992) Adaptations to high hydrostatic pressure. *Annu. Rev. Physiol.* 54, 557-577.

Somero, G. N. (1995) Proteins and temperature. *Annu. Rev. Physiol.* 57, 43-68.

Souchon, H., Béguin, P. and Alzari P. M. (1996) Crystallization of a family 8 cellulase from *Clostridium thermocellum. Proteins: Struct. Funct. Genet.* 25, 134-136.

Spolar, R. S., Livingstoen, J. R. and Record, M. T. Jr (1992) Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water. *Biochemistry,* 31, 3947-3955.

Sriprapundh, D. C., Vieille, C. and Zeikus, J. G. (2000) Molecular determinants of xylose isomerase thermal stability and activity: analysis by site-directed mutagenesis. *Protein Eng.* 13, 259-265.

Stafford, W. R. III (1992) Boundary analysis in sedimentation transport experiments: a procedure for obtaining sedimentation coefficient distributions using the time derivative of the concentration profile. *Anal. Biochem.* 203, 295-301.

Sterner, R., Kleemann, G. R., Szadkowski, H., Lustig, A., Hennig, M. and Kirschner, K. (1996) Phosphoribosyl anthranilate isomerase from *Thermotoga maritima* is an extremely stable and active homodimer. *Protein Sci.* 5, 2000-2008.

Stevens, R. C. and Lipscomb, W. N. (1992) A molecular mechanism for the pyrimidine and purine nucleotide control of aspartate transcarbamylase. *Proc. Natl Acad. Sci.USA,* 89, 5281-5285.

Stevens, R. C., Gouaux, J. E. and Lipscomb, W. N. (1990) Structural consequences of effector binding to the T state of aspartate carbamoyltransferase: crystal structures of the unligated and ATP- and CTP-complexed enzymes at 2.6 Å resolution. *Biochemistry,* 29, 7691-7701.

Subramaniyan, S. and Prema, P. (2002) Biotechnology of microbial xylanases: enzymology, molecular biolgy, and application. *Crit. Rev. Biotechnol.* 22, 33-64.

Summit, M., Scott, B., Nielson, K., Mathur, E. and Baross, J. (1998) Pressure enhances thermal stability of DNA polymerase from three thermophilic organisms. *Extremophiles,* 2, 339-345.

Sun, K., Camardella, L., Di Prisco, G. and Hervé, G. (1998) Properties of aspartate transcarbamylase from *TAD1*, a psycrophilic bacterial strain isolated from Antarctica. *FEMS Microbiol. Lett.* 164, 375-382.

Sun, M. M. C., Caillot, R., Mak, G., Robb, F. T. and Clark, D. S. (2001) Mechanism of pressure-induced thermostabilization of proteins: studies of glutamate dehydrogenases from the hyperthermophile *Thermococcus litoralis. Prot. Sci.* 10, 1750-1757.

Sunna, A. and Antranikian, G. (1997) Xylanolytic enzymes from fungi and bacteria. *Crit. Rev. Biotechnol.* 17, 39-67.

Szewzyk, U., Szewzyk, R. and Stenström, T. A. (1994) Thermophilic, anaerobic bacteria isolated from a deep borehole in granite in Sweden. *Proc. Natl Acad. Sci.USA,* 91, 1810-1813.

Szilagyi, A. and Zavodszky, P. (2000) Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. *Structure,* 8, 493-504.

Tanaka, T., Kawano, N. and Oshima, T. (1981) Cloning of 3-isopropylmalate dehydrogenase gene of an extreme thermophile and partial purification of the gene product. *J. Biochem.(Tokyo)* 89, 677-682.

Tanner, J. J., Hecht, R. M. and Krause, K. L. (1996) Determinants of enzyme thermostability observed in the molecular structure of *Thermus aquaticus* D-glyceraldehyde-3-phosphate dehydrogenase at 2.5 Å resolution. *Biochemistry,* 35, 2597-2609.

Teplyakov, A. V., KUranova, I. P., Harutyunyan, E. H., Vainshtein, B. K., Frommel, C., Hohne, W. E. and Wilson, K. S. (1990) Crystal structure of thermitase at 1.4 Å resolution. *J. Mol. Biol.* 214, 261-279.

Thiry, L. and Hervé, G. (1978) The stimulation of *Escherichia coli* aspartate transcarbamylase activity by adenosine triphosphate. *J. Mol. Biol.* 125, 515-534.

Thoma, R., Hennig, M., Sterner, R. and Kirschner, K. (2000) Structure and function of mutationally generated monomers of dimeric phosphoribosylanthranilate isomerase from *Thermotoa maritima. Structure,* 8, 265-276.

Thomas, A., Hinsen, K., Field, M. J. and Perahia, D. (1999) Tertiary and quaternary conformational changes in aspartate transcarbamylase: a normal mode study. *Proteins,* 34, 96-112.

Tschirret-Guth, R. A., Koo, L. S., Hoa, G. H. and Ortiz De Montellano, P. R. (2001) *J. Am. Chem. Soc.* 123, 3412-3417.

Tsou, C.-L. (1986) Location of the active site of some enzymes in limited and flexible molecular regions. *Trends Biochem. Sci.* 11, 427-429.

Umbarger, H. E. (1956) Evidence for a negative-feedback mechanism in the biosynthesis of isoleucine. *Science,* 123, 848.

Van Boxstael, S. (2004) Asparatate transcarbamylase of the hyperthermophilic archaeon *Pyrococcus abyssi*: structural and kinetic studies. *PhD thesis*

Van Boxstael, S., Cunin, R., Khan, S. and Maes, D. (2003) Aspartate transcarbamylase from the hyperthermophilic archaeon *Pyrococcus abyssi*: thermostability and 1.8 Å resolution crystal structure of the catalytic subunit complexed with the bisubstrate analogue *N*-phosphonacetyl-L-aspartate. *J. Mol. Biol.* 326, 203-216.

Van Boxstael, S., Maes, D. and Cunin, R. (2005) Aspartate transcarbamylase from the hyperthermophilic archaeon *Pyrococcus abyssi. FEBS J.* 272, 2670-2683.

Van de Casteele, M., Chen, P., Roovers, M., Legrain, C. and Glansdorff, N. (1997) Structure and expression of a pyrimidine gene cluster from the extreme thermophile *Thermus* strain ZO5. *J. Bacteriol.* 179, 3470-3481.

Van Dover, C. L. (2000) The ecology of deep-sea hydrothermal vents (Princeton Univ. Press, Princeton).

Van Petegem, F., Collins, T., Meuwis, M. A., Gerday, C., Feller, G. and Van Beeumen, J. (2003) The structure of a cold-adapted family 8 xylanase at 1.3 A resolution. Structural adaptations to cold and investgation of the active site. *J. Biol. Chem.* 278, 7531-7539.

Vedamuthu, M., Sing, S. and Robinson, G. W. (1994) Properties of liquid water: origin of the density anomalies. *J. Phys. Chem.* 98, 2222-2230.

Vickrey, J. F., Hervé, G. and Evans, D. R. (2002) *Pseudomonas aeruginosa* aspartate transcarbamylase. Characterization of its catalytic and regulatory properties. *J. Biol. Chem.* 277, 24490-24498.

Viduguris, G. J. A. and Royer, C. A. (1998) Determination of the volume changes for pressure-induced transitions of apomyoglobin between the native, molten globule, and unfolded states. *Biophys. J.* 75, 463-470.

Vieille, C. and Zeikus, G. J. (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.* 65, 1-43.

Villa, J., Strajbl, M., Glennon, T. M., Sham, Y. Y., Chu, Z. T. and Warshel, A. (2000) How important are entropic contributions to enzyme catalysis? *Proc. Natl Acad. Sci.USA,* 97, 11899-11904.

Villeret, V., Clantin, B., Tricot, C., Legrain, C., Roovers, M., Stalon, V., Glansdorff, N. and Van Beeumen, J. (1998) The crystal structure of *Pyrococcus furiosus* ornithine carbamoyltransferase reveals a key role for oligomerization in enzyme stability at extremely high temperatures. *Proc. Natl Acad. Sci.USA,* 95, 2801-2806.

Vogt, G. and Argos, P. (1997) Protein thermal stability: hydrogen bonds or internal packing? *Fold. Des.* 2, S40-S46.

Vogt, G., Woell, S. and Argos, P. (1997) Protein thermal stability, hydrogen bonds, and ion pairs. *J. Mol. Biol.* 269, 631-643.

Voorhorst, W. G., Eggen, R. I., Luesing, E. J. and De Vos, W. M. (1995) Characterization of the *celB* gene coding for β-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus* and its expression and site-directed mutation in *Escherichia coli. J. Bacteriol.* 177, 7105-7111.

Waldburger, C. D., Schildbach, J. F. and Sauer, R. T. (1995) Are buried salt bridges important for protein stability and conformational specificity? *Nat. Struct. Biol.* 2, 122-128.

Wang, J., Stieglitz, K. A., Cardia, J. P. and Kantrowitz, E. R. (2005) Structural basis for ordered substrate binding and cooperativity in aspartate transcarbamoylase. *Proc. Natl Acad. Sci.USA,* 102, 8881-8886.

Warshel A. (1998) Electrostatic origin of the catalytic power of enzymes and the role of preorganized active sites. *J. Biol. Chem.* 273, 27035-27038.

Whitman, W. B., Coleman, D. C. and Wiebe, W. J. (1998) Prokaryotes: the unseen majority. *Proc. Natl Acad. Sci.USA,* 95, 6578-6583.

Wild, J. R. and Wales, M. E. (1990) Molecular evolution and genetic engineering of protein domains involving aspartate transcarbamoylase. *Annu. Rev. Microbiol.* 44, 193-218.

Wild, J. R., Johnson, J. L. and Loughrey, S. J. (1988) ATP-liganded form of aspartate transcarbamoylase, the logical regulatory target for allosteric control in divergent bacterial systems. *J. Bacteriol.* 170, 446-448.

Wild, J. R., Loughrey-Chen, S. J. and Corder, T. S. (1989) In the presence of CTP, UTP becomes an allosteric inhibitor of aspartate transcarbamoylase. *Proc. Natl Acad. Sci. USA,* 86, 46-50.

Wiley, D. C. and Lipscomb, W. N. (1968) Crystallographic determination of symmetry of aspartate transcarbamylase: studies of trigonal and tetragonal crystalline forms of aspartate transcarbamylase show that the molecule has a three-fold and a two-fold axis. *Nature,* 218, 1119-1121.

Williams, D. H. (1991) *Aldrichimica Acta,* 24, 71-80

Wintrode, P. L. and Arnold, F. H. (2000) Temperature adaptation of enzymes: lessons from laboratory evolution. *Adv. Protein Chem.* 55, 161-225.

Wolfenden, R. (1983) Waterlogged molecules. *Science,* 222, 1087-1093.

Wyman, J. (1948) Heme proteins. *Adv. Protein Chem.,* 4, 407-531.

Xi, X.-G., De Staercke, C., Van Vliet, F., Triniolles, F., Jacobs, A., Stas, P. P., Ladjimi, M. M., Simon, V., Cunin, R. and Hervé, G. (1994) The activation of *Escherichia coli* aspartate transcarbamylase by ATP. Specific involvement of helix H2' at the hydrophobic interface between the two domains of the regulatory chains. *J. Mol. Biol.* 242, 139-149.

Xi, X.-G., Van vliet, F., Ladjimi, M. M., De Wannemaecker, B., De Staercke, C., Glansdorff, N., Piérard, A., Cunin, R. and Hervé, G. (1991) Heterotropic interactions in *Escherichia coli* aspartate transcarbamylase. *J. Mol. Biol.* 220, 789-799.

Xiao, L. and Honig, B. (1999) Electrostatic contributions to the stability of hyperthermophilic proteins. *J. Mol. Biol.* 289, 1435-1444.

Xu, Y. and Glansdorff, N. (2002) Was our ancestor a hyperthermophilic procaryote? *Comp. Biochem. Physiol.* 133, 677-688.

Xu, Y., Feller, G., Gerday, C. and Glansdorff, N. (2003a) Metabolic enzymes from psychrophilic bacteria: challenge of adaptation to low temperatures in ornithine carbamoyltransferase from *Moritella abyssi. J. Bacteriol.* 185, 2161-2168.

Xu, Y., Feller, G., Gerday, C. and Glansdorff, N. (2003b) *Moritella* cold-active dihydrofolate reductase: Are there natural limits to optimization of catalytic effiency at low temperature? *J. Bacteriol.* 185, 5519-5526.

Xu, Y., Nogi, Y., Kato, C., Liang, Z., Rüger, H.-J., De Kegel, D. and Glansdorff, N. (2003c) *Moritella profunda* sp. nov. and *Moritella abyssi* sp. nov., two psychropiezophilic organisms isolated from deep Atlantic sediments. *Internat. J. Syst. Evol. Biol.* 53, 533-538.

Xu, Y., Zhang, Y., Liang, Z., Van de Casteele, M., Legrain, C. and Glansdorff, N. (1998) Aspartate carbamoyltransferase from a psychrophilic deep-sea bacterium, *Vibrio* strain 2693: properties of the enzyme, genetic organization and synthesis in *Escherichia coli. Microbiol.* 144, 1435-1441.

Yano, J. K. and Poulos, T. L. (2003) New understandings of thermostable and peizostable enzymes. *Curr. Opin. Biotech.* 14, 360-365.

Yates, R. A. and Pardee, A. B. (1956) Control of pyrimidine biosynthesis in *Escherichia coli* by a feedback mechanism. *J. Biol. Chem.* 221, 757-770.

Yayanos, A. A. (1986) Evolutional and ecological implications of the properties of deep-sea barophilic bacteria. *Proc. Natl Acad. Sci. USA,* 83, 9542-9546.

Yip, K. S., Stillman, T. J., Britton, K. L., Arymiuk, P. J., Baker, P. J., Sedelnikova, S. E., Engel, P. C., Pasauo, A., Chiaraluce, R., Consalvi, V., Scandurra, R. and Rice, D. W. (1995) The structure of *Pyrococcus furiosus* glutatmate dehydrogenase reveals a key role for io*N*-pair networks in maintaining enzyme stability at extreme temperatures. *Structure,* 3, 1147-1158.

Yip, K. S., Britton, K. L., Stillman, T. J., Lebbink, J., de Vos, W. M., Robb, F. T., Vetriani, C., Maeder, D. and Rice, D. W. (1998) Insights into the molecular basis of thermal stability from the analysis of io*N*-pair networks in the glutamate dehydrogenase family. *Eur. J. Biochem.* 255, 336-346.

Zavodszky, B., Kardos, J., Svingor, A. and Petsko, G. A. (1998) Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc. Natl Acad. Sci. USA,* 95, 7406-7411.

Zechel, D. L. and Withers, S. G. (2000) Glycosidase mechanisms: anatomy of a finely tuned catalyst. *Acc. Chem. Res.* 33, 11-18.

Zhang J., Peng, X., Jonas, A. and Jonas, J. (1995) NMR study of the cold, heat, and pressure unfolding of ribonuclease A. *Biochemistry,* 34, 8631-8641.

Zhou, H.-X. (2002) Toward the physical basis of thermophilic proteins: linking of enriched polar interactions and reduced heat capacity of unfolding. *Biophys. J.* 83, 3126-3133.

Zhu, W., Sandman, K., Lee, G. E., Reeve, J. N. and Summers, M. F. (1998) NMR structure and comparison of the archaeal histone HFoB from the mesophile *Methanobacterium formicicum* with HMfB from the hyperthermophile *Methanothermus fervidus. Biochemistry,* 37, 10573-10580.

Zhu, Z.-Y. and Karlin, S. (1996) Clusters of charged residues in protein three-dimensional structures. *Proc. Natl Acad. Sci. USA,* 93, 8350-8355.

Publications

De Vos, D., Van Petegem, F., Remaut, H., Legrain, C., Glansdorff, N. and Van Beeumen, J. J. (2004) Crystal structure of T state aspartate carbamoyltransferase of the hyperthermophilic archaeon *Sulfolobus acidocaldarius*. *J. Mol. Biol.* 339, 887-900.

De Vos, D., Hulpiau, P., Vergauwen, B., Savvides, S. N. and Van Beeumen, J. (2005) Expression, purification, crystallization and preliminary X-ray crystallographic studies of a cold-adapted aspartate carbamoyltransferase from *Moritella profunda*. *Acta Crystall*. F61, 279-281.

Collins*, T., De Vos*, D., Hoyoux, A., Savvides, S. N., Gerday, C., Van Beeumen, J. and Feller, G. (2005) Study of the active site residues of a glycoside hydrolase family 8 xylanase. *J. Mol. Biol.* 354, 425-435.

Vergauwen, B., De Vos, D. and Van Beeumen, J. J. Characterization of the bifunctional gamma - glutamate-cysteine ligase-glutathione synthetase (GshF) of *Pasteurella multocida*. *J. Biol. Chem.* In press.

De Vos, D.*, Collins, T.*, Nerinckx, W., Savvides, S. N., Claeyssens, M., Gerday, C., Feller, G., and Van Beeumen, J. Oligosaccharide binding in family 8 glycosidases: crystal structures of active site mutants of the β-1,4-xylanase pXyl from Pseudoaltermonas haloplanktis in complex with substrate and product. Submitted for publication.

De Vos, D.*, Vergauwen, B.*, Hulpiau, P., Xu, Y., Glansdorff, N. and Van Beeumen, J.J. Structural basis for cold regulation. Biochemical properties and crystal structure of aspartate carbamoyltransferase from the psychrophilic bacterium Moritella profunda. In preparation.

* both authors contributed equally to this work