HUMAN TESTIS ANGIOTENSIN-CONVERTING ENZYME:

CRYSTAL STRUCTURE OF A GLYCOSYLATION MUTANT AND INVESTIGATION OF A PUTATIVE HINGE-MECHANISM BY NORMAL MODE ANALYSIS

JEAN MARGARET WATERMEYER

A minithesis submitted in partial fulfilment of the requirements for the degree of Magister Scientiae in the Department of Biotechnology,
University of the Western Cape.

Supervisors: Prof. B.T. Sewell & Dr. E.D. Sturrock

November 2004

HUMAN TESTIS ANGIOTENSIN-CONVERTING ENZYME:

CRYSTAL STRUCTURE OF A GLYCOSYLATION MUTANT AND INVESTIGATION OF A PUTATIVE HINGE-MECHANISM BY NORMAL MODE ANALYSIS

Jean Margaret Watermeyer

KEYWORDS

ACE

Angiotensin-converting enzyme

Drug design

Glycosylation

Mutant

Crystallisation

X-ray crystallography

Normal mode

Temperature factor

Hinge-bending

ABSTRACT

HUMAN TESTIS ANGIOTENSIN-CONVERTING ENZYME:

CRYSTAL STRUCTURE OF A GLYCOSYLATION MUTANT AND INVESTIGATION OF A PUTATIVE HINGE-MECHANISM BY NORMAL MODE ANALYSIS

J.M. Watermeyer

M.Sc minithesis, Department of Biotechnology, University of the Western Cape.

Human angiotensin-converting enzyme (ACE) is a key enzyme in the regulation of blood pressure via the renin-angiotensin and kallikrein-kinin systems. A number of orally active drugs have been developed over the years that target somatic ACE, for the treatment of hypertension, myocardial infarction and congestive heart failure. Protein structural information about ACE is an important key for the understanding of the mechanism and substrate-specificity of the enzyme. However, this information has only begun to be elucidated in the past year, with the solution of crystal structures of human testis ACE (tACE), and homologues *Drosophila* AnCE and human ACE2. tACE is identical to the C-terminal domain of somatic ACE, which consists of two homologous domains, each having a slightly different substrate-specificity.

This thesis describes the purification, crystallisation and X-ray crystal structure-determination of a glycosylation-deficient mutant of tACE, tACE-G1,3, to 2.9 Å. The structure of tACE-G1,3 aligns closely with that of native tACE, indicating that the mutations did not alter the conformation. The ability to achieve minimal glycosylation of tACE for crystallisation purposes via mutation, rather than using expensive glycosidase inhibitors,

should prove advantageous for further structural studies, such as the study of the binding of novel inhibitors.

In all of the tACE structures thus far observed, the active site is closed off from the external medium in a deep cleft, so that it is unclear how a large substrate molecule could gain access. However, a hinge motion that opens this cleft has been observed in the structures of ACE2. Temperature factor and sequence comparison between tACE, tACE-G1,3, AnCE and ACE2 suggests the functional conservation of three flexible loop regions, as well as the sequence conservation of three constrained regions, involved in the hinge. Normal mode analysis reveals the intrinsic flexibility of tACE, and further suggests that a putative open form of tACE would behave similarly to the open form of ACE2. Based on these indications, a conservation of the ACE2 hinge-bending mechanism is proposed.

Temperature factor analysis also reveals that subdomain II, containing bound chloride ions, is more structurally rigid than subdomain I, in all structures considered.

Based on these results, lines of investigation are suggested that should yield insight into the mechanisms of action of ACE and its association with various substrates and inhibitors, ideally aiding in the development of novel drugs for the treatment of cardiac disease.

November 2004

DECLARATION

I declare that *Human testis angiotensin-converting enzyme: Crystal structure of a glycosylation mutant and investigation of a putative hinge-mechanism by normal mode analysis* is my own work, that it has not been submitted before for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated and acknowledged as complete references.

Jean Margaret Watermeyer November 2004

Signed: . . . . . .

ACKNOWLEDGEMENTS

This work was funded by the Carnegie Foundation of New York and
the National Research Foundation

Many thanks to:

Ed Sturrock, for his input, encouragement and proofreading

Trevor Sewell, among many other things, for teaching me structural biology

Sylva Schwager for her invaluable advice and assistance in the lab

The Zinc Metalloprotease Group: Ayesha Parker, Trudi van Rensburg, Pierre Redelinghuys, Riyad Domingo, Wendy Kroger

The Structural Biology Group: Muhammed Sayed, Arvind Varsani, Ozlem Bishop, Jason van Rooyen, Tim Frouws, Ndoriah Thuku, James Onyemata, Felix Adusei-Danso, Itai Chitapi, Serah Kimani, Margot Scheffer, Samuel Kwofie

My housemates, Ailsa Leitch, Kat Farquharson and Cari Hudson

My parents, for their interested, enthusiastic support

Andrew Cousins for his support and especially for late-night lifts home

My Lord and saviour, Jesus Christ, who made ACE

CONTENTS

INTRODUCTION:

Human angiotensin-converting enzyme (ACE), a zinc-metallopeptidase, is an important regulator of cardiac function through its C-terminal dipeptidase activity on signalling peptides angiotensin-I and bradykinin. As such, it is the target of a number of drugs used in the treatment of hypertension, myocardial infarction and congestive heart failure.

The first generation of ACE inhibitors currently in clinical use was developed without prior knowledge of the structure of ACE (Acharya *et al*, 2003). Somatic ACE is now known to be a two-domain enzyme, the domains each containing an active site, and having slightly different substrate-specificities (Soubrier *et al*, 1988; Jaspard *et al*, 1993). Domain-specific inhibitors have been designed and used in the *in vitro* and *in vivo* characterisation of ACE activity (Junot *et al*, 2001). The recent solution of a crystal structure of human testis ACE (tACE), identical in sequence to the C-terminal domain of somatic ACE, has revealed the structure of the substrate-binding site and the nature of the binding interactions of known inhibitors (Natesh *et al*, 2003, 2004). In addition, solution of the structures of *Drosophila* and human homologues, ACE2 and AnCE, has given insight into the determinants of substrate-specificity in ACE and ACE homologues (Kim *et al*, 2003; Towler *et al*, 2004).

One reason for the slow elucidation of ACE crystal structures has been the problem of crystallisation. ACE is a glycoprotein, and tACE crystals were eventually obtained by regularising surface glycosylation through the use of the glycosidase-I inhibitor, N-butyl-deoxynojirimycin (Gordon *et al*, 2003). As new potential ACE inhibitors are designed, based on the new structural information now available, it becomes imperative that a reproducible means of crystallisation and structure-determination be developed, for the screening and investigation of these new potential drug leads. The prohibitive cost of glycosidase inhibitors is a barrier to this process.

This study presents the crystallisation and crystal structure-determination of a minimally glycosylated mutant of tACE, tACE-G1,3, to 2.9Å. This mutant was originally produced as one strategy in the attempt to achieve minimal glycosylation for crystallisation purposes (Gordon *et al*, 2003). Structure-determination was carried out in the interests of obtaining new structural information from this mutant. Such novel information was not forthcoming, however, as the structure of tACE-G1,3 does not differ significantly from that of wild-type tACE, despite the fact that measures were taken to minimise model bias. This indicates that the mutations did not disrupt the structure in any way. Moreover, the presence of two fully-processed glycan chains did not have a detrimental effect on crystallisation, causing only a slight rearrangement of lattice contacts. This mutant thus represents a more cost-effective tool for the crystallographic investigation of the binding of inhibitors to tACE.

Both liganded and unliganded structures of tACE and *Drosophila* AnCE are ellipsoid in shape, with the active zinc ion lying deep in a cleft that divides the molecule in half. This cleft is linked with the external milieu by narrow pores that do not provide an adequate path for the entry of bulky substrates (Kim *et al*, 2003). The question thus arises as to the nature of the molecular motions that must occur to allow substrate access to the active site.

In contrast to tACE and AnCE, the unliganded form of ACE2 crystallised in an open conformation, differing from the closed liganded form by a hinge motion of ~16° that opens up the active site by as much as 15Å (Towler *et al*, 2004). This, together with the presence of acetate molecules in the "unliganded" structures of tACE and tACE-G1,3, and calorimetric evidence that entropic factors play an important role in substrate-binding, suggests that a similar hinge mechanism may occur in tACE (Andújar-Sánchez *et al*, 2004).

In this study, the residues important for hinging in ACE2 are identified, and compared with the equivalent residues in tACE, tACE-G1,3 and AnCE, in terms of sequence conservation as well as the Cα-atom temperature factors of the crystal

structures. Conservation of the ACE2 hinge mechanism is suggested by a functional conservation of flexible loop regions, as well as the sequence conservation of structurally constrained regions involved in hinging.

This hypothesis is further supported by normal mode analysis of the open and closed forms of ACE2, the closed structure of tACE, and an open homology model of tACE. Low-frequency vibrational modes for these structures reveal the intrinsic flexibility of tACE about its active site cleft, and suggest that an open form of tACE would behave similarly to that of ACE2.

Analysis of Cα-atom temperature factors also reveals that subdomain II of tACE, containing bound chloride ions, has a high degree of structural rigidity, while subdomain I is more flexible. This domain stability is also evident from the temperature factors of ACE2 and AnCE, indicating a conservation of this characteristic, which may have functional significance.

Because of the importance of ACE as a drug target, the major thrust of further studies based on this work must be in the characterisation and design of novel inhibitors, especially domain-specific inhibitors, with the aim of developing new drugs for clinical use in the treatment of cardiac disease.

Additional studies are also suggested that aim at shedding further light on the mechanism of substrate entry into the active site.

CHAPTER 1: LITERATURE REVIEW

1.1. The importance of ACE as a target for antihypertensive drugs

Human angiotensin-converting enzyme 1 (ACE, peptidyl-dipeptidase A, EC 3.4.15.1) is well known by medical doctors as the target of a variety of drugs commonly used in the treatment of hypertension, myocardial infarction and congestive heart failure. It is a zinc metallopeptidase membrane glycoprotein that cleaves the C-terminal dipeptides off a wide range of oligopeptide substrates. Most tissues produce ACE to some degree, including blood vessels, adrenal glands, heart and brain, but it is most concentrated in the lungs and kidneys (Ehlers & Riordan, 1989).

Among the substrates of ACE are two signalling peptides, angiotensin I (AngI) and bradykinin (BK), both involved in the regulation of blood pressure, and having opposing effects. AngI is the decapeptide (DRVYIHPF / HL) precursor of angiotensin II (AngII), which is a potent vasopressor. The conversion of AngI to AngII, catalysed by ACE, is a key step in the renin-angiotensin system (RAS), which is primarily involved in the maintenance of blood pressure and renal function (reviewed by Zaman *et al*, 2002). BK is a nonapeptide (RPPGFSP / FR), which has vasodilatory activity via the kallikrein-kinin system (KKS). It is inactivated by removal of its C-terminal dipeptide by ACE, with the overall effect of ACE activity on the RAS and KKS thus being an increase in blood pressure. It was because of this effect that ACE became a target for the development of antihypertensives in the 1970's, the result being a first generation of potent ACE inhibitors such as lisinopril, enalaprilat, captopril and their derivatives, which are widely used in cardiovascular therapy today (Zaman *et al,* 2002; Acharya *et al*, 2003).

The involvement of ACE in the regulation of blood pressure and renal function constitutes what is known as its endocrine function. In addition to this, the components of the RAS have been shown to be present in other tissues, such as muscle and bone marrow, where ACE exerts a more localised (paracrine) effect in the maintenance of tissue homeostasis and regulation of cell growth. For

example, ACE has been implicated in haematopoiesis because of its action on N-acetyl-seryl-aspartyl-lysyl-proline (N-AcSDKP), a regulator of haematopoietic stem cell differentiation (Azizi *et al*, 1996; Jones & Woods, 2003; Haznedaroglu & Öztürk, 2003). There has even been the suggestion that ACE acts intracellularly, although this has not been proven (Ehlers & Riordan, 1989).

The discovery of additional functionalities of ACE raises a question as to the effects of ACE inhibitor therapy on systems other than the RAS and KKS. Indeed, ACE inhibitors are not without unpleasant side effects, many of which are not directly linked to their antihypertensive activity (reviewed by Sica, 2004). It is hoped that any new information obtained as to the structure and function of ACE will pave the way for the development of a second generation of more targeted inhibitors, having fewer side-effects.

1.2. Two forms of mammalian ACE

Subsequent to the development of the first generation of inhibitors, human ACE has been found to occur in two active forms, termed somatic and germinal (testis) ACE because of their respective expression profiles.

When somatic ACE (sACE) mRNA was cloned and sequenced in 1988, it became apparent that the coding region comprised two homologous domains, having about 55% overall sequence identity with each other and each containing a putative active site (figure 1.2.1; Soubrier *et al*, 1988; Fernandez *et al*, 2003). sACE is synthesised as a 1306-residue polypeptide, complete with a 29-residue N-terminal signal sequence that is removed during post-translational processing, leaving a 1277-amino-acid mature enzyme (Soubrier *et al*, 1988). The 22 residues at its C-terminus make up a hydrophobic membrane anchor and juxtamembrane region (Soubrier *et al*, 1988). This can be cleaved off between Arg1203 and Ser1204 by a putative ACE-secretase (ACE-sheddase), allowing release of a 1227-residue soluble form of ACE from the cell membrane (Ehlers *et al*, 1991b, 1996; Woodman *et al*, 2000; Santhamma *et al*, 2004).

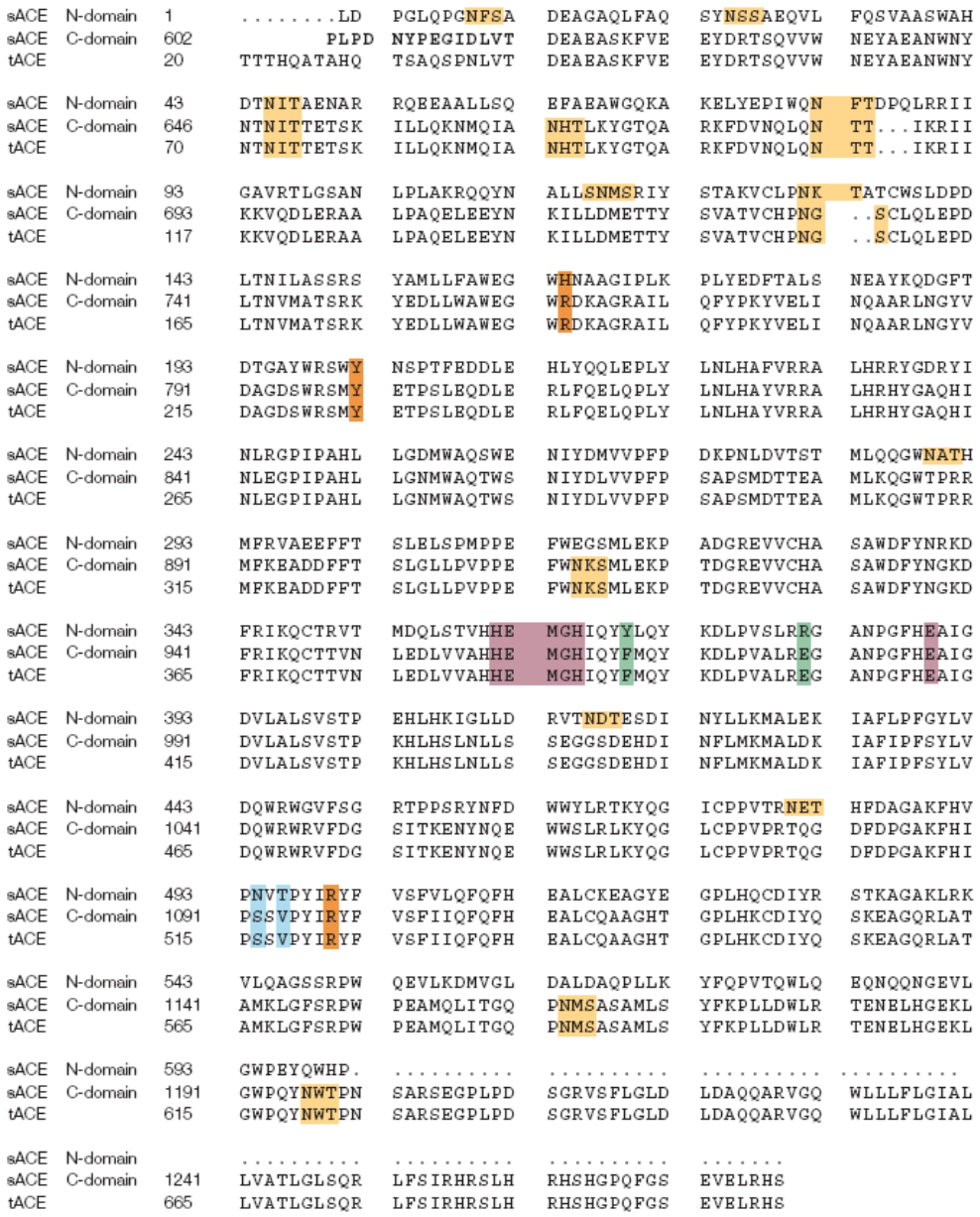


Figure 1.2.1. Sequence alignment of human sACE C-domain, N-domain and tACE (Acharya *et al*, 2003).
The conserved active site zinc-binding motif is highlighted in purple, N-linked glycosylation sites in yellow, chloride ligands in orange and some active site residues in blue/green. Note that tACE is identical to the C-domain of sACE from L37 (613) onwards.

The second ACE form, testis ACE (tACE), found only in adult testis, is 701 amino acids long, and is identical to the C-terminal domain of sACE, with an additional N-terminal signal peptide (figure 1.2.1; Ehlers *et al*, 1989; Kumar *et al*, 1989). tACE is encoded by the same gene as sACE, but is transcribed from an internal promoter located in intron 12 of the full-length gene (Ehlers *et al*, 1989; Ehlers & Riordan, 1989; Hubert *et al*, 1991). It thus contains only one zinc-binding site. This specialised form of ACE has been implicated in the maturation of sperm cells, and has been shown to be necessary for proper sperm function in male mice, in a process that does not involve AngI as a substrate (Hagaman *et al*, 1998; Pauls *et al*, 2003). It should be noted that, while tACE has been shown to be sufficient to carry out the role of sACE in the RAS and in maintaining renal function, sACE is not sufficient to replace tACE in its role in fertility (Kessler *et al*, 2002).

The N-domain of sACE has also been purified as an isolated domain, from human ileal fluid, but it is thought to have been released from membrane-bound sACE by proteolytic cleavage, rather than being produced in a separately-regulated transcription event, as is the case for tACE (Deddish *et al*, 1994).

1.3. ACE homologues

ACE homologues are found across a wide range of species, including mammals and insects, with sequence homology in mammals varying from 80-90% (Bernstein *et* al, 1989; Shai *et al*, 1992; Thekkumkara *et al,* 1992; Cornell *et al,* 1995; Turner & Hooper, 2002; Hens *et al*, 2002).

These mammalian homologues show almost identical activity to human ACE. For example, bovine and mouse ACE (80% and 85% homologous to human ACE, respectively) have been used in catalytic studies that have attempted to elucidate the *in vivo* activity of ACE and ACE inhibitors (Cotton *et* al, 2002; Binevski *et al*, 2003; Georgiadis *et al*, 2003).

In the fly, *Drosophila melanogaster*, two ACE homologues, AnCE and ACEr, have been identified that have about 40% homology to human ACE and are capable of hydrolysing BK and Hippuryl-histidyl-leucine (Hip-His-Leu), a mimic of AngI (figure 1.3.1; Cornell *et al,* 1995; Taylor *et al*, 1996; Coates *et al,* 2000). ACEr appears to share properties with the N-domain of sACE, showing similar catalytic efficiency for Hip-His-Leu, and being capable of binding the N-specific inhibitor, RXP407, while AnCE is catalytically more similar to the C-domain than to the N-domain (Williams *et al*, 2000). The recently solved crystal structures of human tACE and AnCE have revealed a strong structural homology between these two enzymes (Natesh *et al*, 2003; Kim *et al*, 2003).

An ACE homologue that occurs in humans has recently been identified by two distinct strategies (Donoghue *et al*, 2000; Tipnis *et al,* 2000). This homologue, ACE2 (ACEH), is a carboxypeptidase consisting of an N-terminal zinc metallopeptidase domain which is ~42% identical to tACE, and a C-terminal domain which is ~48% identical to human collectrin (figure 1.3.1; Tipnis *et al*, 2000; Zhang *et al*, 2001; Turner *et al*, 2002). The N-terminal domain has also recently been shown to be highly structurally homologous to tACE (Towler *et al*, 2004). The activity of ACE2 differs from that of ACE in that it cleaves only one amino acid residue from the C-terminal end of its substrates (Tipnis *et al*, 2000). Like ACE, it is a regulator of cardiac function, although its activity differs from that of ACE (Crackower *et al*, 2002; Goulter *et al*, 2004). While it is able to act on AngI, it does not cleave BK, is not inhibited by traditional ACE inhibitors, and acts additionally on a several substrates not cleaved by ACE (Tipnis *et al*, 2000; Turner *et al*, 2002). ACE2 is thus thought to share some functions with ACE, but also to play a role in systems in which ACE plays no part (Danilczyk *et al*, 2003).

Figure 1.3.1. Sequence alignment of homologues human ACE2, tACE and *Drosophila* AnCE (Guy *et al*, 2003).
Zinc binding residues are boxed, active site residues are highlighted grey, chloride ligands are underlined. Conserved residues are marked with an asterisk.

tACE also displays structural homology to other members of the zinc peptidase family, notably neurolysin and a carboxypeptidase from the hyperthermophilic archaeon *Pyrococcus furiosus,* which are quite different with respect to sequence (Arndt *et al*, 2002; Natesh *et al*, 2003). The presence of ACE homologues in such a broad range of species supports the hypothesis that this enzyme is critical for the maintenance of tissue homeostasis and regulation of cell growth.

1.4. Substrates and catalytic activity of sACE

The M2 family of zinc metallopeptidases of which ACE is a member, is one of several families of related zinc metallopeptidases, all containing a characteristic His-Glu-Xaa-Xaa-His (HEXXH) zinc-binding motif, and sharing a common reaction mechanism (Lipscomb & Sträter, 1996). This motif is present in both domains of sACE, and it has been shown that both domains have similar catalytic activities, are inhibited by the traditional ACE inhibitors, and can function as isolated domains (Wei *et al.* 1991, 1992; Jaspard *et al*, 1993).

Detailed studies, however, have shown that the substrate- and inhibitor-specificities and catalytic activities of the two domains are slightly different, as is to be expected considering the differences in their amino-acid sequences. For example, the substrates N-AcSDKP, Angiotensin$_{1\text{-}7}$ (Ang$_{1\text{-}7}$), and luteinising hormone-releasing hormone (LH-RH) are preferentially cleaved by the N-domain, while the phosphinic peptides RXP407 and RXPA380 have been shown to be selective inhibitors of the N- and C-domains respectively (Rousseau *et al,* 1995; Deddish *et al*, 1998; Junot *et* al, 2001; Cotton *et* al, 2002; Georgiadis *et al*, 2003). Interestingly, Ang$_{1\text{-}7}$, an N-domain-specific substrate, inhibits C-domain activity (Deddish *et al*, 1998). The two domains also have different affinities for the traditional inhibitors trandolaprilat, lisinopril, enalaprilat and captopril (Wei *et al*, 1992). Similar domain-specific preference for some substrates and inhibitors over others has also been recorded for bovine ACE *in vitro* (Binevski *et al*, 2003).

The use of the domain-specific inhibitors RXP380 and RXP407 on mouse ACE has shown that the N-and C- domains of this enzyme are responsible for 30% and 70% of AngI cleavage, respectively, with a similar ratio applying to BK inactivation (Cotton *et* al, 2002; Georgiadis *et al*, 2003). Studies in mice, using sACE constructs having inactive or deleted N- or C-domains, have shown that the C-domain of human sACE cleaves the substrates Hip-His-Leu and AngI at a faster rate than the N-domain, despite the fact that the binding constants are similar (Wei *et al*, 1991). These data, supported by the recent finding that the C-domain alone is sufficient to maintain RAS function in N-domain knockout mice, suggest that the main function of the N-domain is in an activity other than its haematoregulatory role in the RAS (Fuchs *et al*, 2004).

A further difference between the domains is that of chloride-dependence. ACE undergoes a substrate-dependent enzyme activation in the presence of chloride (Shapiro *et al*, 1983). This activation of ACE is the result of strong chloride-dependent activation of the C-domain, while the N-domain is only very weakly activated by chloride ions (Wei *et al*, 1991, 1992; Acharya *et al*, 2003). However, since ACE has only been observed to function *in vivo* at physiological chloride concentrations, at which the activity of the C-domain is maximal, the significance of this observation is not clear.

Also unclear is what effect the N- and C- domains have on each other's activity. Indeed, it would seem that the two domains are relatively independent, since they are both capable of catalysis in the absence of the other domain (Wei *et al*, 1991, 1992; Cotton *et* al, 2002; Georgiadis *et al*, 2003; Fuchs *et al*, 2004). However, the *in vivo* inhibition of BK hydrolysis in mice requires that both domains be inhibited, despite the fact that this substrate is specific to the N-domain (Georgiadis *et al*, 2003). This, together with an experiment that showed that truncation of N-domain altered the kinetic properties of the C-domain, would seem to suggest that some measure of cooperativity is involved (Marcic *et al*, 2000). Moreover, there is evidence for the negative cooperativity of the two domains of bovine ACE, *in vitro*, and a 1:1 stoichiometry for inhibitor-binding

has been observed in microcalorimetry experiments of pig lung ACE (Binevski *et al*, 2003; Andújar-Sánchez *et al*, 2004).

The C-domain of sACE is more prone to thermal denaturation than the N-domain, and it has been hypothesised that the N-domain may play a role in stabilising the structure of the C-domain (Voronov *et al*, 2002). Similar evidence for the greater stability of the N-domain is its survival in ileal fluid, when the C-domain appears to be completely degraded, although this could be mediated by its higher levels of glycosylation (Deddish *et al*, 1994).

1.5. ACE glycosylation

The N-terminal 36 residues of human tACE are heavily O-glycosylated (Soubrier, *et al*, 1988; Ehlers *et al*, 1992). This region seems to be important for its function in the binding of sperm to the oviduct epithelium (Hagaman, 1998). However, this glycosylation is not essential for expression or enzymatic activity (Ehlers *et al*, 1992; Kasturi *et al*, 1994).

While sACE is not O-glycosylated, each domain has several potential N-linked glycosylation sites, nine in the N-domain and seven in the C-domain (and in tACE) (figure. 1.2.1). In tACE, of these seven potential N-linked glycosylation sites, the three most N-terminal are fully glycosylated, the following three partially glycosylated, and the C-terminal-most site is unglycosylated (Yu *et al*, 1997). These sites have been shown to have differential importance for the expression of stable, enzymatically active tACE (Kasturi *et al*, 1994; Gordon *et al*, 2003).

In general, glycosylation poses a problem for crystallisation, since the heterogeneity and flexibility of glycans within a sample of purified glycoprotein makes the formation of a uniform crystal lattice impossible (Hogg *et al*, 2003; Mehndiratta *et al*, 2004). It was thus necessary to reduce the level of glycosylation in order to obtain a sample of ACE or a homologue that was

suitable for crystallisation. Two approaches to this problem have recently shown encouraging results. tACE was crystallised as a truncated enzyme lacking the O-glycosylated N-terminal sequence, after expression in the presence of a glycosidase-I inhibitor, N-butyl-deoxynojirimycin (NB-DNJ), which enforces uniform glycosylation (Butters *et al*, 1999; Gordon *et al*, 2003). The *Drosophila* homologue AnCE, and the human homologue ACE2, were crystallised after expression in active, unglycosylated form using baculovirus systems (Kim *et al*, 2003; Towler *et al*, 2004). In addition, promising initial crystallisation results were observed for tACE mutants having various combinations of N-linked glycosylation sites knocked out by Asn-Gln point mutation (Gordon *et al*, 2003).

1.6. Crystal structures

Crystal structures of tACE, AnCE and ACE2 were published within a short space of time, bringing a sudden flood of structural information after years during which no progress had been made on this front. The structure of tACE to 2.0Å was published in January 2003, followed shortly afterward by AnCE to 2.4Å in February 2003, and ACE2 to 2.2Å in January 2004 (Natesh *et al*, 2003; Kim *et al*, 2003; Towler *et al*, 2004). Based on the structure of tACE, a number of homology models of the N-domain of sACE (sACEn) have been calculated (Fernandez *et al*, 2003; Tzakos *et al*, 2003).

All four structures are ellipsoid in shape, consisting mostly of α-helices with only a few short beta strands, and are divided into two subdomains by a deep cleft in which the active site is shielded from the external milieu (figure 1.6.1; Natesh *et al*, 2003). This cleft effectively cuts the structure in half, as it extends almost all of the way across the molecule, so that the two subdomains are connected only by the few α-helices that form the floor of the cleft. Similar shielding of the active residues is a common feature of proteases, serving to limit access of larger or folded substrates to the active site (Brown *et al,* 2001; Towler *et al,* 2004).

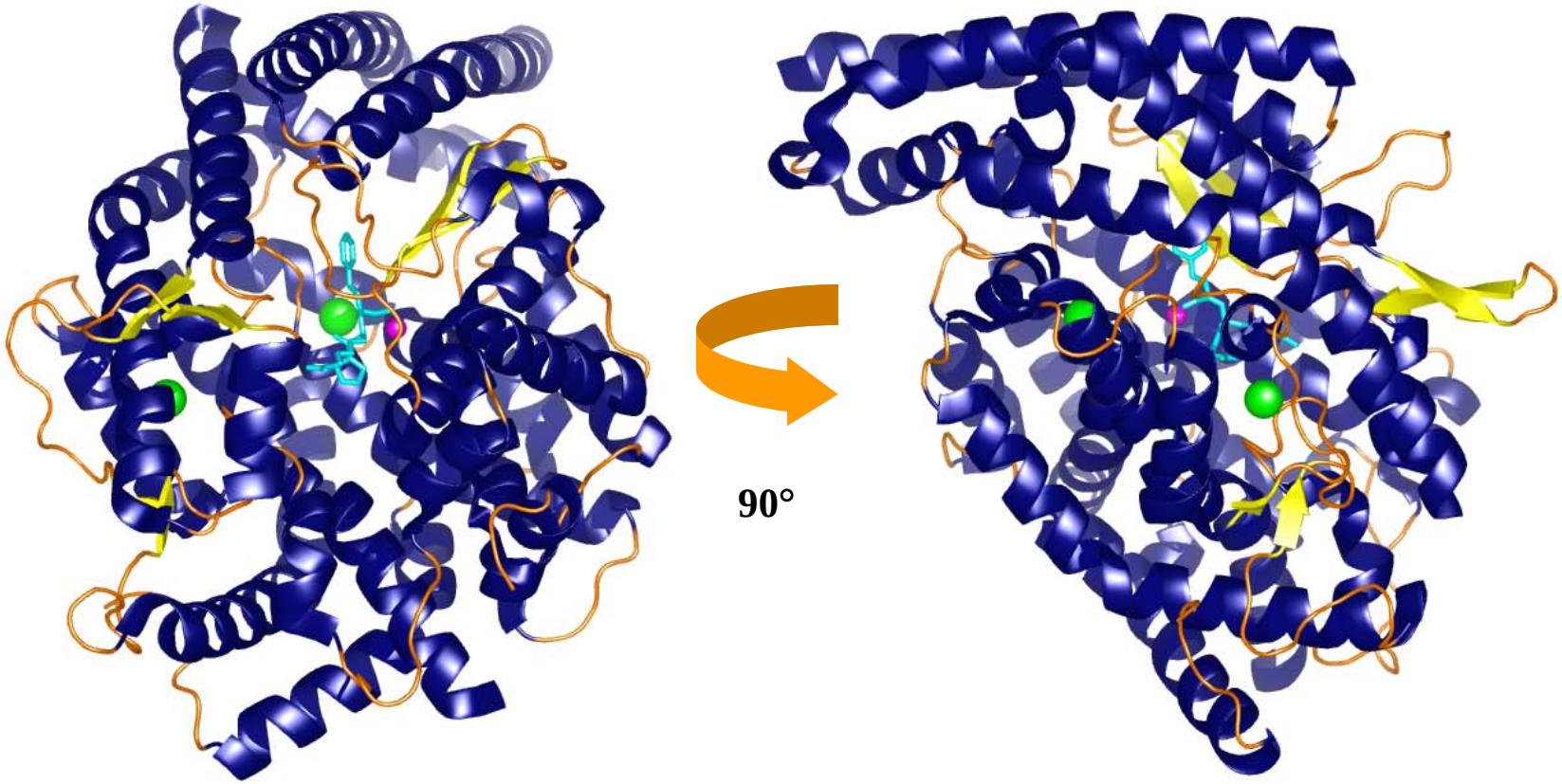


Figure 1.6.1. Overview of the crystal structure of human tACE (PDB ID 1o8a) with bound inhibitor, lisinopril. The two views differ by a 90° rotation. α-helices are coloured navy, β-strands yellow, the zinc ion magenta, chloride ions light green and lisinopril cyan.

The arrangement of the residues of the HEXXH zinc-binding motif of the active site is highly conserved throughout the zinc-binding metallopeptidases that carry this motif, despite weak sequence or even structural homology outside of the motif. For example the root mean square deviation (rmsd) between the active site residues of ACE2 and tACE is 0.53Å, and the rmsd between those of tACE and thermolysin is 0.52Å (Natesh *et al*, 2003; Towler *et al*, 2004).

In this common coordination geometry, the zinc ion is tetra-coordinated by the amide groups of the two His residues of the HEXXH-motif, the carboxyl oxygens of an additional conserved Glu located 24 residues downstream of the motif, and the inhibitor (mimicking the carbonyl oxygen of the scissile peptide bond) or bound water molecule in the case of unliganded structures (figure 1.6.2).

The zinc-binding site forms a bottleneck in the central groove, to one side of which lies a small cavity in which the substrate C-terminus binds, and to the other side of which is a larger cavity that accommodates the substrate N-terminus (figure 1.6.3; Kim *et al*, 2003). It is at this bottleneck that the scissile peptide bond (or equivalent inhibitor mimic) lies, with the carboxyl dipeptide to one side and the remaining N-terminal residues to the other.

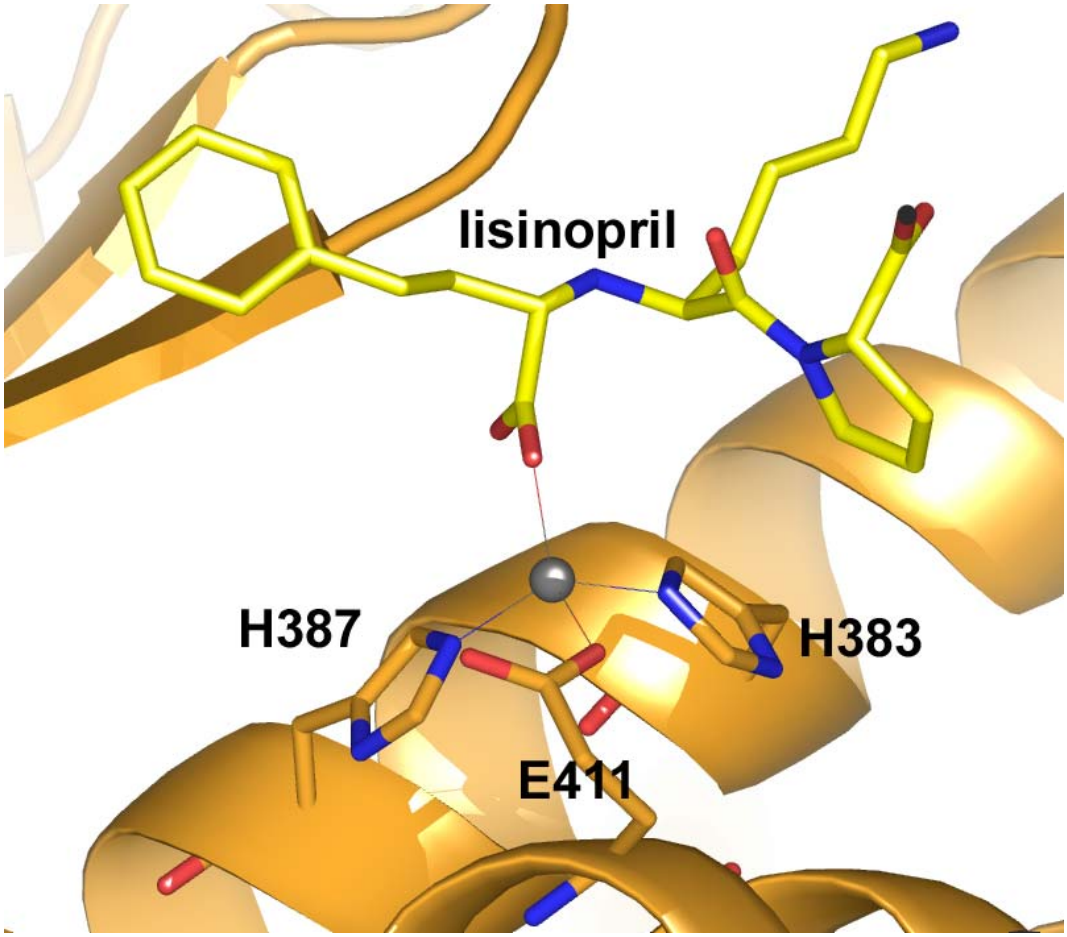


Figure 1.6.2. Tetra-coordination of the zinc ion at the active site of human tACE by inhibitor lisinopril and His383, His387 and Glu411 of the HEMGH + E zinc-binding motif (PDB ID 1o8f).

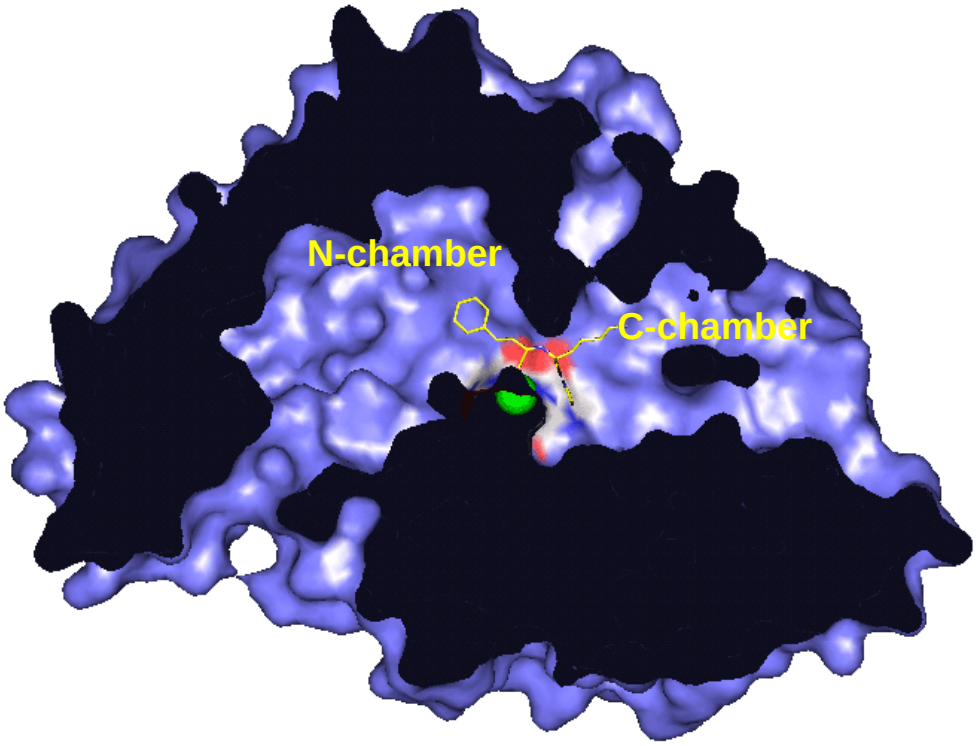


Figure 1.6.3. Cut-away view of human tACE (1o8a), showing the inhibitor, lisinopril, in the active site, with the zinc ion (green sphere) located at the bottleneck between N- and C- chambers.

1.7. Catalytic Mechanism

Considering the similarity of the active sites, it is likely that the catalytic mechanism is also conserved. This mechanism has been elucidated by extensive catalytic and structural studies on ACE, as well as on thermolysin and other zinc metallopeptidases (Williams *et al*, 1994; Lipscomb & Sträter, 1996; Pelmenschikov *et al*, 2002).

Briefly, the proposed mechanism is as follows (figure 1.7.1): The zinc ion and the Glu within the HEXXH motif form H-bonds with and polarise a nucleophilic water molecule at the active site, while the zinc ion also binds to and polarises the scissile peptide carbonyl group. The Glu abstracts a water proton, and the nucleophilic water oxygen attacks the carbonyl carbon, leading to the formation of a carboxyl anion intermediate. This intermediate is stabilised by the presence of positively charged residues (H-bond donors) such as His and Tyr.

A neighbouring protonated His transfers a proton onto the peptide nitrogen, making this group positively charged. This unstable state decomposes with breakage of the peptide bond, abstraction of a proton from the Glu by the leaving amide group, and transfer of a proton from the leaving carboxyl group back to the donor His (Lipscomb & Sträter, 1996; Pelmenschikov *et al*, 2002; Towler *et al*, 2004).

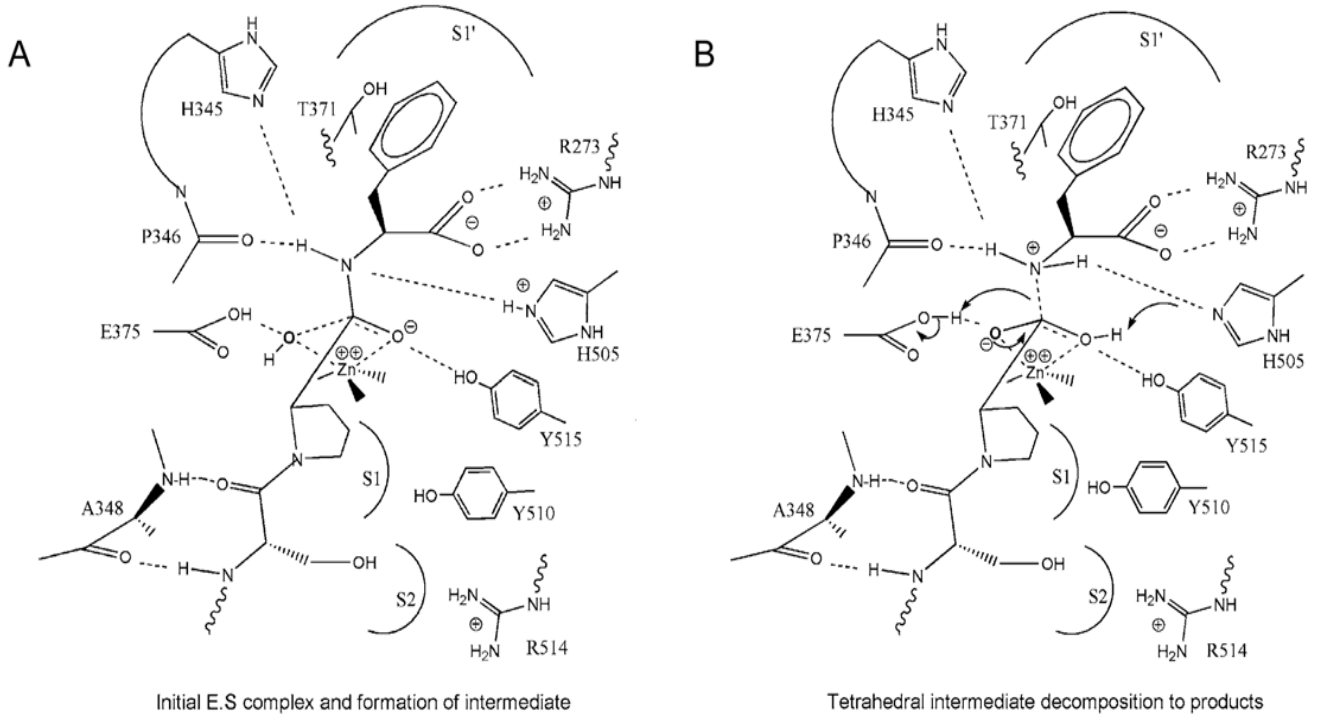


Figure 1.7.1. Catalytic mechanism example: Proposed catalytic mechanism for ACE2.
A) Formation of the ES complex: polarisation of the Zn-bound nucleophilic water by the Zn ion and the abstraction of a water proton by E375 leads to the attack of the carbonyl carbon by the nucleophilic water oxygen, resulting in the formation of a carboxyl anion intermediate. Simultaneous transfer of a proton from H505 to the peptide nitrogen makes this group positively charged. H505 and Y515 stabilise the intermediate.
B) The tetrahedral intermediate collapses with breaking of the unstable peptide bond, abstraction of a proton from the leaving amide by E375, and transfer of a proton from the leaving carboxyl group to H505. (Lipscomb & Sträter, 1996; Towler *et al*, 2004)

1.8. Inhibitor binding and substrate specificity

The first-generation ACE inhibitors was designed based on active site models that were taken from those of carboxypeptidase A (CPA) and thermolysin, both well-characterised enzymes then thought to be mechanistically similar to ACE (Acharya *et al*, 2003).

Captopril, the first to be put to clinical use, is a dipeptide mimetic (of carboxy-Ala-Pro, with a sulphydryl group replacing the zinc-binding scissile carbonyl group) that occupies the S_1' (Ala) and S_2' (Pro-COO^-) subsites of the enzyme. Lisinopril and enalaprilat are tripeptide mimetics (of Phe-Lys-Pro and Phe-Ala-Pro, respectively), additionally occupying the S_1 subsite, and having *N*-carboxyalkyl groups replacing the scissile carbonyl group (figure 1.8.1; Patchett *et al*, 1980; Turner & Hooper, 2002; Acharya *et al*, 2003). While it has become clear that ACE is not homologous, or closely related in terms of structure, to CPA, the crystal structures show that these predictions for the inhibitor-binding sites were correct.

The first level of control for substrate-specificity is probably at the level of the shape and size of the substrate-binding cleft. The small size of the cleft C-terminal to the zinc-binding site limits the number of residues that can be cleaved off the peptide, while the dimensions of the N-terminal region provide a length limitation, as well as preventing the binding of folded substrates (Brown *et al*, 2001). This effect is illustrated by the difference in activities between tACE and ACE2. The former accommodates two substrate residues C-terminal to the zinc ion (in the S1' and S2' subsites) and thus has a dipeptidase activity, while the latter accommodates only one residue in the S1' position due to crowding of the potential second site by a bulky Arg, and thus has monopeptidase activity (Natesh *et al*, 2003; Towler *et al*, 2004).

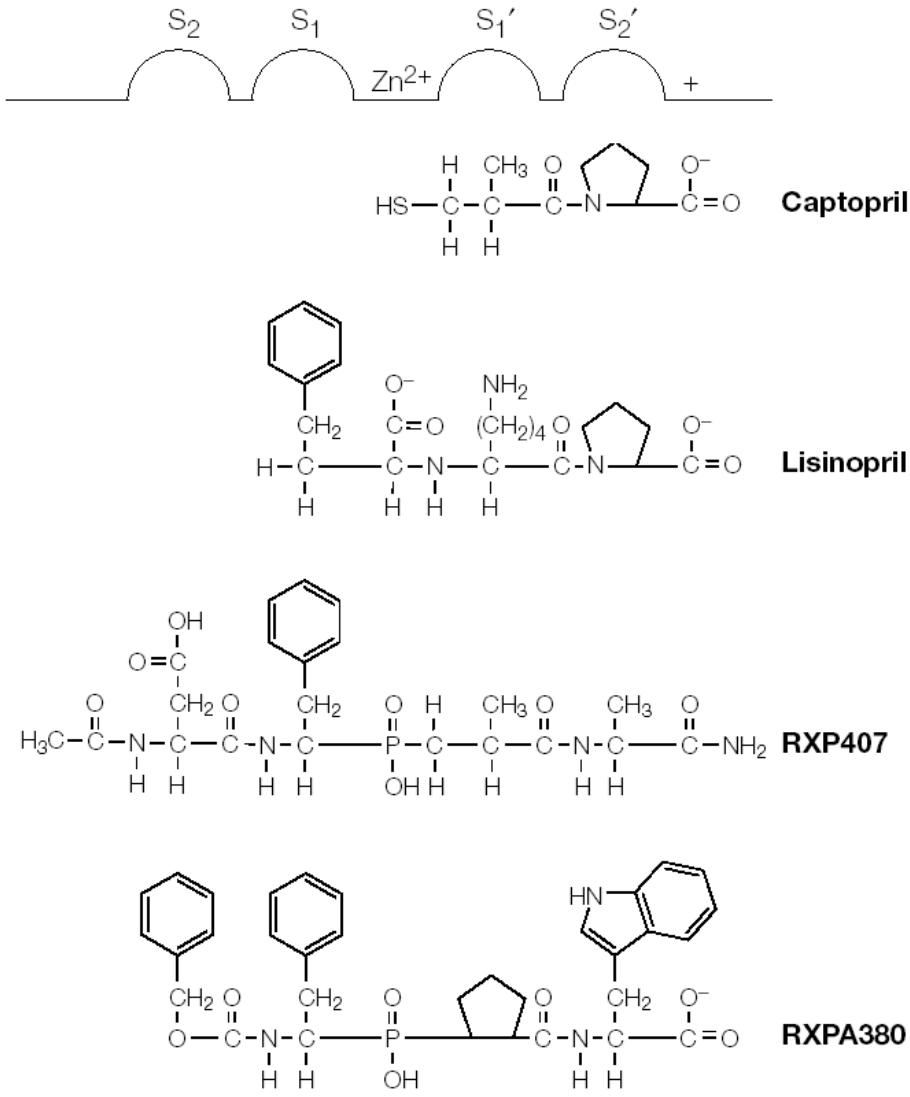


Figure 1.8.1. Traditional representation of the active site of ACE, with ACE inhibitors captopril and lisinopril, and phosphinic peptide inhibitors RXP407 (N-domain-specific) and RXPA380 (C-domain-specific) (Acharya *et al*, 2003).

More specific control of substrate-binding occurs at the level of the structure of the active site, and the residues involved in the interaction. In addition to the unliganded structures of tACE, AnCE and ACE2, several inhibitor-bound structures have been published, and these shed light on the interaction of inhibitors with the active site residues. Structures of both tACE and AnCE have been solved with lisinopril and captopril in the active site, and tACE has additionally been solved bound to enalaprilat (Kim *et al*, 2003; Natesh *et al*, 2003, 2004). ACE2 was also solved bound to a specific inhibitor, (*S*,*S*) 2-{1-carboxy-2-[3[(3,5-dichloro-benzyl)-3*H*-imidazole-4-yl]-ethylamino}-4-methyl-pentanoic acid (MLN-4760; Towler *et al*, 2004).

The residues comprising the S_2' site, in which the C-terminal residue of the substrate or inhibitor binds, are conserved in tACE, $sACE_{N\text{-}domain}$ and AnCE. The carboxyl oxygens of the substrate C-terminus form H-bonds and ionic interactions with a conserved Lys and Tyr that appear to serve as an anchor for the substrate (figure 1.8.2, K511 & Y520; table 1.8.1). The S2' residue side chain occupies a hydrophobic pocket consisting of three conserved Phe and two Tyr. This explains the preference of ACE for bulky hydrophobic residues such as Pro or Trp in this position (Wei *et al*, 1992; Bala *et al*, 2002). These residues are all different in ACE2, and the entire S_2' subsite is occluded (Towler *et al*, 2004).

The S_1' subsite of tACE, AnCE and $sACE_{N\text{-}domain}$, occupied by the Lys residue of lisinopril or the Ala residue of captopril and enalaprilat, is lined by predominantly negative and H-bonding residues such as Gln, Glu, Thr and Asp (table 1.8.1). These residues interact with the Lys side-chain by charge-charge interactions and by water-mediated H-bonds (figure 1.8.2). The greater affinity of tACE for lisinopril than for enalaprilat has been attributed to the replacement of the shorter Ala sidechain of enalaprilat with a longer Lys sidechain which interacts more extensively with the deep S_1' site (Acharya *et al*, 2003).

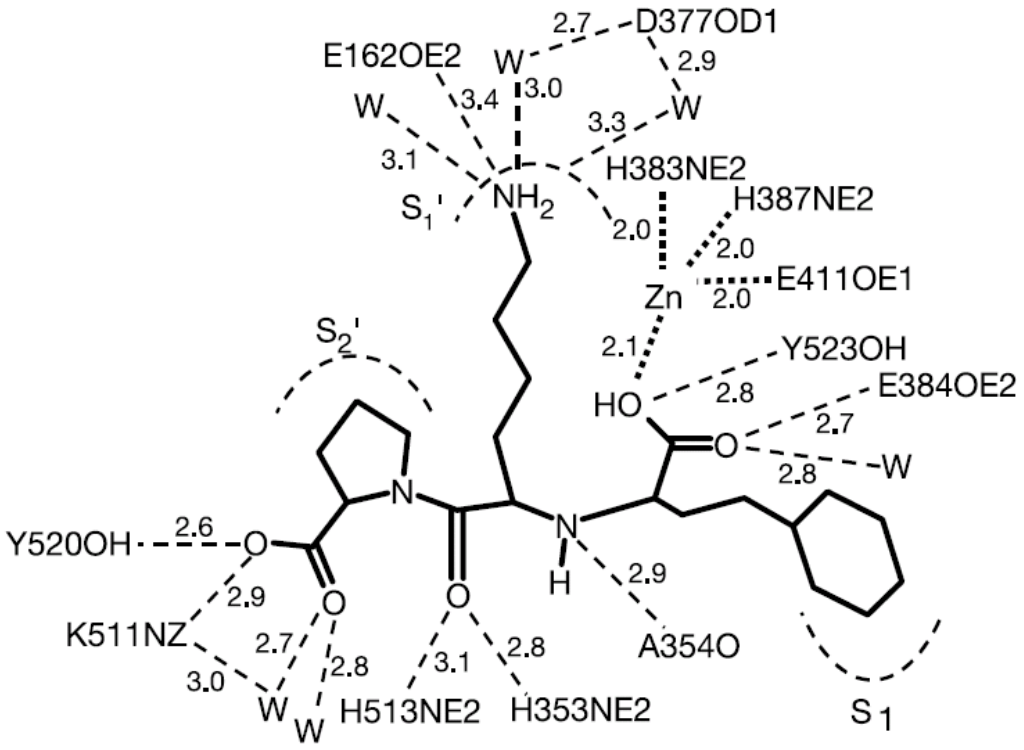


Figure 1.8.2. Diagrammatic representation of the layout of lisinopril-binding residues in the active site of tACE, showing the S1, S1' and S2' subsites (Natesh *et al*, 2003). W: water molecule.

Table 1.8.1. Residues of the C-domain of human sACE making contact with lisinopril versus those in the N-domain model, according to Fernandez *et al* (2003). Residues that are different in the S1' subsite, and are thought to be important for substrate-selectivity, are highlighted in grey.

Position	Lisinopril	sACE-C	Distance (Å)	sACE-N	Distance (Å)	sACE-N minimized	Distance (Å)
S1	Aromatic ring	Ring-Phe_{1117}	3.7	Ring-Phe_{521}	3.8	Ring-Phe_{521}	3.6
0	COOH-(C=O)	W_{747}	2.8[a]	W_{632}	2.8	—	—
0	COOH-(C=O)	OE2-Glu_{989}	2.7[a]	OE2-Glu_{393}	2.7	OE2-Glu_{393}	3.6
0	COOH-(C=O)	W_{747} OE2-Glu_{989}	2.8 3.2	W_{385} OE2-Glu_{393}	2.8 3.2	—	—
0	COOH-(O^-)	OH-Tyr_{1128}	2.8[a]	OH-Tyr_{532}	2.7	OH-Tyr_{532}	2.9
0	COOH-(O^-)	Zn	2.1[a]	Zn	2.3	Zn	2.0
S1′	Lys-(NH_2)	W_{385}–OD1-Asp_{982}	3.3–2.9[a]	W_{671}–OD1-Gln_{386}	3.0–2.9	OD1-Gln_{386}	4.6
S1′	Lys-(NH_2)	W_{787}–OD1-Asp_{982}	3.0–2.7[a]	W_{718}–OD1-Asp_{385}	3.2–2.7	OD1-Asp_{385}	3.7
S1′	Lys-(NH_2)	OE2-Glu_{767}	3.4[a]	OE2-Asp_{171}	5.0	OE2-Asp_{171}	4.3
S1′	Lys-(NH_2)	W_{1038}	3.5[a]	W_{915}	3.1	—	—
S1′	Cα-(NH)	O-Ala_{959}	2.9[a]	O-Ala_{363}	2.9	O-Ala_{363}	2.3
S1′	Cα-(NH)	OE-Glu_{989}	3.4	OE-Glu_{393}	3.4	OE-Glu_{393}	2.3
S1′	Cα-(C=O)	NE2-His_{958}	2.8[a]	NE2-His_{362}	2.7	NE2-His_{362}	2.1
S1′	Cα-(C=O)	NE2-His_{1118}	3.1[a]	NE2-His_{522}	3.1	NE2-His_{522}	2.5
S2′	Pro-(C=O)	W_{866}	2.8[a]	W_{749}	2.8	—	—
S2′	Pro-(C=O)	W_{759}–NZ-Lys_{1116}	2.7–3.0[a]	W_{644}–NZ-Lys_{520}	2.7–3.0	NZ-Lys_{520}	3.1
S2′	Pro-(O^-)	NZ-Lys_{1116}	2.9[a]	NZ-Lys_{520}	2.9	NZ-Lys_{520}	2.0
S2′	Pro-(O^-)	OH-Tyr_{1125}	2.6[a]	OH-Tyr_{529}	2.7	OH-Tyr_{529}	3.2
S2′	Pro-ring	Ring-Tyr_{1125}	3.6	Ring-Tyr_{529}	3.6	Ring-Tyr_{529}	3.7

The presence of a Gln and Asp in this site in $sACE_{N\text{-}domain}$ (Q386 and D171 in $sACE_{N\text{-}domain}$ versus Asp377/982 and Glu162/767 in tACE/$sACE_{c\text{-}domain}$), seems to result in a preference of the N-domain for inhibitors having an Ala at this position, over lisinopril (table 1.8.1; Wei *et al*, 1992; Fernandez *et al*, 2003; Natesh *et al*, 2004). The reason for this preference is unclear.

In ACE2 the S_1' site is much larger than in the other ACE structures, and is lined by numerous H-bonding and hydrophobic side-chains that accommodate the bulky 3,5-dichloro-benzyl imidazole group of MLN-4760. The C terminus of the inhibitor is anchored by an interaction with Arg273, the side-chain of which occludes the space corresponding to the S_2' subsite in tACE and AnCE (figure 1.7.1; Towler *et al*, 2004).

The S_1 site appears to accommodate predominantly large hydrophobic groups, such as the phenylpropyl group of lisinopril, which makes van der Waal's contacts with Val518 of tACE, and with Phe521 in the N-domain model (Fernandez *et al*, 2003). This could also contribute to the preference of the N-domain for captopril over lisinopril, since the bulky Phe521 residue may make binding of the bulky phenylpropyl group of lisinopril unfavourable (Wei *et al*, 1992).

Acharya *et al* (2003) drew on the structure of tACE and kinetic studies to derive some requirements for domain-specificity, which could be used in the design of second-generation inhibitors of human ACE. In this scheme, C-selectivity is determined by bulky P_2' and P_1' groups with a large neutral or basic P_2 group, while N-selectivity is determined by an amidated C-terminal carboxyl group in the P_2' position, and an acidic P_2 group. Domain-specific inhibitors are predicted to provide some solutions to the side-effects observed for ACE inhibitors currently in clinical use, and perhaps to exert new activities based on N-domain selectivity (Acharya *et al*, 2003).

1.9. Chloride binding sites and chloride activation

The structure of tACE revealed two highly ordered, buried chloride ions, consistent with the strong chloride-dependence that has been observed for the C-domain of sACE (Natesh *et al*, 2003).

One of the ions, CL2 (~10Å from the active site), is bound to R522, which is part of the same α-helix as two important inhibitor ligands, Y520 and Y523 (figure 1.8.2). Thus it was suggested that the binding of chloride to R522 has some effect on the active site, via Y520 and Y523 (Natesh *et al*, 2003). This is supported by a previous study that revealed that R522 (R1098 in $sACE_{c\text{-}domain}$) is critical for chloride-dependence (Liu *et al*, 2001).

A modelling investigation by Tzakos *et al* (2003) revealed a cavity filled with a network of five water molecules that links CL2 with the surface of the protein, and could serve as an ion-conduction pore and putative ion-selectivity filter, based on sequence similarity to the ion-selectivity channel of *Salmonella typhimurium* serovar *typhimurium* CLC chloride channels. In this model, CL2 was proposed to serve as an "ionic switch" that plays a role in the positioning of Y523 in the active site.

However, it has also been observed that W279, which flanks the binding site of CL1 (~20Å from the active site) is a ligand of lisinopril, and that this chloride is also in close proximity to K511, which binds the inhibitor C-terminus (figure 1.8.2; Tzakos *et al*, 2003).

ACE2 is also activated by chloride anions, and a single bound chloride ion was observed in the crystal structure of unliganded ACE2, bound to R169, W477 and K481 (via a water), which correspond to R186, W485 and R489 of the CL1 site in tACE (Towler *et al*, 2004). The important ligands in the CL2 site of this enzyme are also conserved, but the nearby Pro residues of tACE are replaced by E398 and S511, the side chains of which extend into the chloride-binding site, eliminating it

(Towler *et al*, 2004). Due to the low resolution of this structure, the possibility of a second unresolved chloride-binding site in ACE2 could not be ruled out.

Thus, the mechanism of chloride activation remains unclear, and the binding or release of either of these two chlorides could alter the configuration of the active site.

The weaker chloride-activation of the N-domain has not yet been adequately explained. The residues directly involved in chloride-binding in tACE are largely conserved in $sACE_{N\text{-}domain}$; with the single (and possibly significant) mutation of R186 to H195 (figure 1.2.1); as are the putative ion-selectivity filter residues (Fernandez *et al*, 2003; Tzakos *et al*, 2003).

1.10. The hinge-bending mechanism of ACE2

In the structures of tACE and AnCE, inhibitor-bound and otherwise, the active site residues are buried deep in the active cleft, being closed off from the external milieu to such an extent that it is unclear how a bulky substrate molecule might gain access. No major conformational differences were observed between inhibitor-bound and unliganded structures (Natesh *et al*, 2003; Kim *et al*, 2003). Based on this evidence, it was proposed that a conformational change, or "breathing motion" must occur in the residues surrounding the potential access channels, allowing for substrate-binding (Natesh *et al*, 2003; Fernandez *et al*, 2003; Kim *et al*, 2003).

The unliganded structure of ACE2, however, is significantly different from its inhibitor-bound counterpart. The active site cleft in unliganded conformation is as much as 15Å wider open at the outer edge than in the inhibitor-bound structure, and the enzyme appears to undergo a ligand-dependent hinge-bending movement of ~16° that closes the substrate into the active site (figure 1.10.1; Towler *et al*, 2004).

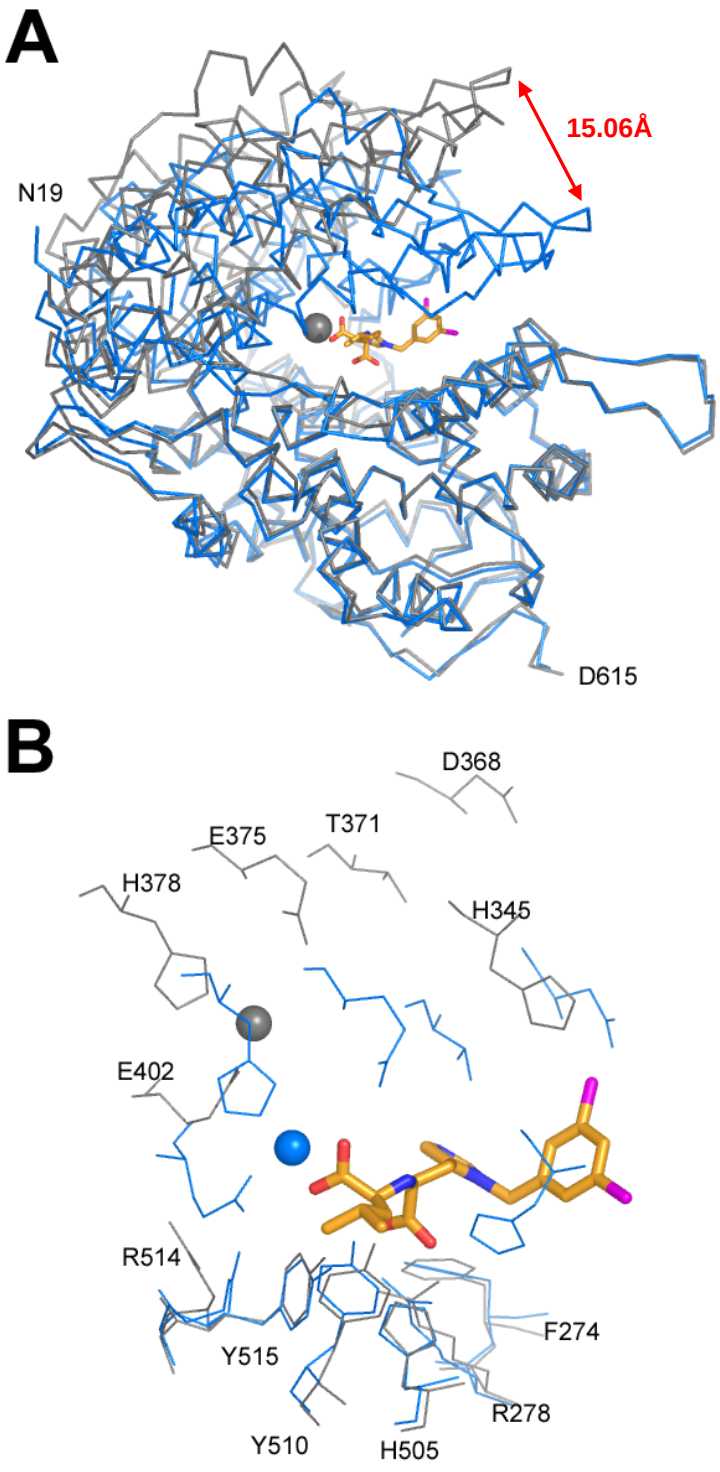


Figure 1.10.1. The hinge movement of ACE2.
A) Movement of one half of the molecule relative to the other closes the active site on substrate-binding, by as much as 15Å. **B)** Residues at the active site move into contact with the bound inhibitor MLN-4760.
Black lines: unliganded state; blue lines: inhibitor-bound state; spheres: zinc ion. (Towler *et al*, 2004)

This type of substrate-dependent hinge-bending mechanism has been observed in a variety of other enzymes, including T4 lysozyme and the *Escherichia coli* D-allose-binding and L-arabinose-binding proteins, all of which have their active sites occluded deep within a cleft in the substrate-bound state (Newcomer *et al*, 1981; Gerstein *et al*, 1994; Mchaourab *et al*, 1997; Magnusson *et al*, 2002). Notably, the same mechanism has been observed in other zinc metallopeptidases, including thermolysin, *Bacillus cereus* neutral protease and *Pseudomonas aeruginosa* elastase; all members of the M4 class of zinc metallopeptidases that have the same (HEXXH + E) zinc binding site motif as ACE and ACE2 (Zhang *et al*, 1995; Lipscomb & Sträter, 1996; Hausrath & Matthews, 2002).

This suggests that the hinge-bending motion observed in ACE2 may be a common feature of this class of peptidases, and considering the close structural homology of human ACE to ACE2, it is likely that it uses a similar mechanism. This possibility seems to be negated by the structure of unliganded tACE solved by Natesh *et al* (2003), which is in a closed conformation. However, the presence of an acetate molecule and another unknown molecule (modelled as N-carboxyalanine) in the active site suggest that this structure could be another "inhibitor-bound" closed form, rather than a true unliganded state (Natesh *et al*, 2003; Towler *et al*, 2004).

Further evidence for a hinge-mechanism in ACE is the indication from microcalorimetry experiments that the energetics of inhibitor binding by ACE are largely entropically-driven (Andújar-Sánchez *et al*, 2004). This could be expected in the case of a mechanism such as hinge-closing where large numbers of solvent-exposed residues become associated with protein residues during binding.

1.11. Conclusions

Human ACE is a C-terminal peptidyl-dipeptidase that occurs in most tissues, either in a membrane-anchored or in a soluble form. It belongs to the M2 family of zinc metallopeptidases, which bear the HEXXH + E zinc-binding motif. Among a range of substrates, it acts on AngI and BK, which are important regulators of blood pressure via the renin-angiotensin and kallikrein kinin systems, respectively. This activity has made ACE the target of focussed drug-development studies.

Our understanding of the structure and function of ACE has come a long way since the development of the first-generation inhibitors. Human ACE occurs in two isoforms, two-domain sACE, and tACE, which is identical to the C-terminal domain of sACE. The two domains of sACE (which are 55% identical) have been shown to exhibit different binding specificities, chloride activation and thermostability. Notably, the C-domain alone appears to be largely responsible for the conversion of AngI to AngII, which constitutes the hypertensive effect of ACE. This suggests the possibility of designing therapeutic inhibitors that are specifically targeted to one domain or the other, with a corresponding increase in the specificity of the response. For example, a C-domain-selective inhibitor that allowed the continued cleavage of BK by the N-domain might be expected to alleviate side-effects such as cough and angioedema, which are thought to be associated with elevated levels of BK, while inhibiting the cleavage of AngI by the C-domain (Zaman *et al*, 2002; Acharya *et al*, 2003).

The process of designing such specific inhibitors has recently been aided by the solution of crystal structures of human tACE in complex with lisinopril, captopril and enalaprilat, and of two homologues, *Drosophila* AnCE and human ACE2. These structures have given us a clear view of ACE as an ellipsoid, mostly α-helical protein, divided into two subdomains by a deep active-site cleft.

The active site is strongly homologous; with respect to sequence and structure; to those of other zinc metallopeptidases, suggesting that the catalytic mechanism is conserved. Differences between the active sites of tACE, AnCE, ACE2, and the N-domain of sACE have allowed some deductions to be made as to the factors contributing to substrate-specificity, which should in turn allow for the development of new domain-specific inhibitors of sACE.

The crystal structures have also revealed the presence of chloride anions in the structure of tACE and ACE2, consistent with the chloride-activation observed in these enzymes.

The structure of unliganded ACE2 reveals a hinge-bending mechanism that closes the active site in the substrate-bound form. This mechanism may be applicable to tACE, as it has also been observed in several other zinc metallopeptidases.

Further investigations are required into such problems as that of the influence of the two domains of sACE on each other; since different studies seem to point either to cooperativity or a lack thereof; or the mechanism of substrate entry into the inaccessible active site of tACE. The mechanism and role of chloride-activation of the C-domain of sACE also remain unclear.

Light could be shed on these questions by the elucidation of additional crystal structures, for example of a truly unliganded form of tACE, the N-domain of sACE or the whole enzyme. In the absence of such crystal structures, the application of computer modelling studies could reveal insights based on the data currently available.

CHAPTER 2: MATERIALS AND METHODS

2.1. Sample preparation:

Glycosylation mutants

Minimally glycosylated mutants of tACE, in which selected glycosylation sites are knocked out by Asn-Gln mutation, were previously constructed and expressed in Chinese hamster ovary (CHO) cells in Prof. Sturrock's laboratory (Gordon *et al*, 2003). Mutants tACE-G3 (lacking all but the third glycosylation site), tACE-G1,3 (lacking all but the first and third glycosylation sites), and tACE-G1234 (lacking all but the first four glycosylation sites) were selected for expression and purification, based on their ACE activity, and on their high levels of expression in CHO cells (figure 2.1.1).

tACE mutant expression

tACE mutants were expressed in CHO cells (Ehlers *et al*, 1991b). Mutant cell lines were grown in 10% FCS medium (50% DMEM (Highveld Biological, South Africa), 50% Ham's medium (Highveld Biological, South Africa), supplemented with 10% foetal calf serum (Invitrogen Life Technologies, United Kingdom)). Expression was induced with 40μM $ZnCl_2$ in 2% FCS medium (50% DMEM, 50% Ham's medium supplemented with 2% foetal calf serum). Mutant tACE-G1234 was also expressed in the presence of glycosidase 1-inhibitor, N-butyl-deoxynojirimycin (NB-DNJ; Toronto Research Chemicals Inc., Canada) at a concentration of 1.5mM.

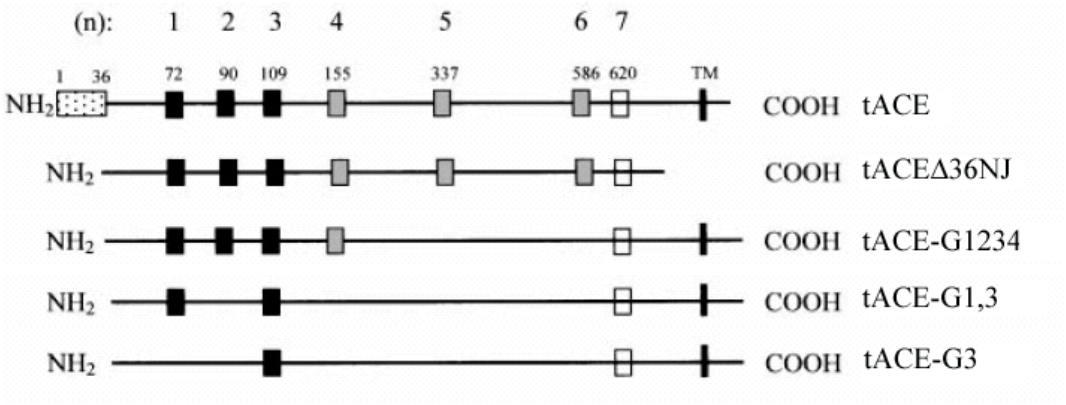


Figure 2.1.1. Schematic diagram of tACE glycosylation mutants. (Gordon *et al*, 2003)
tACE: full-length tACE, with N-terminal signal sequence (1-36) and C-terminal membrane anchor (TM). N-linked glycosylation sequons are represented by blocks. Black blocks indicate full glycosylation, grey blocks indicate partial glycosylation and white blocks indicate no glycosylation. **tACEΔ36NJ:** the truncated native tACE that was crystallised after expression in the presence of glycosidase-I inhibitor, NB-DNJ.
tACE-G1234, tACE-G1,3, tACE-G3: glycosylation mutants used in this study, lacking the N-terminal signal sequence but having the transmembrane domain.

Purification of tACE mutants by lisinopril affinity chromatography

Soluble tACE glycosylation mutants were purified from harvested medium by lisinopril affinity chromatography (Ehlers *et al*, 1989, 1991b). Briefly, pooled medium was applied to sepharose-28-lisinopril affinity columns and unbound protein was eluted with wash buffer (0.5M NaCl, 20mM HEPES, pH7.5). tACE mutants were eluted as a single peak using 50mM borate, pH 9.5. Fractions were assayed for protein concentration and ACE activity, and those containing ACE were pooled.

Buffer exchange and sample concentration

Pooled fractions were exchanged from borate buffer into 5mM HEPES, and concentrated by centrifugation using an Amicon Ultra-4 spin column (Millipore). Buffer exchange was either accomplished by dialysis followed by concentration, or by simultaneous concentration and buffer exchange during the centrifugation step. Volumes were reduced to a final sample concentration of 2-5mg/ml.

The purity of the concentrated samples was assessed by 10% denaturing SDS-PAGE with Coomassie staining. Samples were loaded in ~5μg and ~12μg amounts, in the first case to check the molecular weight and in the second case, overloaded to check for the presence of contaminants.

2.2. Assays:

ACE activity assay

ACE activity of the harvested medium, elution peak fractions, column flow-through and dialysate was determined by measuring the hydrolysis of Hip-His Leu (Sigma) by a fluorimetric assay (Friedland & Silverstein, 1976).

Briefly, samples were incubated for 15 minutes at 37°C in the presence of 5.65mM Hip-His-Leu, 5.2mM NaOH, 0.052M HEPES (pH 7.5) and 0.31M NaCl, where after the hydrolysis reaction was stopped by the addition of 750μl 0.28M NaOH. This mixture was then incubated at room temperature for ten minutes in the presence of 8mM o-phthaldialdehyde, allowing the formation of the fluorescent adduct of o-phthaldialdehyde and L-His-L-Leu. 3N HCl (100μl) was added to stop the reaction,

and the amount of L-His-L-Leu present was determined by fluorimetric analysis (excitation wavelength, 360nm; emission wavelength, 485nm) and comparison of the fluorescence with that of L-His-L-Leu standards. 1mU of ACE activity was defined as the production of 1nmole L-His-L-Leu per minute.

Protein Concentration

The concentration of the various protein samples was determined by a Bradford protein assay (Bio-Rad protein micro-assay).

2.3. Crystallisation

Crystals were grown by the hanging drop vapour diffusion method. A narrow range of crystallisation conditions were screened, based on those determined for minimally glycosylated native tACE (Gordon *et al*, 2003). Reservoir buffer consisted of 15% polyethylene glycol 4000 (Fluka), 10μM $ZnSO_4$ and 50mM sodium acetate (Merck) at pH 4.5, 4.7 or 4.9, filter-sterilised (0.22μm filter).

Samples of purified tACE mutants were centrifuged for 5 minutes in a bench-top centrifuge to clear precipitated material before being used for crystallisation trials. Protein concentration was re-assessed by Bradford assay before each screen was set up. The samples were then combined at varying concentrations in a ratio of one part protein solution to one part reservoir buffer in 4μl drops, sealed over 1ml reservoirs in Linbro plates (Hampton) and incubated at 16°C in a low temperature incubator (United Scientific). In order to reduce vibration from the incubator refrigeration unit, crystallisation trays were placed on a 2cm layer of low-density foam.

To prevent water vapour from condensing on cover slips, screens were set up in a room with an ambient temperature of 18-19°C, all solutions were pre-incubated at 16°C, and Linbro plates were incubated for several minutes at 16°C after every 4-6 wells had been set up. Screens were also carried out with 500μl of various combinations of paraffin and silicon oils (Hampton) covering the reservoirs (Chayen, 1997).

2.4. Data Collection and Processing

Data Collection

X-ray diffraction data were collected at 100K from a single crystal of tACE-G1,3 (visible dimensions 0.4 X 0.05mm), grown using a reservoir buffer of 15% PEG, 50mM NaAc, pH4.5, 10μM $ZnSO_4$, with 500μl of a 3:2 combination of paraffin to silicon oils over the reservoir. The crystal was soaked briefly in 80% reservoir buffer, 20% glycerol and flash-frozen in the cryostream prior to data collection.

Data were collected to a nominal resolution of 2.9Å at the in-house X-ray source in the Department of Biotechnology, University of the Western Cape [comprising a Rigaku RUH3R copper rotating-anode X-ray source operated at 40 kV, 22mA; a Rigaku R-axis IV+ image plate camera; an X-stream 2000 low-temperature system; and an AXCOPX50 glass capillary optic with a 0.1 mm focus]. The crystal-to-detector distance was set at 102mm, and data frames were recorded over 1100-second exposures, with an oscillation angle of 0.5°, from 60°-225°, covering a total rotation of 165.5°.

Data processing

From these images, 308966 diffraction spots, to 2.9Å, were integrated, scaled and merged to 14275 unique reflections, using DENZO/SCALEPACK (Otwinowski & Minor, 1997).

Molecular replacement was carried out using *EPMR2.5* with the structure of minimally glycosylated native tACE (PDB ID, 1o8a) as a model (Kissinger *et al*, 1999; Natesh *et al*, 2003). A resolution range of 15.0-4.0Å was used.

Phase refinement was initially carried out on 11181 reflections to 3.2Å, using CCP4i v5.0 (Collaborative Computational Project, Number 4, 1994). The refinement procedure consisted of one cycle of rigid body refinement, followed by 20 cycles of least squares restrained refinement using the maximum-likelihood method as implemented in *REFMAC5* (Murshodov *et al*, 1997). After this first refinement step, a further 3094 reflections were added, extending the resolution to 2.9Å. Additional

rounds of model building and phase refinement were then carried out. A relative weighting term of 0.3 was initially applied to experimental data. This was decreased to 0.05 towards the end of the refinement in order to improve the geometric parameters of the model (see results section). Experimental sigma's were included in this term. Isotropic temperature factors were refined. A test set, comprising 5% of reflections, was excluded from the refinement for the calculation of R_{free}. Simple bulk-solvent scaling was used.

Model building was carried out using O 7.0.0, using 1o8a as a starting model, with $2F_{obs}$-F_{calc} and F_{obs}-F_{calc} maps. After several steps of model building and refinement, a map was calculated by prime-and-switch phasing using RESOLVE and a round of model-building was performed using this map as a guide in order to reduce model bias (Terwilliger, 2001). This was followed by further rounds of refinement and model-building, carried out as before (see results section for details).

Zinc and chloride ions were added when clear density became apparent in the F_{obs}-F_{calc} map. Water molecules were added gradually to the model, manually and using ARP/wARP (subject to manual examination), being positioned where well-defined positive peaks were present in both the $2F_{obs}$-F_{calc} and maps, and where they could form hydrogen bonds with protein, glycan or other water molecules.

Co-ordinates of the glycan residues were obtained from HIC-UP and added to the model when clear density appeared at the glycosylated residues. Water oxygens that had been modelled into disordered or partially-ordered regions of the glycan chains were identified based on their location in density that was either connected with or proximal to that occupied by ordered glycan residues. These atoms were initially refined as water molecules, but later deleted, since the density they occupied did not resolve into clear glycan residues.

Multiple conformations of the side-chains of several surface residues were introduced when successive rounds of model-building and refinement using different conformations failed to resolve a single high-occupancy rotamer, while at the same time density was present that suggested the simultaneous presence of another rotamer

in the structure. Side-chains were truncated in cases where no density was resolved after several rounds of model-building and refinement.

Iterated refinement with model-building was carried out until no further decrease in R_{free} was achieved. During the final rounds model-building and refinement, PROCHECK was used to identify residues that had main-chains or side-chains in unfavourable conformations, and these were corrected.

Once the structure had been refined to convergence (no further information was apparent in successive rounds of refinement), the test set of reflections was included in the data set and a further pass of refinement was carried out, followed by one model-building phase and a last refinement step.

Structure Validation and Analysis

Structure analysis was carried out using CCP4i: PROCHECK and SFCHECK were used to analyse the geometry and agreement of the structure with the experimental data. Lattice contacts were identified using CONTACT.

Secondary structure was assigned using the secondary structure definition program of Fodje (2002).

Alignment with wild-type tACE (1o8a) was determined for C-α residues, using ALIGN (Cohen, 1997). All-atom pair-wise alignments of active site residues and chloride ligands were carried out using PyMOL (DeLano).

2.5. Identification of hinge regions in ACE2

The open and closed structures of ACE2 (ACE2o, PDB ID 1r4l, and ACE2c, PDB ID 1r42, respectively) were compared in order to identify residues involved in the hinge-bending mechanism.

Structure alignments and root mean square deviations for Cα atoms were computed using ALIGN (Cohen, 1997). Hinge regions in ACE2 were defined as those residues to one side of which the Cα atoms of the ACE2o deviated from those of ACE2c by

>3Å, and to the other side of which the structures were closely aligned. These residues were compared with the corresponding residues in tACE (PDB ID 1o8a) and AnCE (PDB ID 1j38), based on the alignment of Guy *et al* (2003).

2.6. Analysis of temperature factors

As an indication of the order or disorder of the regions, temperature factors (B-factors) were obtained from the PDB entries. <B> and σ(B) were calculated for all Cα atoms, and the B-factor of each individual Cα atom was expressed as a function of σ(B).

2.7. Modelling of a putative open form of tACE:

In order to perform NMA of an open form of tACE, open and closed homology models of tACE (tACEo and tACEc) were built based on ACE2o and ACE2c.

Amino acid sequences were aligned according to Towler *et al* (2004) and 25 open and 25 closed models were generated using MODELLER6v2 (Sali & Blundell, 1993). Procheck was used to analyse the models. Based on their Ramachandran plots, one model was selected from each group as having the fewest residues falling into unfavourable regions. SCWRL was used to optimise the side-chain packing of these models (Canutescue *et al*, 2003).

2.8. Normal mode analysis:

Normal mode analysis (NMA) is a statistical technique for exploring the intrinsic flexibility of a molecule. The basis of this technique is the assumption that molecular motions can be described as simple harmonic oscillations about a local energy minimum (Van Wynsberghe *et al*, 2004). The molecule's potential energy surface is expressed as a matrix (the Hessian matrix) of the second derivatives of its potential energy function, in terms of its mass-weighted atomic co-ordinates (Tama *et al*, 2000). The Hessian matrix is then diagonalised, yielding eigenvectors describing directions of collective atomic oscillation, and eigenvalues expressing the frequencies of these oscillations (Tama *et al*, 2000; Van Wynsberghe *et al*, 2004).

Numerous studies have shown that large-scale macromolecular motions, such as the hinge mechanism observed in ACE2, can be described accurately by one, or a combination of a few, of the low frequency vibrational modes calculated by this method (for example, Brooks & Karplus, 1983; Mouawad & Perahia, 1996; Jääskeläinen *et al*, 1998; Ma & Karplus, 1998; Tama & Sanejouand, 2001; Shen *et al*, 2002; Van Wynsberghe *et al*, 2004).

A potential energy function using an empirically-based molecular mechanics force field has the advantage of allowing the study of a macromolecule in its native electrostatic environment, as well as the determination of probable amplitudes of motion in the absence of a crystal structure (Van Wynsberghe *et al*, 2004). However, the complexity of such a potential function makes NMA of macromolecules computationally intensive. Consequently, a number of approximations have been introduced that reduce the computational time required for NMA. Among these, the block algorithm for diagonalisation of the Hessian matrix treats the residues in a protein in rigid groups, with results being similar to those obtained by a more thorough treatment of the problem (Tama *et al*, 2000). The elastic network model (ENM), using a simplified force field that describes interactions between spatially-related Cα-atoms in terms of a simple Hookean spring potential, has likewise been shown to yield accurate results, without the requirement of an energy-minimisation step prior to NMA (Tama & Sanejouand, 2001). These approximations do not appear to have a detrimental effect on the correlation between the low-frequency modes obtained and experimentally-derived structures (Tama & Sanejouand, 2001). Moreover, they allow the rapid determination of protein flexibility and the identification of the regions important in these motions (Suhre & Sanejouand, 2004).

NMA of ACE2 and tACE, using the simplified elastic network model force field and a block diagonalisation algorithm, was performed using elNémo (Suhre & Sanejouand, 2004). The block size was set to three residues, and the Cα-Cα linkage cutoff to 8Å, being the recommended default values for these parameters.

The crystallographic structures of ACE2o, ACE2c (minus the collectrin-like domain) and unliganded wild-type tACE were analysed, as well as tACEo and tACEc. In each

case the ten lowest-frequency normal modes were calculated, and the degree of cumulative overlap between the normal modes and the structural change was determined. Overlap is a measure of the degree of similarity between the direction of the observed conformational change and the one described by the normal mode in question, with a value of 1 indicating complete agreement between the normal mode and the observed change (Valadié *et al*, 2003).

For the modes showing the highest overlap with the observed structural change, the amplitude of perturbation required for maximal agreement between the resulting structure and the comparison structure was determined, and perturbed models were generated accordingly. Cα-Cα distance fluctuations were calculated by elNémo and used as an indication of the residues involved in the motion described by the mode in question.

The normal modes were also assessed in terms of collectivity, a measure of the proportion of atoms displaying large-amplitude displacements in the motion described by the mode (Suhre & Sanejouand, 2004).

CHAPTER 3: CRYSTAL STRUCTURE OF A GLYCOSYLATION MUTANT OF HUMAN TESTIS ACE

3.1. Sample Preparation:

Samples were eluted from the lisinopril affinity column as a single peak. SDS-PAGE demonstrated that all samples were of the correct molecular weight (~75kDa) and did not contain protein contaminants (figure 3.1.1). Mutants that were glycosylated to different degrees could be seen to have different molecular weights (e.g. tACE-G1234 versus tACEG3). tACE-G1234 expressed in the presence of NB-DNJ, appeared as a sharper band than that expressed in the absence of NB-DNJ, indicative of the homogeneity of glycosylation resulting from the action of this glycosidase 1-inhibitor.

3.2. Crystallisation

An initial crystallisation screen revealed extensive, concentration-dependent, granular precipitation for tACE-G1234 (with or without NB-DNJ) and tACE-G3 under all of the conditions screened. Starting concentrations screened ranged from 0.7-6.0mg/ml for tACE-G1234 (+NB-DNJ), 0.8-4.0mg/ml for tACE-G1234 (-NB-DNJ), and 0.4-2.0mg/ml for tACE-G3.

In this same initial screening, crystals of tACE-G1,3 grew at pH4.5, 4.7 and 4.9, with a 1:1 combination of silicon and paraffin oils over the wells, in drops with starting protein concentrations ranging from ~0.6-3.0mg/ml. At starting concentrations of 3.0mg/ml, numerous tiny-to-medium-sized crystals grew and twinning was evident, while at 0.6mg/ml crystals were sparse and did not grow large enough for diffraction.

Further screens were set up using 1:1, 3:2 and 2:3 ratios of silicon to paraffin oils over the reservoirs. The presence of the oils was seen to reduce the amount of sample precipitation, as well as the amount of water vapour condensing on the cover slides, with a combination of 2 parts silicon to 3 parts paraffin oil favouring the development of few, large crystals.

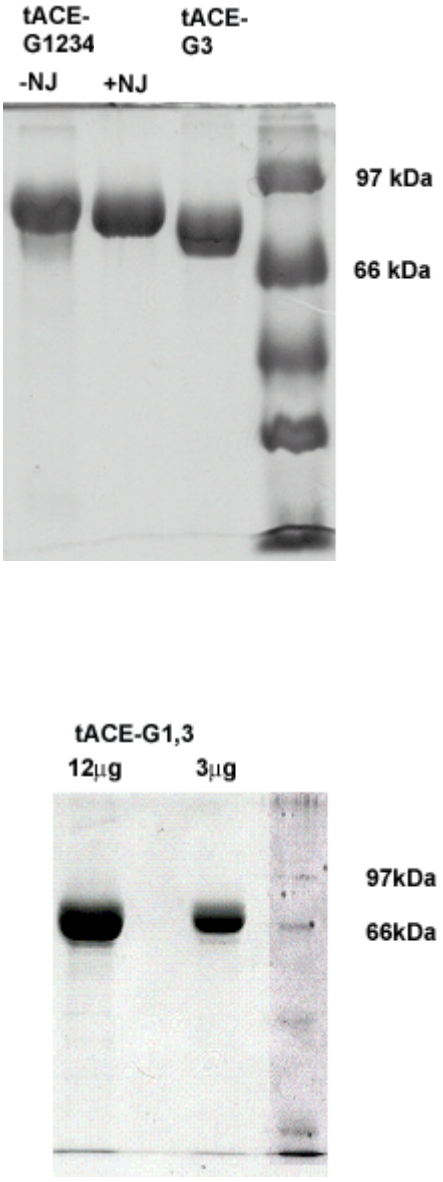


Figure 3.1.1. SDS-PAGE with Coomassie staining to determine sample homogeneity. Wells were overloaded for assessment of sample purity and approximate molecular weight. +/-NJ: tACE-G1234 sample expressed in the presence or absence of N-butyl-deoxynojirimycin. The molecular weight of tACE is ~75kDa.

tACE-G1,3 crystals of a suitable size and separation were grown in wells with a starting protein concentration of 1.5mg/ml, having 2 parts silicon to 3 parts paraffin oil over the reservoir, at pH4.5, 4.7 and 4.9. While the crystals grown at pH4.9 were generally larger than those grown at a lower pH, they also showed large cracks. All of the crystals had the same morphology, and similar results were obtained for samples of tACE-G1,3 from independent batches of expression and purification.

Data was collected from a crystal grown at pH4.5 with dimensions (visible in the plane of the drop) of ~0.4 X 0.05mm.

3.3. Data collection and processing

Data collection

Data collection statistics are shown in table 3.3.1. tACE-G1,3 crystallised in the same space group as wild-type tACE (Natesh *et al*, 2003). However, the unit cell dimensions differ slightly from those of the structures of tACE solved previously, being ~3Å longer in the a-axis and ~2Å longer in the c-axis (table 3.3.2).

At the crystal-to-detector distance of 102mm , a zone of diffraction spots was lost during integration, due to overlap of integration boxes. Despite the loss of these reflections, the wide oscillation range ensured that sufficient observations were made for each unique reflection, as can be seen from the redundancy (3.9). A longer crystal-to-detector distance was therefore not required.

Data were initially scaled and merged to 3.2Å, yielding 11181 unique reflections with an overall R_{merge} of 18.3% (30.29% in the highest resolution shell). After an initial step of phasing, phase refinement and model-building, however, 3094 additional reflections were included up to 2.9Å resolution. Although R_{merge} for the 3.02-2.90Å shell was 40.8%, these reflections were included based on their reasonable signal-to-noise ratio and redundancy.

Table 3.3.1. Reflection data collection, scaling and merging. Values in parentheses apply to the highest resolution shell (3.0-2.9Å).

Space group	P212121
Mosaicity	0.4
Resolution range (Å)	20.0-2.9
Completeness (%)	96.7 (99.0)
R_{merge} (%)	20.0 (36.2)
$\langle I/\sigma(I) \rangle$	5.88 (1.86)
$\langle I/\sigma(I) \rangle$ (systematically absent reflections)	0.5
Redundancy	3.9 (3.0)
No. of observations	308966
No. of unique reflections	14275
Unit-cell dimensions (Å)	a = 59.81, b = 85.17, c = 135.58

Table 3.3.2. Comparison of unit cell dimensions (in Ångströms) of tACE-G1,3 with those of the published structures of tACE. PDB accession codes are given in brackets for each structure.

Structure	**a**	**b**	**c**	**Volume ($Å^3$)**
Unliganded tACE (1o8a)	56.620	85.060	133.790	644345
Lisinopril-bound tACE (1o86)	56.470	84.470	133.990	639135
Enalaprilat-bound tACE (1uze)	56.720	85.351	133.726	647382
Captopril-bound tACE (1uzf)	56.648	84.899	133.500	642049
tACE-G1,3	59.810	84.900	135.990	690539

3.4. Phase refinement and model-building

Molecular replacement yielded a single solution having a correlation-coefficient of 0.703 and an R-value of 0.565, indicating the presence of only one molecule in the asymmetric unit. This was as expected from the similarity of the unit cell dimensions to those obtained previously (table 3.3.2).

Model building and phase refinement were alternated in a number of build-refine steps. After each model-building phase, cycles of least-squares restrained refinement were carried out until the R_{free} value ceased to decrease. Refinement statistics that were monitored throughout the phase-refinement and model-building process are presented in figure 3.4.1. It should be noted that because of the high degree of similarity between the molecular replacement solution and the final structure, the initial phasing was accurate and thus the changes observed in these quality-control statistics over the course of the experiment were slight.

Initially, phase refinement was carried out to 3.0Å, by one cycle of rigid body refinement, followed by 20 cycles of least-squares restrained refinement. During The subsequent model-building phase, the mutated asparagines were replaced with glutamine residues and their side-chains adjusted to fit the electron density maps. Density was present in the F_{obs}-F_{calc} map at the zinc and chloride binding sites, allowing the insertion of these ions into the model (figure 3.4.2).

After this step, additional reflections were added to the data set to extend the resolution to 2.9Å. This was followed by 8 cycles of restrained refinement using the partially-adjusted model with mutated residues and zinc and chloride ions. During the subsequent model-building phase, the N72- and N109-linked glycan residues, as well as additional C-terminal and loop residues, were modelled in where clear density appeared in the F_{obs}-F_{calc} map (figure 3.4.3).

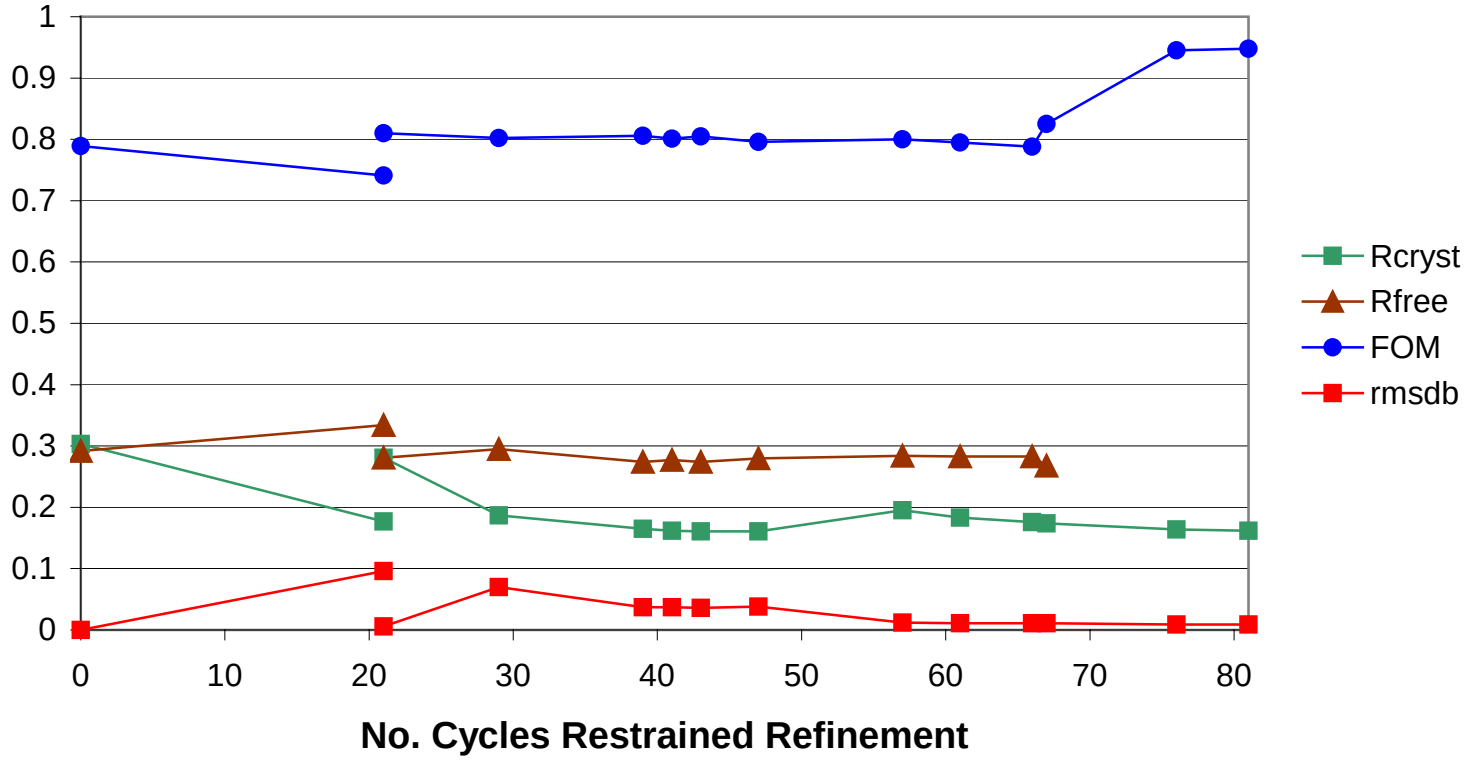


Figure 3.4.1. Refinement statistics. Each data point represents the refinement statistics at the end of one step of model-building and refinement.
Rcryst = crystallographic R-factor; Rfree = free R-factor; FOM= mean phasing figure of merit; rmsdb = r.m.s. deviation of bond length from ideal values.

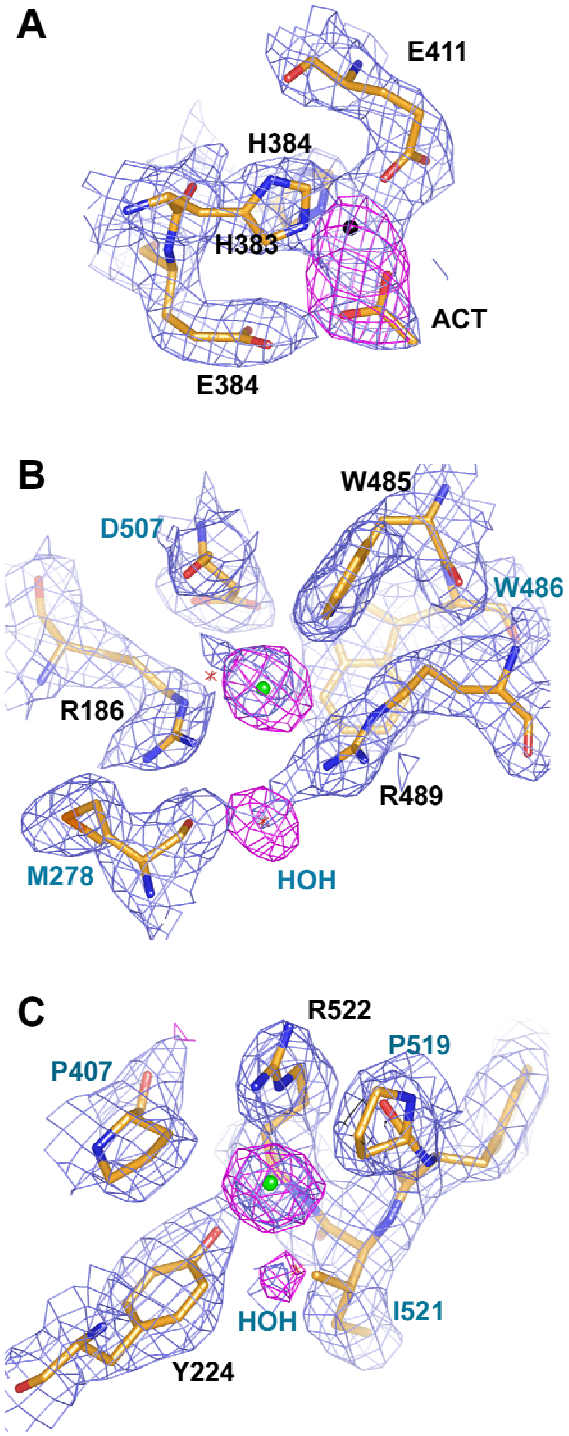


Figure 3.4.2. $2F_{osbs}$-F_{calc} and F_{obs}-F_{calc} maps after the first refinement step, showing the positions of **A)** the active site zinc ion (black) and acetate molecule (ACT), **B)** chloride ion CL1 (green) and associated water molecule, **C)** chloride ion CL2 (green) and associated water molecule. Blue density: $2F_{obs}$-F_{calc} map; magenta density: F_{obs}-F_{calc} map. The final model of tACE-G1,3 is shown in the density. Black labels indicate chloride or zinc ligands.

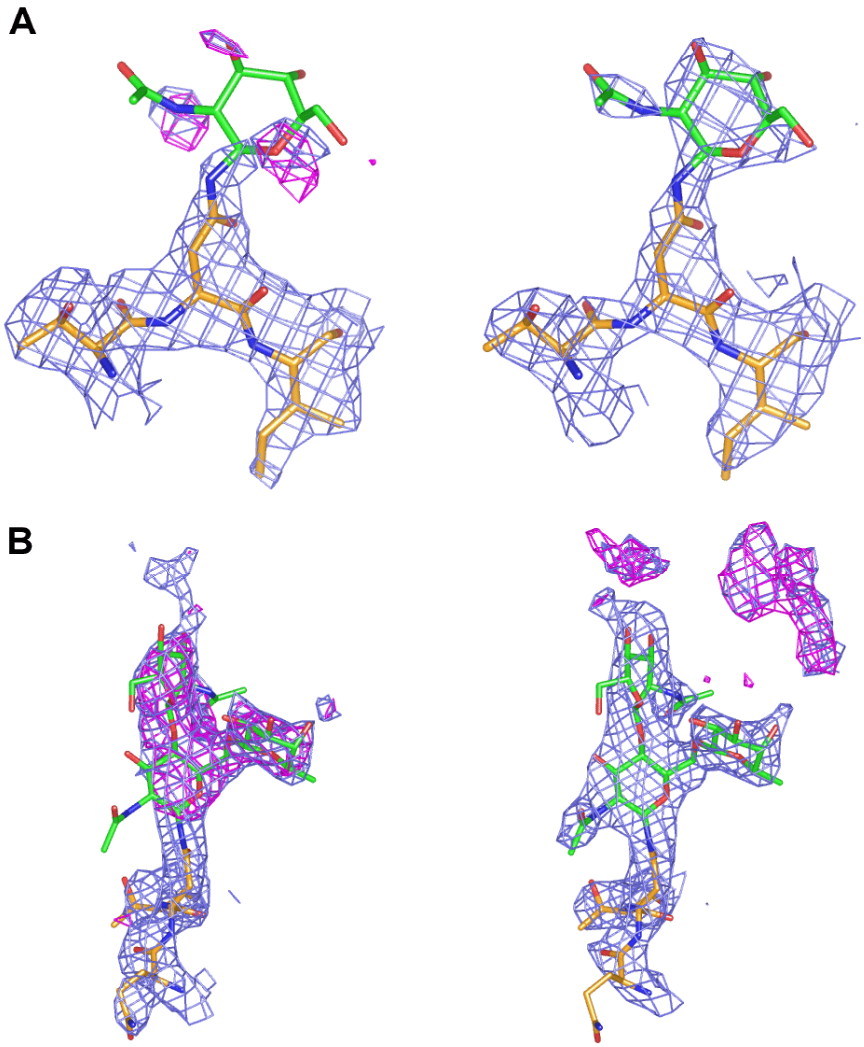


Figure 3.4.3. Modelling of N-linked glycan residues at **A)** N72 and **B)** N109. Left-hand-side: omit maps after the first round of refinement; right-hand side: maps after the final round of refinement.

Density for the N109-linked glycan was present in the omit-maps from the first refinement step, with further unmodelled density being evident after the final round of refinement, indicating the presence of more ordered or partially-ordered atoms in the crystal structure than were built into the model. At the N72-linked glycosylation site, however, only a limited amount of ordered density was present for the N72-linked glycan, indicating a greater degree of disorder at this site. This density did not appear in the initial omit maps, and the residue could only be modelled in after several steps of refinement and model-building.

The next three steps of refinement and model-building (steps 3-5, 39-43 cycles of restrained refinement in figure 3.4.1) were carried out using ARP/wARP to add water molecules during the refinement phases. Each water molecule was examined manually and only those found to be in reasonable positions and within H-bonding distance of a protein, water or glycan residue were kept. ARP/wARP was abandoned when no further sensible water molecules could be added automatically, and a refine-build step (step 6) was carried out with the manual addition of water molecules. In addition to the water molecules, one glycerol molecule was placed in the active cleft.

This was followed by a phase modification step, to reduce model bias. Phases were modified by the prime-and-switch technique, in which phases are adjusted according to solvent boundary information, without reference to the starting phases (Terwilliger, 2001). The resulting phase information had a bias ratio of 0.96 (as opposed to 1.66 prior to phase modification), indicating a substantial decrease in model bias. However, the mean figure of merit of these phases was low (0.64). This was probably a result of the low solvent content of the crystals, since the prime-and-switch technique works best for solvent contents of 50% and higher (Terwilliger, 2001). Because of the poor quality of these phases, the subsequent model-building step made use of $2F_{obs}$-F_{calc} and F_{obs}-F_{calc} maps, calculated using the unmodified phases, with the modified phase map as a guide. Adjustments were only made where this map deviated clearly and substantially from the weighted difference map.

Up until this stage (6 build-refine steps, 47 refinement cycles), the relative weighting on the experimental and theoretical terms in the refinement had been set to 0.3, with the result that the geometry of the model had become distorted. This can be seen from the high r.m.s. bond length deviation values (0.038Å) in figure 3.4.1. During the steps following the phase modification step, this problem was corrected by adjusting the relative weighting from 0.3 to 0.05 as was appropriate for data to this resolution. The result was a decrease in r.m.s. bond length deviation (to 0.011Å) and a slight increase in phasing figure of merit (0.796 to 0.800). However, this also resulted in an increase in R_{cryst} from 0.161 to 0.195, since the previous rounds of incorrect weighting had amounted to an over-fitting to the experimental data. R_{free} also increased slightly during these steps (0.280 to 0.284).

During the subsequent two steps of refinement and model building, PROCHECK was used to identify residues having unfavourable main chain and side chain conformations. These were then corrected during model-building. Also during these steps, an active site acetate was added and unresolvable or disordered side chain atoms were deleted. At this stage (9 build-refine steps, 66 refinement cycles), R_{free} appeared to have reached a plateau at 0.283, indicating the convergence of the refinement and model-building procedure.

In the step prior to the penultimate step, water molecules that had been modelled into probable disordered glycan density were identified as such and deleted from the model. The inter-residue links for the glycan residues and zinc ion were also regularised. This resulted in a decrease in R_{free} to 0.268 and an increase in FOM from 0.788 to 0.825.

For the penultimate step, the test set (5%) of reflections was included in six cycles of refinement. A model-building phase was then carried out during which a few minor adjustments were made, and some water atoms added or deleted. This was followed by a final four cycles of refinement. During these last two steps, the

phasing figure of merit increased dramatically, as could be expected due to the inclusion of additional experimental data, while the r.m.s. bond deviation decreased from 0.011Å to 0.09Å.

The final refinement statistics are shown in Table 3.4.1.

3.5. Structure validation:

The Ramachandran plot of tACE-G1,3 output by PROCHECK is shown in figure 3.5.1. 92.9% of non-glycine residues lie inside the most favoured regions of the Ramachandran plot, with 6.5% in additionally allowed, and 0.6% in the generously allowed regions. No non-glycine residues were in disallowed regions.

The G-factor is greater than -0.5, indicating normal stereochemistry, while the r.m.s. deviations of the geometric parameters from ideality are all lower than the target values. From the Luzzati plot, the estimated maximal co-ordinate error for low temperature factor atoms is 0.278Å, and the average temperature factor is low. σ(B) is high, indicating a large variance in temperature-factors. This is largely due to the presence of loop L433-D440 and the glycan residues, which show a high degree of disorder in the crystal lattice.

3.6. The structure of tACE-G1,3

The structure of tACE-G1,3 is very similar to the tACE structures solved previously, with the r.m.s. deviation between the Cα atoms of tACE-G1,3 and the published wild-type structure being only 0.30Å.

Like the other structures, this mutant is largely α-helical, with a few short regions of β-strand (figure 3.6.1). The molecule is ellipsoid in shape and divided down the middle by a deep cleft, in which the active site zinc ion is located. This cleft is closed off from the external milieu by the N-terminal two α-helices or lid-helices.

Table 3.4.1. Refinement statistics at the end of model-building and phase refinement.

Solvent content (%)	46.74
Final R_{cryst} (%)*	16.2
R_{free} (%)#	26.8
R_{free}/R_{cryst}	1.65
Phasing figure of merit	0.948
G-factor	0.08
R.m.s. deviation from ideality (target values)	
Bond length (Å)	0.009 (0.021)
Planar groups (Å)	0.004 (0.020)
Bond angles (°)	1.205 (1.934)
<u> (co-ordinate error from Luzzati plot)	0.273
No. of atoms (protein)	5339 (4823)
No. of solvent atoms	445
No. of zinc ions	1
No. of chloride ions	2
No. of glycan atoms	55
Average temperature factor (A^2)	22.8
σ(B)	11.85

* R_{cryst} after inclusion of the test set of reflections.
# R_{free} prior to the inclusion of the test set.

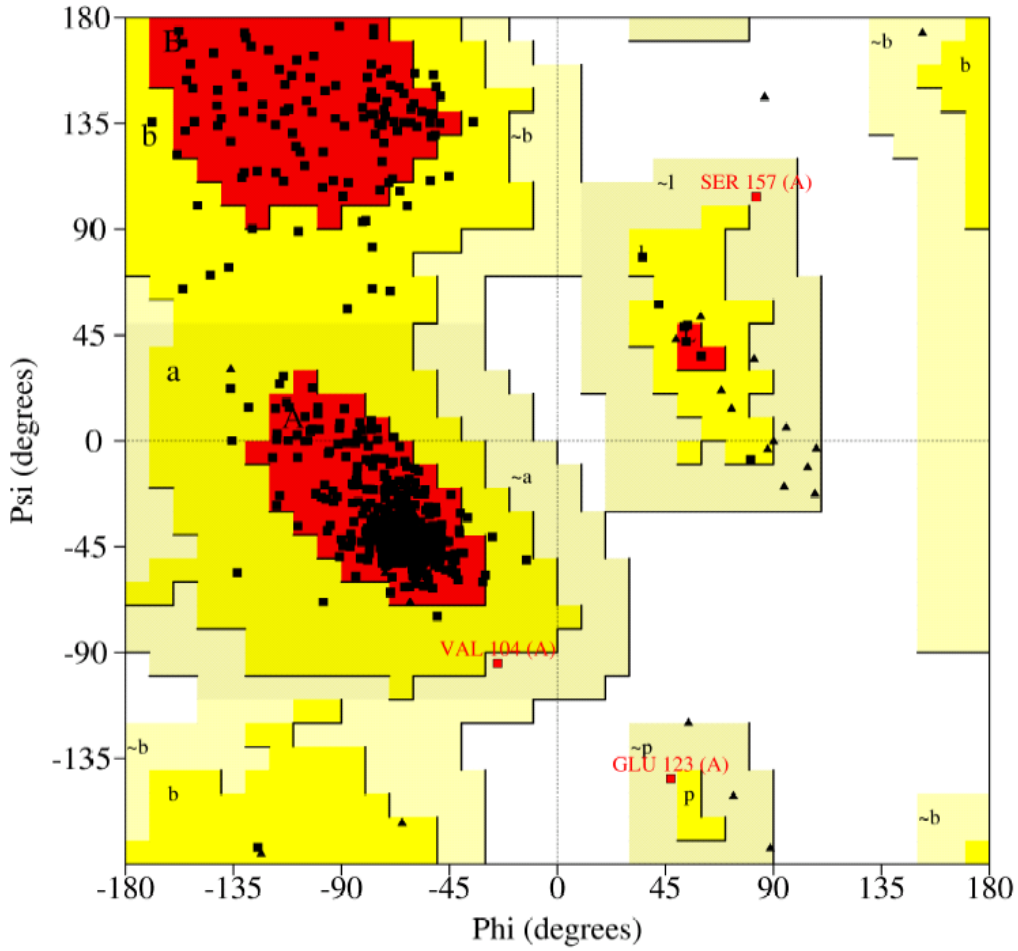


Figure 3.5.1. Ramachandran plot of tACE-G1,3 (PROCHECK). The most favoured regions are coloured red, additionally allowed regions yellow, generously allowed regions tan, and disfavoured regions white. Non-glycine residues are represented by blocks (■) and glycines by triangles (▲). Non-glycine residues falling into generously allowed regions are labeled in red text. No non-glycine residues fell into disallowed regions.

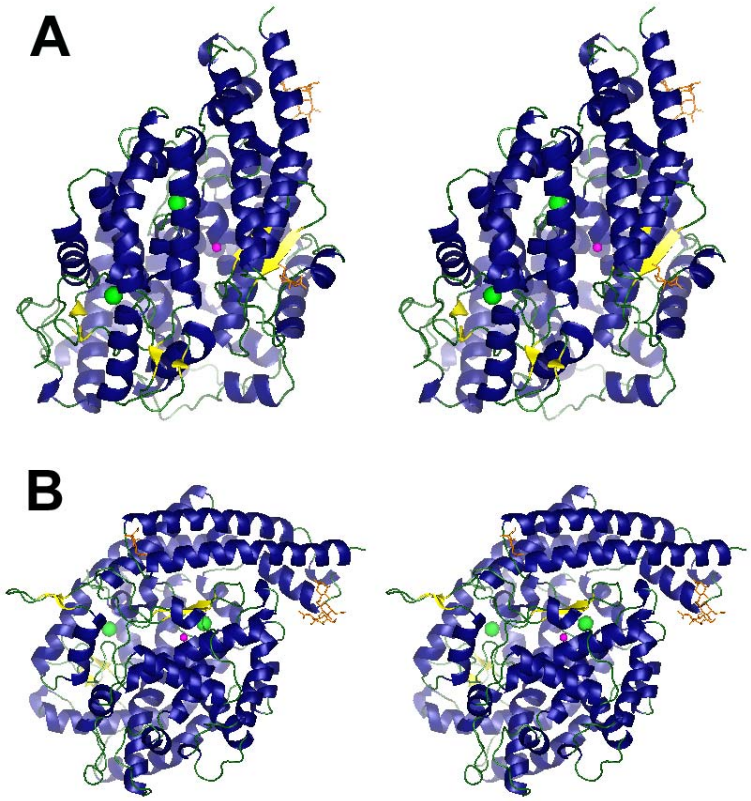


Figure 3.6.1. Overview of tACE-G1,3 structure, stereo views. **A)** View looking down into the active cleft with lid helices to the top right. **B)** Side view with lid helices on top. Navy: α-helices; yellow: β-strands; orange: glycan residues; green spheres: chloride ions; magenta sphere: zinc ion.

Two chloride ions are present in the structure, at the same sites as in wild-type tACE (figure 3.4.2). The r.m.s. deviations for all atoms in the residues surrounding CL1 (D186, M278, W485, W486, R489, D507) and CL2 (W220, Y224, P407, P519, I521, R522) are 0.673Å and 0.652Å, respectively. The conserved water molecules proposed by Tzakos *et al* (2003) to represent an entrance pore for CL2 are also present in this structure.

The active site zinc ion is tetra-coordinated by H383, H387, E411 and the carboxyl oxygen of an acetate molecule (figures 3.4.2 and 3.6.2). The r.m.s. deviation between the zinc-binding residues and those of 1o8a is 0.221Å for all atoms, and 0.181Å excluding the acetate molecule.

The active site acetate makes hydrogen-bonding contact with E384 of the zinc-binding motif, and superimposes onto the zinc-binding carboxyl group of the inhibitor lisinopril in the lisinopril-bound structure (1o8f). In addition to the acetate molecule, the active site of the unliganded wild-type tACE structure contains an N-carboxyalanine moiety, which was modelled into unknown density. This position is occupied by water molecules in tACE-G1,3, as is the position occupied by a glycine moiety that was modelled into unknown density in the active site of the lisinopril-bound structure (figure 3.6.3).

In addition to these two sites, a string of three water molecules in tACE-G1,3 occupy density which may correspond to a ligand which is present in the lattice at low occupancy (figure 3.6.3). However, at 2.9Å resolution, it is not possible to distinguish between such a ligand and water molecules.

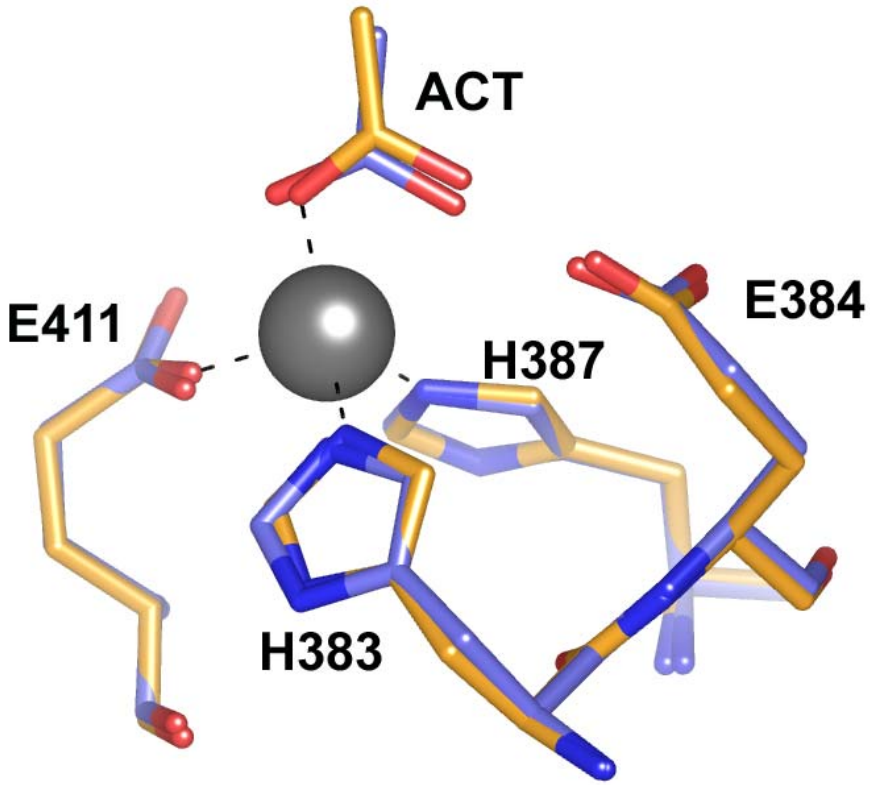


Figure 3.6.2. Alignment of zinc ligands in tACE-G1,3 and wild-type tACE (PDB accession code 1o8a). Yellow sticks: tACE-G1,3; blue sticks: wild-type tACE; ACT: acetate. R.m.s. deviation for all atoms = 0.221Å. R.m.s. deviation for all atoms excluding ACT = 0.181Å.

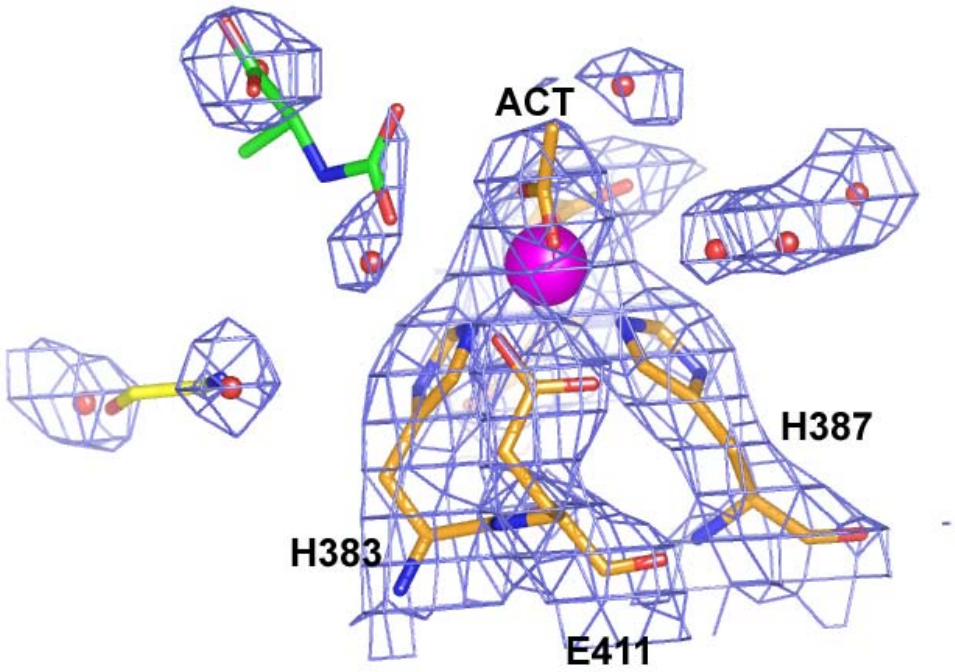


Figure 3.6.3. Unknown density in the active site. Orange sticks: zinc-binding residues in tACE-G1,3; green sticks: N-carboxyalanine from the aligned wild-type structure (1o8a); yellow sticks: glycine from the aligned lisinopril-bound structure (1o8f); magenta sphere: zinc ion; ACT: acetate molecule; red spheres: water molecules in the active site of tACE-G1,3 that may occupy density corresponding to low-occupancy ligand.

3.7. Temperature factor analysis

The residues having the highest temperature factors (greater than 2σ(B)) are located around glycosylation site N109 at the end of lid helix α2, and in the exterior loop region L433-D440, which was disordered in the wild-type structure (1o8a). Other regions having higher-than average temperature factors (greater than σ(B)) include the ends of the lid helices (including the N72-glycosylation site), exterior loop P314-S318, the N-terminal ends of α10 and α11, loop E342-V350 (stretched out into the surroundings and probably interacting with the glycan chain at N72), and the C-terminal loop.

Residues having lower-than-average temperature factors (less than -σ(B)) are all located in the interior of one half of the molecule, subdomain II, surrounding the CL1 binding-site and including chloride and substrate ligands. This is illustrated in figure 3.7.1, where it can be seen by contrast that the residues having high temperature factors are all on the surface, and fall largely into subdomain I. The only high temperature factor residues in subdomain II are at the ends of α4 (flanking the lid helices) and the C-terminal loop residues, which were disordered in the wild-type structure.

3.8. Disordered residues:

Of the four mutation sites, the side-chains of three (Q90, Q155 & Q618) are present in the structure in two (modelled) low-occupancy conformations. Q155 lies in the loop between β1 and β2, a section of β-sheet that stretches out from the molecule into the surrounding medium, probably interacting with the N72-linked glycan. This loop has lower temperature factors in tACE-G1,3 than in wild-type tACE, either because the absence of the glycan at Q155 increases structural stability, or because the presence of a longer glycan chain at N72 stabilises the structure.

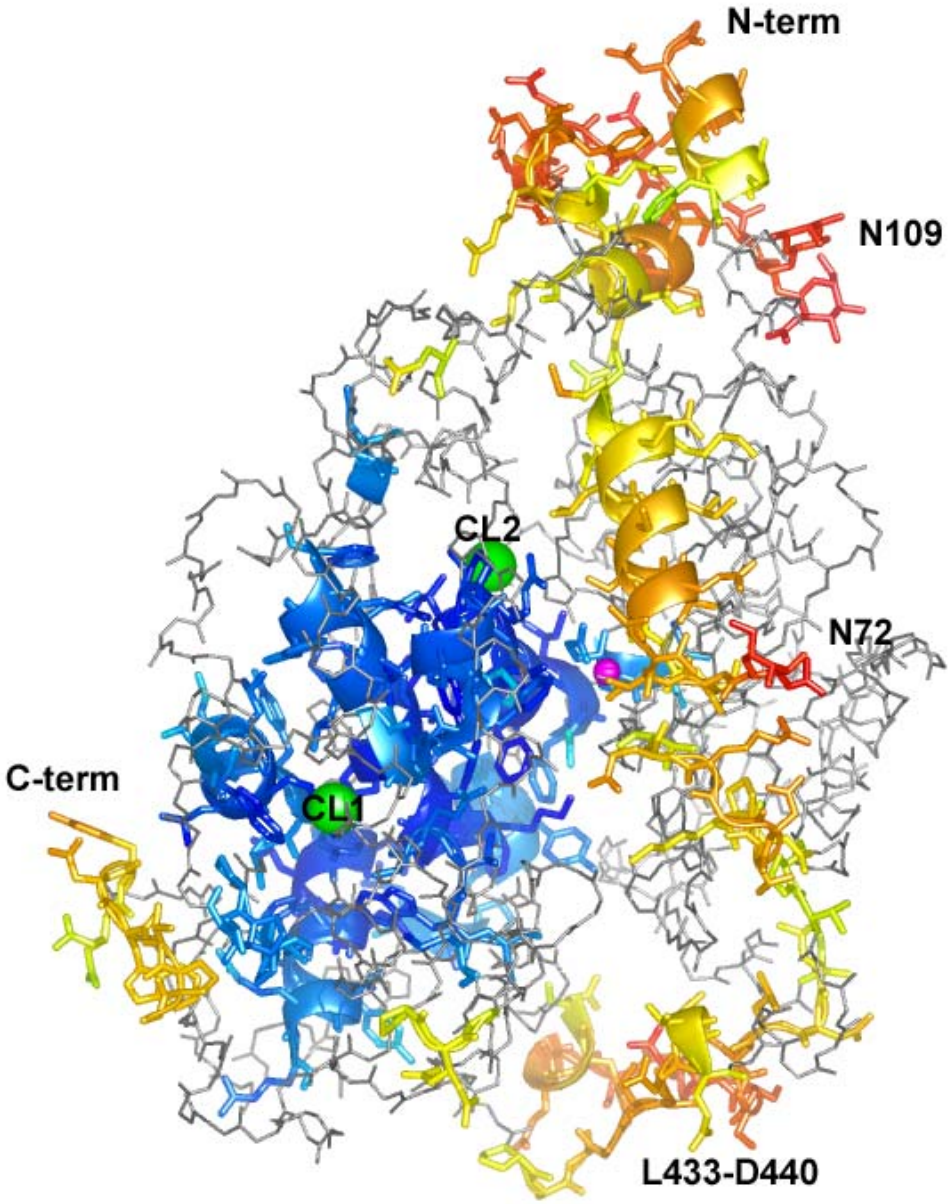


Figure 3.7.1. Temperature factors in tACE-G1,3. View looking down into the active site with the lid helices to the top right. Subdomain II is on the left, and subdomain I on the right. Residues having temperature factors higher or lower than average are shown in cartoon representation, with low temperature factor residues coloured blue and high temperature factor residues coloured yellow-to-red. Chloride ions = green spheres, zinc ion = purple sphere.

Other residues modelled as having two rotameric forms are R53, E143, R314, E403, H543 and Q586. These are all on the exterior surface of the structure, apart from E143 and E403, which are on the interior surface of the active cleft, distant from the zinc ion (16.74Å and 12.29Å, respectively).

Residues having unresolvable side-chain atoms are T39, E43, K46, E64, K79, L107, S295, P297, M299, Q539 and K613, all on the surface of the molecule. S295, P297 and M299 lie on one of two flexible loops that stretch across the end of the active cleft. The second of these two loops, L433-D440, is disordered in the unliganded wild-type tACE structure, but could be modelled into density in tACE-G1,3, although the temperature factors remain high.

3.9. Glycans in the crystal lattice

The unit cell of tACE-G1,3 is slightly larger (690539 $Å^3$) than that of the wild-type structure determined using NB-DNJ (644345$Å^3$; table 3.3.2). This is probably because NB-DNJ-treatment of wild-type tACE led to the production of shorter glycan chains, which occupy smaller volumes than those of tACE-G1,3.

In the crystal structure of tACE-G1,3, 32 residues lie <3.6Å from symmetry-related protein atoms and are thus probably involved in lattice contacts, while in the wild-type structure of tACE, 47 residues were identified as making close contacts. Most of these contact residues lie in subdomain II, the half of the molecule containing zinc and chloride ions and having low temperature factors (figure 3.9.1).

Only one of the contact sites in the wild-type structure involves an ordered glycan residue (NAG696, on N586), while neither of the ordered regions of the glycan chains in tACE-G1,3 are involved in lattice contacts.

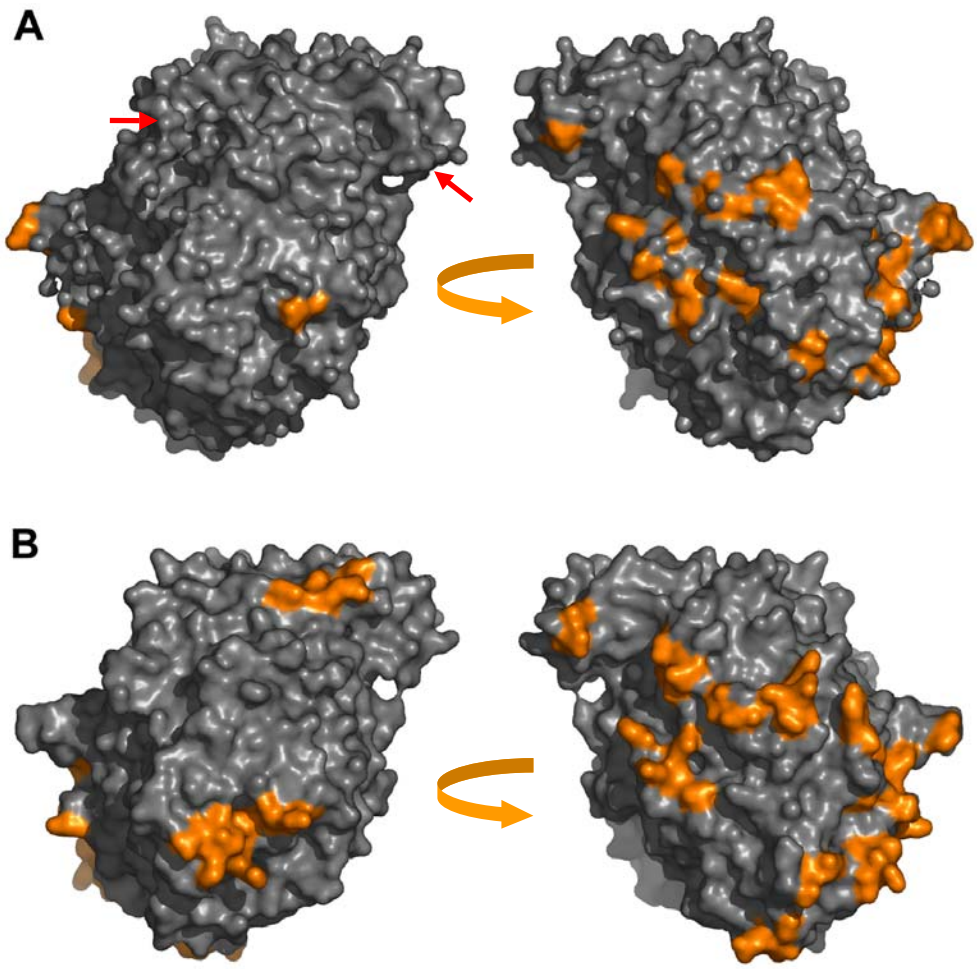


Figure 3.9.1. Lattice contacts: residues closer than 3.6Å to symmetry-related molecules. **A)** tACE-G1,3, with N-linked glycans indicated by red arrows; **B)** wild-type tACE (1o8a). Views are shown from each side of the molecule, with the lid helices on top and the active cleft in the plane of the page. Images on the left-hand-side show subdomain I, while those on the right show subdomain II. Differences in surface texture are largely due to the presence of water molecules in the structures.

Notably, subdomain I, on which side of the molecule the glycan chains of tACE-G1,3 lie, has only one residue making lattice contacts (Q554), as opposed to eight in wild-type tACE (K425, H426, H428-N431, Y553, Q554, NAG696). It is likely, however, that the disordered glycan residues make further contacts in tACE-G1,3.

Differences in lattice contacts between the two structures result not only from an increase in unit cell volume, but also from a slight change in the relative orientations of symmetry-related molecules. This can be seen, for example, in figure 3.9.2, where a channel is visible in the middle of the unit cell, between symmetry-related molecules. This channel is occupied by the N72-linked glycan residues, which take up more space in tACE-G1,3 than in wild-type tACE. Consequently, in tACE-G1,3, the molecules on the right-hand side of the unit cell are shifted slightly across and down relative to those on the left, making the channel larger.

The N109-linked glycan is located in a smaller inter-molecular pocket, which could introduce constraint, resulting in its higher degree of order in the crystal structure.

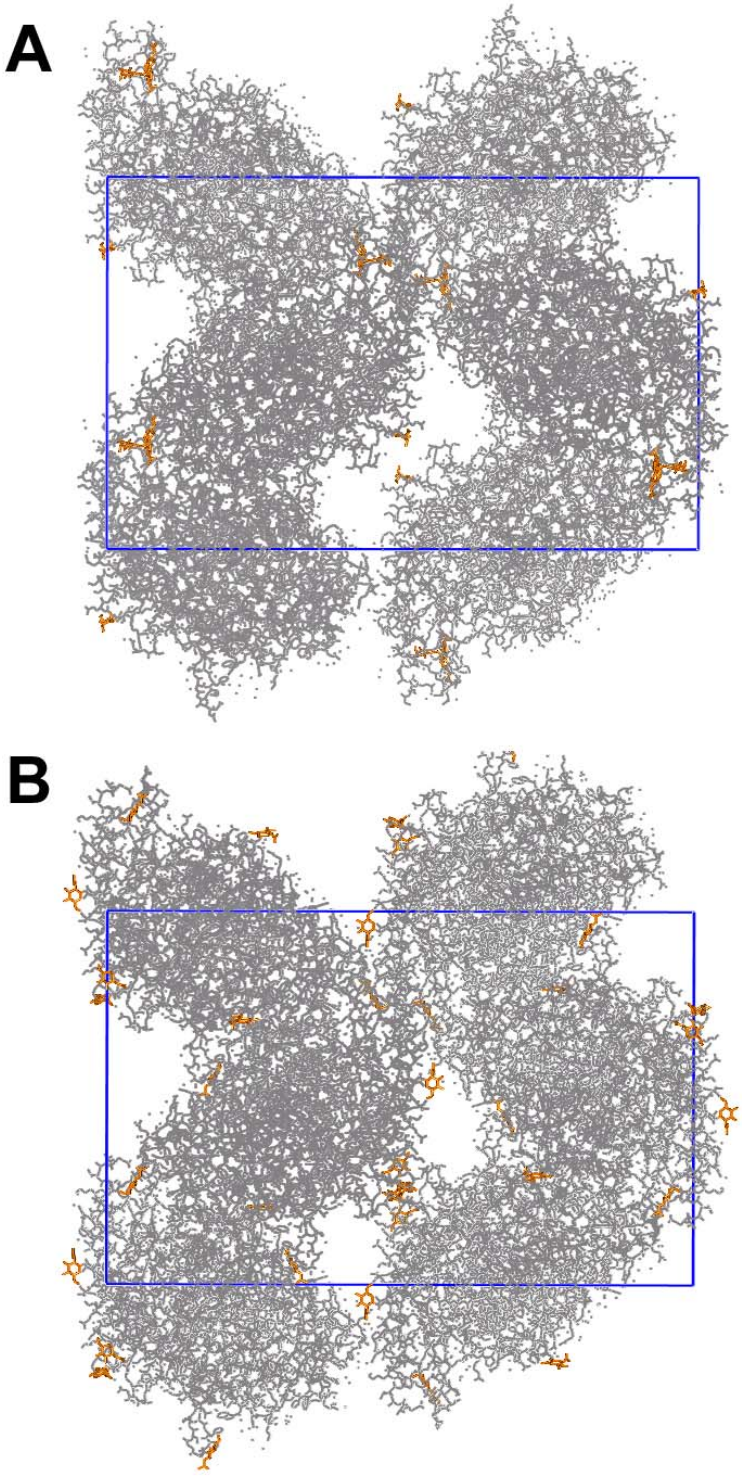


Figure 3.9.2. Lattice packing in **A)** tACE-G1,3; B) wild-type tACE (1o8a). View from the side of the unit cell, with the c-axis horizontal and the b-axis vertical. Glycan residues are coloured orange; the unit cell dimensions are shown in blue. A channel containing N72-linked glycan residues is shown in center of the lower half of the unit cell.

CHAPTER 4: INVESTIGATION OF A PUTATIVE HINGE MECHANISM IN HUMAN TESTIS ACE

4.1. ACE2 hinge regions:

The motion observed between ACE2c and ACE2o can be described as a hinge movement that opens up the active site. The lid helices, α1 and α2, swivel to one side, and the cleft opens up more on one end than the other. Looking down onto the active site cleft with the lid helices (α1 and α2) on top, the hinge axis stretches from under the N-terminus to the middle of the underside of the active site.

Based on the structural alignment of ACE2o with ACE2c, five hinge regions were identified: 79-106, 288-289, 391-400, 429, 527-529 and 561-566 (Fig 4.1.1). A comparison of the Cα-atom temperature factors of ACE2c, ACE2o, wild-type tACE, tACE-G1,3 and AnCE is presented in figure 4.1.2.

The hinge region closest to the N-terminus (L79-S106) occurs between the α2 lid helix and α4. The overall sequence conservation of this region between tACE, AnCE and ACE2 is low (~7%), although the α-helical structure of the region is conserved. The temperature factors of this region are considerably higher than average in ACE2o, ACE2c, AnCE, tACE-G1,3 and the closed unliganded structure of tACE, indicating a degree of disorder in the structure as would be expected from a flexible region.

Next, K288-P289 lie in a loop region that crosses the more open end of the active site cleft. P289 is conserved in tACE (P298) and AnCE, but the preceding Gly (G286), which presumably lends flexibility to the region, is mutated to a Pro. The temperature factors of these residues are higher than average for all structures considered.

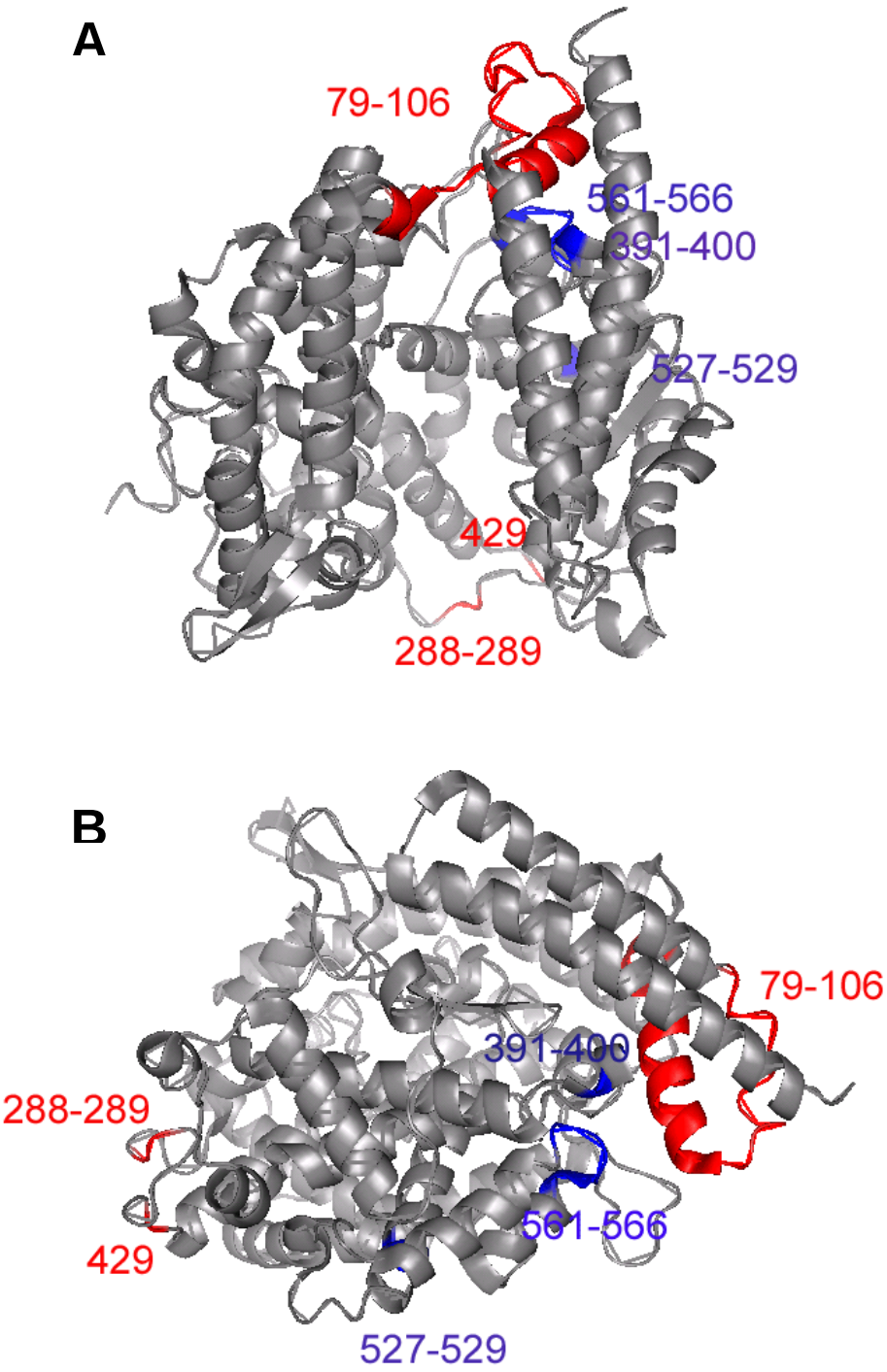


Figure 4.1.1. Two views of ACE2o, showing hinge regions.
A) Looking down into the active site cleft, with the N-terminal lid helices in the foreground on the right half of the molecule. **B)** From the side, with the lid helices diagonal, center top to center right. Hinge regions with low temperature factors and high sequence-conservation (391-400, 527-529, 561-566) are coloured blue; those with high temperature factors and low sequence-conservation (79-106, 288-289, 429) are coloured red.

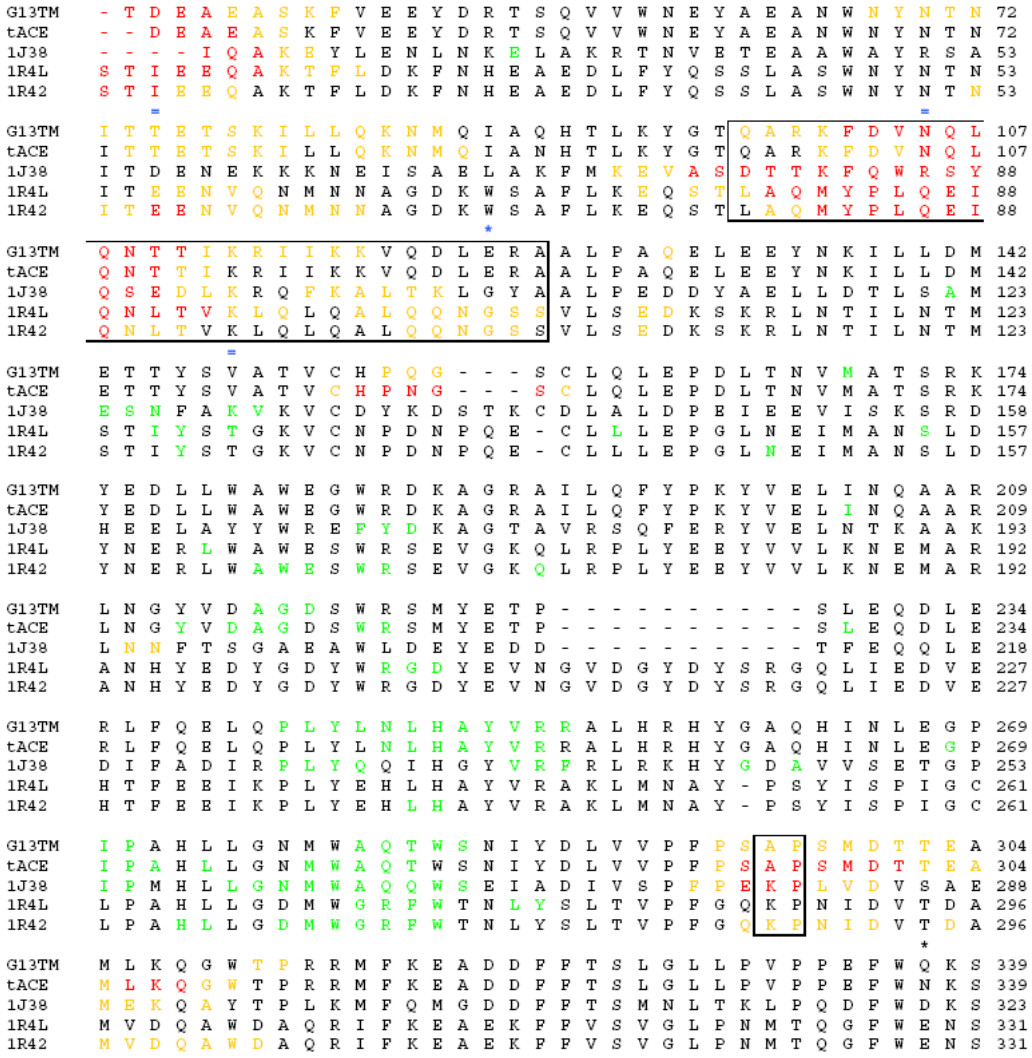
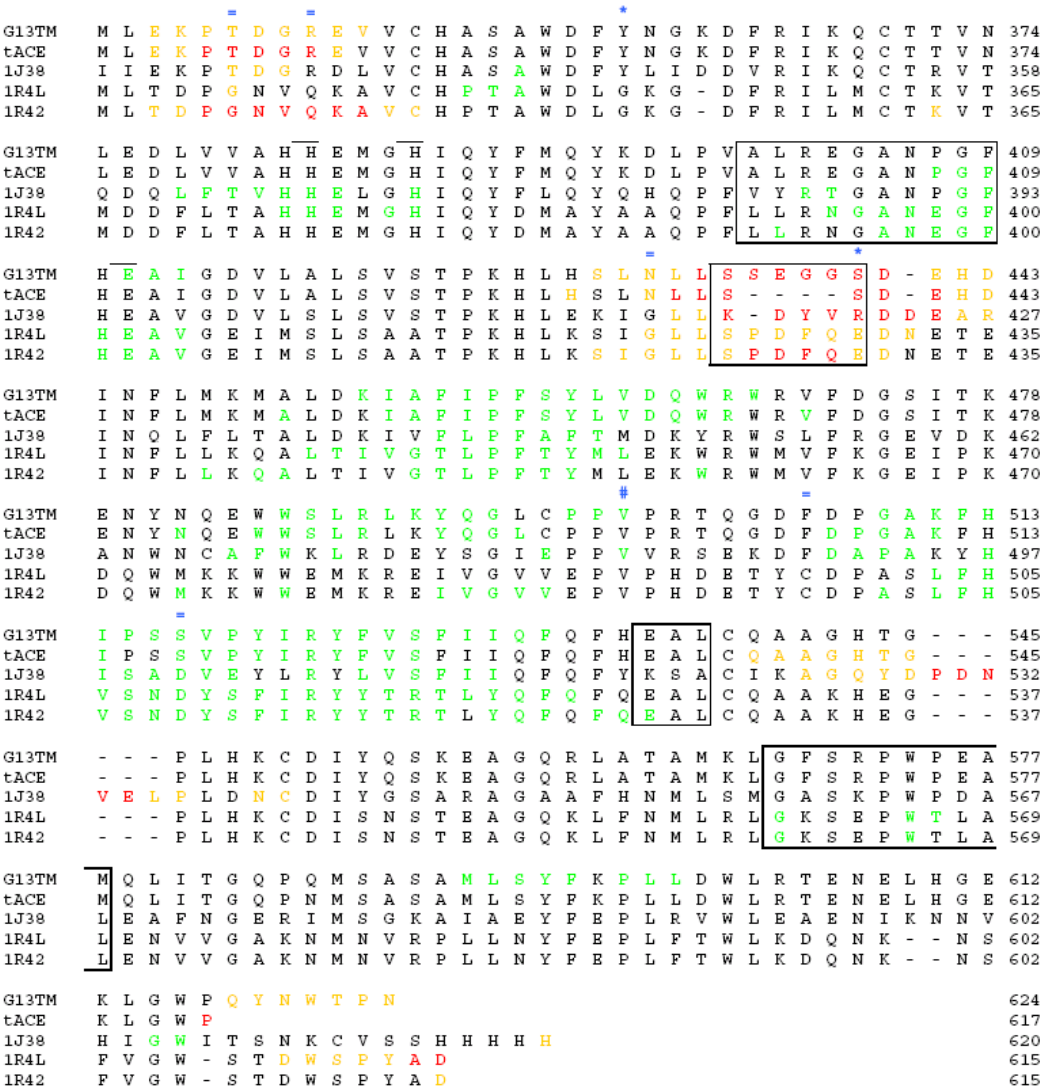


Figure 4.1.2, and overleaf. Temperature factor (B-factor) comparison between tACE glycosylation mutant tACE-G1,3, unliganded wild-type tACE (tACE, 1o8a), *Drosophila* AnCE (1j38) and the open (1r42) and closed (1r4l) structures of ACE2. Sequences are aligned according to Guy *et al* (2003). Note that residues 343-437 of tACE are not shown as they are disordered in the crystal structure.
Residues are coloured according to their temperature factors:
ORANGE: B-factors greater than σ(B); RED: B-factors greater than 2σ(B); GREEN: B-factors less than -σ(B). Boxes indicate hinge regions in ACE2. Lines over residues indicate zinc ligands. Light blue = over a residues indicates residues that are in unfavourable conformations in the closed model of tACE, * indicates residues in unfavourable conformations in both the open and closed models, and # indicates residues that are in unfavourable conformations in the open model only.



Running parallel to this loop, and also crossing the end of the active site cleft, the loop region containing Q429 (S425-E430) is also a hinge region. The sequence of this region is not conserved. The equivalent loop in tACE (S435-S439) is one amino-acid shorter than in ACE2, but in the published structure of tACE, this glycine- and serine-rich loop is disordered and could not be modelled, indicating that its flexibility is conserved despite its shorter length. While it is ordered in tACE-G1,3, AnCE, ACE2c and ACE2o, the temperature factors in this region are high.

E527-L529 lie on the underside of the active site cleft. These residues are conserved in tACE and ACE2, but not in AnCE, although they are surrounded by residues that are conserved in all three homologues. The temperature factors for this region are average-to-low indicating a high degree of order in all structures.

G561-W566 lie under the N-terminus, on the opposite side of the active site cleft from the K288-P289 and S425-E430 loops. These residues are 80% conserved in all three homologues, and the temperature factors in this region are average in tACE-G1,3, tACE, and AnCE, and low in both ACE2 structures.

In addition to these five hinge regions, a sixth region of close alignment (L391-F400) lies between two stretches of structural divergence. This region is in close proximity to the G561-W566 loop, under the N-terminal α1 lid helix. Like the G561-W566 loop, it is conserved (60%) in all three homologues, includes two conserved glycines and a conserved alanine, and has low temperature factors.

Thus six hinge regions were identified based on the domain motion observed in the structures of ACE2. Three of these are highly conserved in tACE, ACE2 and AnCE, rich in glycines, alanines and serines, and display a high degree of order in the crystal structures, as evidenced by their temperature factors. The other three regions show low sequence-conservation and a high degree of disorder or flexibility in the crystal structures.

4.2. Modelling of a putative open form of tACE:

An open model of tACE was generated for normal mode analysis, based on the unbound ACE2o.

A model of tACE (tACEc) based on ACE2c served as a control for errors introduced by the modelling process. The best model of tACEc had a r.m.s. deviation of 1.19Å (572 pairs of Cα-atoms) from the structure of tACE-G1,3, which is the same as the mean deviation between tACE-G1,3 and the closed form of ACE2 (1.20Å for 568 pairs of Cα-atoms).

The Ramachandran plot for this model revealed thirteen residues that have unfavourable main chain conformations. All but two of these correspond either to glycines in ACE2 (E123, D164, T345, Y360, N431), to residues neighbouring glycines in ACE2 (V148, R348), to residues neighbouring an insertion in ACE2 relative to tACE (T226, D440), or to residues which also have unfavourable main-chain torsion angles in the ACE2 structure (N105, F506). Of the remaining two unfavourable residues, T75 (E56 of ACE2) occurs on the surface of the structure, at the start of the α2 helix and S517 (D509 of ACE2) occurs in the active site next to V518 (Y510 of ACE2), an important residue for substrate-recognition (Natesh *et al*, 2003; Towler *et al*, 2004).

Four of the residues in unfavourable conformations, N105, E123, N431 and D440, lie in or near a putative hinge region (figure 4.1.2). N105 and E123 occur in the N-terminal-most hinge region, which has low homology to ACE2. While N105 is in a favourable conformation in the open model, E123 replaces a glycine in ACE2. N431 and D440 flank the flexible S435-S439 loop of tACE, which has low sequence homology to the equivalent loop in ACE2, appearing to retain its flexibility by virtue of a greater proportion of glycines.

The open model of tACE (tACEo) only had six residues with unfavourable main-chain torsion angles. These residues were also, bar one (V499), in unfavourable conformations in the model of the inhibitor-bound form and include D440, near

the S435-S439 loop. The region surrounding V499 (V491 of ACE2) is on the surface of the molecule and contains an extra Pro in tACE, compared with ACE2.

4.3. Normal Mode Analysis:

The results of the normal mode analysis are presented in table 4.3.1 and figure 4.3.1. NMA of both ACE2o and tACEo yielded a single low-frequency normal mode (mode 1, having the lowest frequency, in both cases) that had high overlap with the observed or proposed structural change. In both cases, this mode has moderate collectivity and describes a closing of the active site cleft, with similar amplitudes of perturbation required for maximal overlap with the closed forms. The low r.m.s. deviations between the resulting perturbed models and the comparison structures indicate that the end-states of the normal modes are in very similar conformations to the observed (or modelled) closed structures.

In the case of tACEo, this same low-frequency mode showed high overlap not only with tACEc, the model based on ACE2c, but also with wild-type tACE, the unliganded closed structure. Thus the closed state occurring in this proposed structural transition is similar to the structure of tACE that has been observed, and not merely an artefact of the modelling process.

In contrast with these results, the closed structures (ACE2c, tACEc and wild-type tACE) did not yield normal modes having high overlap with the structural change. This indicates a poor modelling of the proposed structural change by NMA of these structures. While the motions described by the normal modes showing the highest overlap do involve partial opening of the active site cleft, the amplitudes of displacement required for maximum agreement with the comparison structure are low, and the resulting structures do not align closely with the open forms.

Table 4.3.1. Normal mode analysis of tACE and ACE2 - statistics for the mode having the highest overlap for each structure studied.

Structure	Comparison structure	Mode	Overlap	Collectivity	RMSD[a] (Å)	|Amplitude| (DQ[b])	Cumulative Overlap[c]
ACE2o	ACE2c	1	0.733	0.386	1.77	224	0.893
ACE2c	ACE2o	2	0.117	0.132	3.43	85	0.436
tACEo	tACEc	1	0.709	0.427	1.86	248	0.857
tACEc	tACEo	6	0.133	0.500	3.62	96	0.451
tACEo	wild-type tACE	1	0.643	0.427	2.19	245	0.825
wild-type tACE	tACEo	2	0.129	0.468	3.89	107	0.371

[a] R.m.s. deviation for all Cα-atoms, between the comparison structure and a model perturbed along the direction of the mode so as to have the best alignment with the comparison structure.
[b] The amplitude of this perturbation in arbitrary units.
[c] Total overlap for the ten lowest-frequency normal modes calculated for the structure.

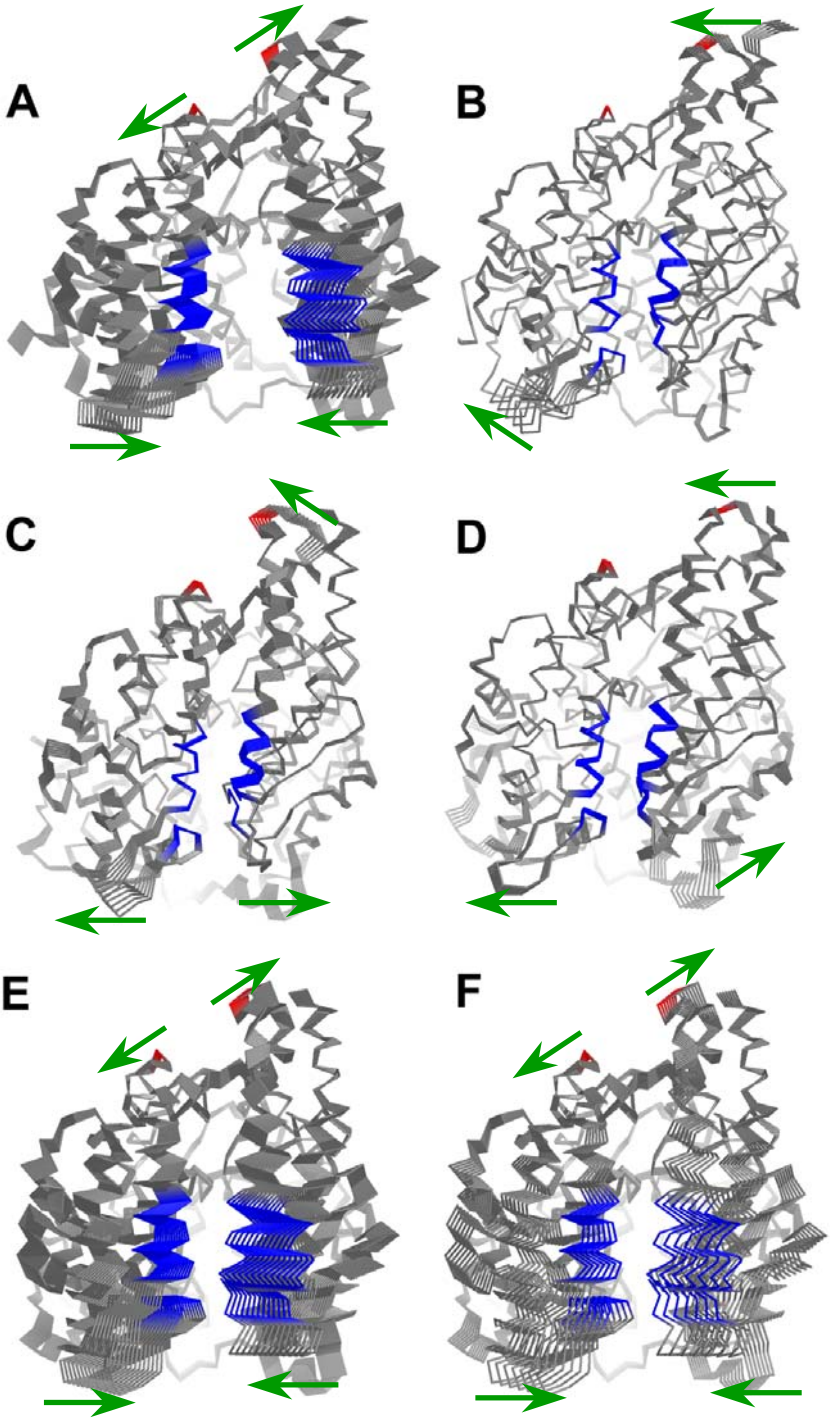


Figure 4.3.1. Normal mode analysis of tACE and ACE2: models perturbed according to the modes showing highest overlap with the putative structural change, the amplitude of perturbation being that which resulted in the closest alignment with the comparison structure. **A)** ACE2o (vs. ACE2c); **B)** ACE2c (vs. ACE2o); **C)** wild-type tACE (vs. tACEo); **D)** tACEc (vs. tACEo); **E)** tACEo (vs. wild-type tACE); **F)** tACEo (vs. tACEc). Colours are according to the residues showing the greatest relative displacements in ACE2o and tACEo: red residues move apart, and blue residues move together as the cleft closes in the modes for these two structures. Green arrows indicate the major directions of movement in the mode in question.

In the low frequency normal mode obtained for tACEo, the residues showing greatest movement apart, based on their Cα-Cα distance fluctuations, are I73-L82 and E349-V351 on one side of the active site cleft, and L140-V148 and E162 on the other. These residues all lie at the opening of the active site cleft, towards the end of the cleft observed to open widest in ACE2 (figures 4.3.1 and 4.3.2). Residues on the under-side of the active site cleft, away from the lid helices, show little or no movement relative to their neighbours. The distance fluctuations are illustrated by the differences in Cα-Cα distances between representative residues in table 4.3.2. The equivalent residues in ACE2 are affected in a similar way by the structural change between ACE2o and ACE2c, and by perturbation along mode 1 for ACE2o (table 4.3.2, figure 4.3.2).

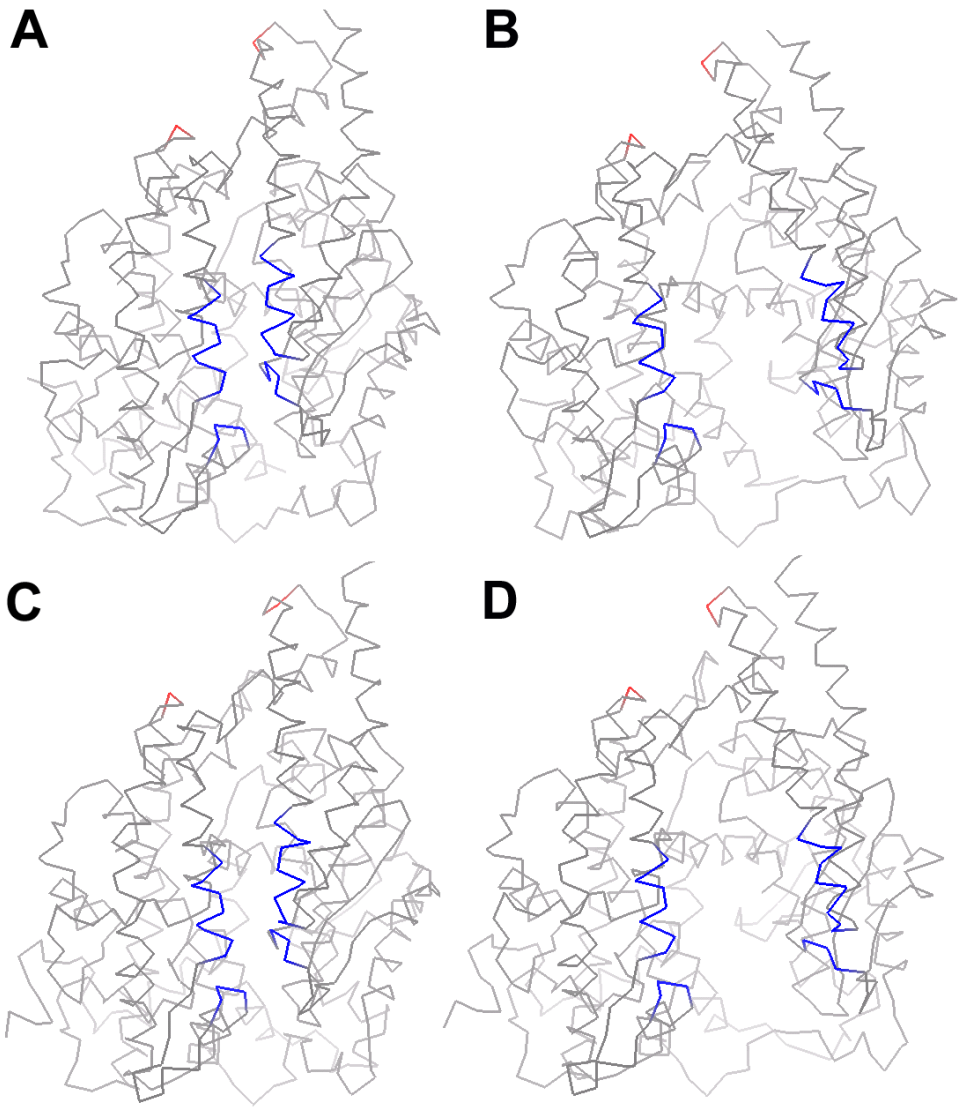


Figure 4.3.2. Residues showing the greatest relative motion during hinging, according to mode 1 of tACEo. Residues moving closer together by are coloured red, and those moving further apart are coloured blue. **A&B)** wild-type tACE (1o8a) and tACEo; **C&D)** the equivalent residues in ACE2c (1r4l) and ACE2o (1r42).

Table 4.3.2. Residues showing large movement relative to one another in mode 1 of tACEo. Comparison of Cα-Cα distances in Ångströms between these residues, with those between the equivalent residues in mode 1 of ACE2o and in the crystal structures of ACE2.

Structure	**Moving Residues**	**Distance when Closed (Å)**	**Distance when Open (Å)**	**Difference (Å)**
tACEo mode 1	N105, G212	16.3	13.8	2.5
ACE2o vs. ACE2c	L85, H195	15.7	13.9	1.8
ACE2o mode 1	L85, H195	16.5	13.9	2.6
tACEo mode 1	N155, D346	24.9	36.8	-11.9
ACE2o vs. ACE2c	D136, N338	23.1	37.8	-14.8
ACE2o mode 1	D136, N338	20.6	37.8	-17.2

CHAPTER 5: DISCUSSION

5.1. Mutation as a potential solution to the glycosylation problem

The effect of N-linked glycosylation on a protein structure is usually to increase its stability by entropic effects (Imperiali & O'Connor, 1999). However, the presence of glycan residues on the surface of a protein is well known to have a negative influence on its ability to form an ordered crystal lattice. This effect is thought to be due partly to the increased conformational entropy introduced by the glycan chain, and partly to the heterogeneity involved in glycan-processing (Hogg *et al*, 2002; Mehndiratta *et al*, 2004).

Human testis ACE contains six N-linked glycosylation sites, having differential importance for the activity and structural integrity of the enzyme (Yu *et al*, 1997; Gordon *et al*, 2003). The reduction of surface glycosylation has been shown to be a useful strategy for the crystallisation of this enzyme, leading to the recent solution of a crystal structure of tACE expressed in CHO cells in the presence of the glycosidase 1-inhibitor, N-butyl-deoxynojirimycin (NB-DNJ) (Gordon *et al*, 2003; Natesh *et al*, 2003).

In the past, the prohibitive cost of glycosidase inhibitors such as NB-DNJ has led other groups to seek out new methods for the reduction of surface glycosylation (Grueninger-Leitch *et al*, 1996). To this end, glycosylation mutants of human tACE, having all but a minimum number of glycosylation sites knocked out by Asn-Gln point mutation, were constructed and expressed in CHO cells, with preliminary crystallisation screens indicating that this approach might yield some results (Gordon *et al*, 2003).

However, it remained to be seen whether a glycosylation mutant of tACE could be crystallised reproducibly for the determination of a crystal structure, or whether such a structure might yield any novel information as to the mode of activity of tACE.

5.2. Reproducible crystallisation of a tACE glycosylation mutant

Crystallisation trials using conditions based on those that had previously yielded promising results, showed extensive precipitation of all mutant tACE samples but tACE-G1,3. This minimally glycosylated mutant crystallised under all of the narrow range of conditions screened, with results being reproducible from one batch of purified protein to the next. The absence of all but the first (N72) and third (N109) glycosylation sites in this mutant and the presence of two fully-processed glycan chains thus does not appear to have an adverse or prohibitive effect on crystal formation.

The use of oils to slow the rate of vapour diffusion was found to improve crystal quality. This is probably due in part to a limitation of the effects of temperature fluctuations on the crystallisation conditions (e.g. by minimising condensation on the cover slide) (Chayen, 1997).

It is thus possible to crystallise a minimally-glycosylated mutant of human tACE, tACE-G1,3, reproducibly for X-ray diffraction purposes.

5.3. Collection of diffraction data and refinement of an unbiased model

Diffraction data collected from a crystal of tACE-G1,3 had a satisfactory R_{merge}, signal-to-noise ratio and redundancy, to 2.9Å. The crystallographic space group was the same as that of the structures of wild-type tACE, and unit cell dimensions were similar, with a slight increase in the volume of the unit cell in this case, probably due to the presence of longer N-linked glycan chains.

In keeping with these indications, the agreement of the molecular replacement solution with the structural data was extremely good. Rather than re-build the model from scratch, this solution was used as a starting point for model-building and phase refinement. Care was taken to remove model-bias by working from weighted difference maps, and by carrying out a round of model-building against

a prime-and-switch-phased map. The over-weighting towards experimental data in the first few steps of refinement would also have had the effect of pulling the structure away from the starting model in cases where the two differed. Thus it can be assumed that the structure solution presented here represents a reasonable fit to the experimental data, without being overly biased towards the starting model.

5.4. The structure of tACE-G1,3

The structure of tACE-G1,3 does not differ in any major respects from the tACE structures solved previously. It is globular, largely alpha-helical, and divided down the middle by an active site cleft. Main-chain deviations between tACE-G1,3 and wild-type tACE are very slight, as are the deviations between the zinc- and chloride-binding sites. As was the case with the unliganded wild-type structure, a fourth zinc ligand is present, in the form of an acetate molecule.

The side chains of several residues were disordered in tACE-G1,3, as was also the case for wild-type tACE. These included three of the four mutant Gln residues, which were present in two rotameric forms, indicative of a disruption of the local hydrogen-bonding environment by the introduction of a longer Gln side chain. However, all of the disordered side chains lie on the surface of the molecule and their degree of order is probably not of functional significance.

High temperature factors, indicative of a high degree of structural disorder or flexibility, were observed in the two flexible loops that cross the end of the active cleft, the ends of the lid helices (near the glycosylated residues), and in the C-terminal loop. Despite the high Cα temperature factors of the residues surrounding N109, the glycan chain at this residue seems to be more ordered than that at N72, as can be seen by comparing the F_{obs}-F_{calc} and $2F_{obs}$-F_{calc} map densities at these sites (figure 3.4.3). The N72-linked glycan projects into a large channel between symmetry-related molecules, while the N109-linked glycan

appears to be in a more constrained environment, which probably explains its higher degree of order in the lattice.

Interestingly, the difference in glycosylation appears to have lead to a slight rearrangement of the crystal lattice relative to that of the wild-type structure, with the expansion of the N72 channel, and the loss of several crystal contacts. This might have been expected to destabilise the lattice, however the loss of protein-protein contacts may be compensated for by glycan-mediated contacts.

Because all of the potential glycosylation sites are on the surface of the molecule, projecting into inter-molecular spaces in the lattice, it is difficult to make any predictions regarding the effect of full glycosylation of the other glycosylation sites on the lattice arrangement. However, it might be expected that different combinations of glycosylation sites could lead to different re-arrangements of lattice contacts. Based on this observation, future crystallisation experiments using the other glycosylation mutants might be expected to yield some novel structural information by virtue of a lattice rearrangement.

From this structural evidence, it appears that the loss of glycosylation sites did not disrupt the structure of tACE-G1,3. Since this mutant can be expressed without the need for expensive glycosidase inhibitors, it promises to be an invaluable tool for further crystallographic studies, in particular for co-crystallisation studies using new potential ACE inhibitors.

5.5. Subdomain II is more rigid than subdomain I

Examination of Cα-atom temperature factors as an indication of the degree of order or structural constraint of a residue, revealed that subdomain II of tACE-G1,3, containing two chloride ions and the active zinc ion, is significantly more stable than subdomain I, which in turn contains most of the more disordered residues.

This pattern of temperature factors is conserved in wild-type tACE, AnCE and ACE2, indicating a conserved structural stability. Since these homologues all crystallised in different space groups and under different conditions, the observed stability appears to be a function of the structure itself. The observation that subdomain II contributes most of the crystal contacts in tACE-G1,3 and wild-type tACE is thus probably a consequence of the stability of this domain, rather than the cause thereof.

This domain stability is likely to be contributed by the bound chloride ions, and may have functional significance. For example, is seems likely that the N-domain of somatic ACE would associate with the C-domain along this relatively immobile surface, leaving subdomain I free to open and close as substrate is bound and products are released.

5.6. Hinge-bending is a possible mechanism for substrate entry

All of the structures of tACE solved to date are in the same conformation, with the active site residues buried deep in a cleft in the molecule (Natesh *et al*, 2003, 2004). The same is true for the insect homologue, AnCE (Kim *et al*, 2003). As it is seen in these structures, the active site cleft is not entirely isolated, being linked by pores to the external milieu. However, these pores are of small diameter (<3Å), and it has been suggested that some kind of flexibility or "breathing" motion must occur in order for a bulky substrate such as AngI to gain access to the active residues (Kim *et al*, 2003).

The first evidence for such a breathing motion is seen in the structures of the human homologue, ACE2, solved recently (Towler *et al*, 2004). While the inhibitor-bound form of this ACE homologue is in a similar conformation to tACE and AnCE, the unbound open form differs considerably. In this form, the active site is opened up by a hinge movement of ~16°, which results in the displacement of some surface residues by as much as 15Å. Based on this hinging, ACE2 can be divided into two domains, which move relatively independently of each other (Towler *et al*, 2004).

Given this evidence, and the fact that some kind of opening action is required in order for substrates to access the active site, it is likely that tACE and AnCE both undergo a similar hinging motion upon substrate-binding, being found in an open state when not bound to inhibitor. Such a hinge mechanism would also explain the large contribution of entropy to the energetics of inhibitor-binding by ACE, since closing of the active site would result in the association of numerous residues that would be bound to ordered solvent molecules when in an open conformation (Andújar-Sánchez *et al*, 2004).

Why then are the unliganded crystal structures of these enzymes not in an open conformation? A clue to a possible explanation lies in the presence of electron density in the active site of unliganded wild-type tACE that was modelled as an acetate and an N-carboxyalanine (Natesh *et al*, 2003). While no unknown ligands were modelled into tACE-G1,3, the resolution of this structure was not high enough to distinguish a low-occupancy ligand from water molecules in the active site. The possibility thus remains that what has been interpreted as a true unliganded form of tACE is more accurately equivalent to another inhibitor-bound form (Towler *et al*, 2004). However, it should be noted that the structure of AnCE (PDB ID 1j38), which is also in a closed conformation in the absence of inhibitor, contains only water molecules as additional zinc-ligands.

If the mechanism for opening the active site is conserved in ACE homologues, then one would expect at least a measure of functional conservation among the residues involved in this motion.

5.7. Hinge-bending regions of ACE2

The transition between the open and closed structures of ACE2 can be described as a hinge-bending movement that opens up the active site (Towler *et al*, 2004). The lid helices, α1 and α2, swivel to one side, with the hinge axis stretching from under the lid helix N-terminus (which lies on the top of one end of the active site cleft), to the middle of the underside of the cleft.

Six regions are involved in the hinge motion, and these can be divided into two groups. The first group comprises four regions (L79-S106, L391-F400, E527-529 and G561-W566) that are concentrated on the N-terminus side of the active site cleft and underneath it. This side of the active cleft is more static during the observed motion, and this group could thus be described as the hinge itself. Three of these regions are highly conserved between tACE, AnCE and ACE, show low temperature factors (a high degree of structural constraint) in both open and closed forms of ACE2, and contain glycines, serines and alanines. The fourth region of this group (L79-S106), showing low sequence conservation and a high degree of disorder, lies on the surface of the molecule, below the N-terminus. Hinging seems to occur about these four regions, with the active site cleft opening more on one side than the other (figure 4.1.1).

G561-W566 of this group lie between the conserved HEMGH of the zinc-binding motif (H383-H384 of tACE) and the additional conserved downstream zinc-binding Glu (E411 of tACE). Although hinging about this region does not alter the relative orientation of the zinc-coordinating residues, it does move the zinc ion away from the opposite wall of the active site cleft by ~3.8Å, thus opening up the catalytic site (figure 1.9.1).

The second group of residues involved in hinging comprises two loop regions (S425-E430 and K288-P289) that stretch one above the other, across the end of the active site cleft that opens out more in the unbound form. The sequence of these residues is not conserved, but their temperature factors are high in all of the structures considered, indicating a degree of flexibility that is functionally conserved. In the case of tACE the S435-S439 loop (corresponding to S425-E430 of ACE2) was disordered in the wild-type crystal structure and could not be modelled.

Because of their structural constraint, it is unclear how the more conserved hinge regions influence the hinge mechanism, although the presence of glycines, serines and alanines in these regions suggests that they may exist in more than one energetically favourable conformation. Their sequence conservation, however,

together with the conservation of functional flexibility in the loop regions, suggests a conservation of the hinge mechanism among these homologues.

5.8. Investigation of tACE hinging by normal mode analysis

A modelling technique that has proven useful in describing protein domain motions is that of normal mode analysis (NMA). The basis of this technique is the assumption that the atoms in a molecule behave as simple harmonic oscillators, and that all of the possible motions of the molecule can be described by combinations of these oscillations. Diagonalisation of the Hessian matrix (the mass-weighted second derivatives of the molecule's potential energy matrix) yields so-called "directions" of oscillation, or normal modes. Each of these normal modes describes a harmonic oscillation of the structure about a local energy minimum, with the low-frequency modes describing large-amplitude collective motions, such as those undergone during a hinge-bending event in a protein (Suhre & Sanejouand, 2004). It has been shown that in many cases a single low-frequency normal mode accurately describes the domain movements observed in the crystallographic structures of proteins (Tama & Sanejouand, 2001).

In order to investigate the possibility of hinge-movement in tACE further, NMA of tACE and ACE2 was carried out.

For comparison, a model of a putative open structure of tACE (tACEo) was made, based on the structure of the open form of ACE2. Because ACE2 exhibits only 42% sequence identity to tACE, this model was only a rough one, as indicated by the presence of several residues having unfavourable main chain torsion values, and by the deviation (1.19Å, Cα atoms) of the closed control model (tACEc) from the crystal structure of tACE-G1,3. However, since the force-field used to compute low-frequency normal modes was a simplified one, based on the spatial proximity of Cα atoms, the modes determined can be expected to depend on macroscopic more than on local structure (Tama *et al*, 2000; Tama & Sanejouand, 2001).

Four of the residues in unfavourable conformations in tACEc were located in or near putative hinge regions. However, only one of these was also present in tACEo, and this residue (S449) is in a region that is highly disordered (i.e. flexible) in the crystal structures of tACE. The errors in the model were therefore not expected to affect the outcome of the analysis adversely.

NMA of both tACEo and ACE2o yielded a single normal mode having a high degree of overlap with the putative structural change. These modes describe a closing of the active site cleft that corresponds closely to the hinging observed in the ACE2 crystal structures, with moving residues behaving similarly in both cases, and with perturbed structures aligning closely with the observed closed crystal forms. The residues on the underside of the active site cleft, including most of subdomain II, show little or no movement relative to neighbouring residues.

The NMA for the closed structures (tACEc, ACE2c and wild-type tACE) however, did not yield modes having high overlap with the structural change. This is probably because the moving residues are close together in the starting structure, becoming linked in the force field so that large Cα-Cα fluctuations between them are not favoured. This tendency of closed structures to yield normal modes that do not describe the structural transition as well as those of the corresponding open form has been documented for a number of cases (Tama & Sanejouand, 2001). However, since these results were similar for both wild-type tACE and ACE2c, they do not exclude the possibility that tACE might undergo a hinge movement. Rather, the fact that NMA of the wild-type structure of tACE does yield some normal modes having high collectivity is a further indication that some kind of concerted movement about the active site is likely to occur.

It could be argued that the similarity in behaviour of the normal modes of tACEo and ACE2o is an artefact due to the modelling of the open structure based on ACE2o. However the close alignment of the perturbed model of tACEo with the wild-type tACE structure suggests that this model may represent a reasonable approximation of a real open form of tACE.

This evidence, taken together with the requirement for opening of the active cleft to allow substrate access, the entropically-driven nature of substrate binding, the hinging of ACE2, the presence of an acetate molecule in the active site of tACE-G1,3 and the wild-type tACE structure, and the conservation of the proposed hinge regions in ACE2, tACE and AnCE at the level of flexibility, suggests that tACE does have an open form and that this open form is similar to ACE2o.

CHAPTER 6: CONCLUSIONS AND DIRECTIONS FOR FURTHER STUDY

6.1. Conclusions

Apart from a slight rearrangement of lattice contacts, the structure of glycosylation mutant tACE-G1,3 does not differ significantly from the structure of wild-type tACE. This means that it is possible to solve tACE structures without the use of expensive glycosidase inhibitors, which in turn has important implications for future structural studies of ACE inhibitor binding.

Based on analysis of temperature factors, subdomain II of tACE-G1,3, containing bound chloride ions and the zinc-binding site, displays a high degree structural rigidity while subdomain I appears to be more flexible. This domain stability is conserved in wild-type tACE, ACE2 and AnCE, and may have functional significance.

The residues involved in the hinge motion of ACE2 are conserved, at least at the level of functional flexibility, in tACE and AnCE. Normal mode analysis of the closed tACE structure and a modelled open form demonstrated that the intrinsic flexibility of this structure about the active site cleft is similar to that of ACE2. This suggests that hinging is a common mechanism for substrate entry in ACE homologues. This hypothesis is further supported in the literature by the observation that some kind of motion must occur to allow substrate access, the presence of unidentified ligand in the active site of the wild-type tACE structure, and the calorimetric evidence for a large entropic contribution to substrate binding (Kim *et al*, 2003; Natesh *et al*, 2003; Andújar-Sánchez *et al*, 2004).

6.2. Directions for further study

By establishing a more cost-effective means of determining tACE structures, this work provides a platform for further crystallographic studies of this enzyme.

Because of the importance of ACE as a drug target, the primary thrust of these studies must be in the area of drug-design and characterisation. Crystallisation of new potential inhibitors bound to tACE-G1,3 would provide insight into the mode of binding of these possible drug leads, yielding structure-based insight to aid the

design process. The likelihood that a large-scale hinge motion is involved in the binding of substrates means that co-crystallisation is likely to be a more successful strategy than soaking pre-grown crystals in a solution of the ligand.

Further investigation into the possibility of a hinge mechanism could entail the introduction of mutations into the putative hinge regions of tACE. For example, mutants having shorter or more rigid loop regions might be expected to have different catalytic characteristics from wild-type tACE.

NMA could also be carried out in more detail. For instance, the application of an empirical force field would allow a more in-depth comparison to be made between ACE2o and the tACEo homology model, while determination of the magnitude of the energy barriers to be overcome in the proposed hinge event would shed further light on its energetic favourability (Van Wynbergher *et al*, 2004). Such studies could be verified experimentally, by calorimetric comparison of inhibitor-binding by tACE and ACE2.

Crystallisation screens could also be carried out using a broader range of conditions. For example, a structure lacking the active site acetate observed in unliganded wild-type tACE and tACE-G1,3 might be obtained by using a different precipitant buffer. Alternatively, the use of different combinations of glycosylation site mutations might yield a mutant tACE that crystallised preferentially in an alternative crystal form, perhaps having a different conformation.

The aim of all such studies would be the elucidation of the activity and substrate-binding capabilities of ACE, in the interests of designing new drugs for the treatment of cardiac disease.

ABBREVIATIONS

Å	Ångströms
ACE	Angiotensin-converting Enzyme
ACE2	human homologue of ACE
ACE2c	closed, inhibitor-bound structure of ACE2
ACE2o	open, unbound structure of ACE2
ACEr	*Drosophila* homologue of ACE
AnCE	*Drosophila* homologue of ACE
Ang_{1-7}	Angiotensin, residues 1-7 (DRVYIHP)
AngI	Agiotensin-I (DRVYIHPFHL), precursor of AngII
AngII	Agiotensin-II (DRVYIHPF)
BK	Bradykinin (RPPGFSPFR)
CHO's	Chinese Hamster Ovary cells
DMEM	Dulbecco's Modified Eagle Medium
ENM	Elastic Network Model
FCS	Foetal Calf Serum
FOM	phasing Figure of Merit
HEPES	*N*-(2-Hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)
Hip-His-Leu	Hippuryl-histidyl-leucine
KKS	Kallikrein-kinin System
N-AcSDKP	*N*-Acetyl-seryl-aspartyl-lysyl-proline
NB-DNJ	*N*-butyl-deoxynojirimycin
NMA	Normal Mode Analysis
r.m.s.	root mean square
RAS	Renin-Angiotensin System
R_{cryst}	crystallographic R-factor = $\Sigma\|\|F_{obs}\|-\|F_{calc}\|\|/\Sigma\|F_{obs}\|$
R_{free}	crystallographic R-factor for the test set (5% of reflections)
R_{merge}	merging R-factor = Σ ($\|(I - \langle I \rangle)\|$) / Σ (I)
RXP407	N-domain-specific substrate
RXPA380	C-domain-specific substrate
sACE	somatic ACE
$sACE_{C\text{-}domain}$	C-terminal domain of sACE
$sACE_{N\text{-}domain}$	N-terminal domain of sACE
SDS-PAGE	Sodium Dodecylsulphate Polyacrylamide Gel Electrophoresis

tACE	testis ACE
tACEc	closed model of human tACE, based on ACE2c
tACE-G1,3	mutant of human tACE, lacking all but the first and third N-glycosylation sites
tACE-G1234	mutant of human tACE, lacking all but the first four N-glycosylation sites
tACE-G3	mutant of human tACE, tacking all but the third N-glycosylation site
tACEo	open model of human tACE, based on ACE2o
wt-tACE	wild-type human tACE

REFERENCES

1. Acharya K, Sturrock E, Riordan J & Ehlers M (2003). *ACE revisited: A new target for structure-based drug design.* **Nature Reviews Drug Discov.** 2: 891-903.

2. Andújar-Sánchez M, Cámara-Artigas A & Jara-Pérez V (2004). *A calorimetric study of the binding of lisinopril, enalaprilat and captopril to angiotensin-converting enzyme.* **Biophys. Chem.** 111: 183-189.

3. Arndt J, Hao B, Ramakrishnan V, Cheng T, Chan S & Chan M (2002). *Crystal structure of a novel carboxypeptidase from the hyperthermophilic archaeon Pyrococcus furiosus.* **Structure (Camb).** 10: 215-224.

4. Azizi M, Rousseau A, Ezan E, Guyene T, Michelet S, Grognet J, Corvol P & Lenfant M (1996). *Acute angiotensin-converting enzyme inhibition increases the plasma level of the natural stem cell regulator N-acetyl-seryl-aspartyl-lysyl-proline.* **J. Clin. Invest.** 97: 839-844.

5. Bala M, Pasha M, Bhardwaj D & Pasha S (2002). *Novel peptidomimics as angiotensin-converting enzyme inhibitors: A combinatorial approach.* **Bioorg. Med. Chem.** 10: 3685-3691.

6. Bernstein K, Martin B, Edwards A & Bernstein E (1989). *Mouse angiotensin-converting enzyme is a protein composed of two homologous domains.* **J. Biol. Chem.** 264:11945-11951.

7. Binevski P, Sizova E, Pozdnev V & Kost O (2003). *Evidence for the negative cooperativity of the two active sites within bovine somatic angiotensin-converting enzyme.* **FEBS Lett.** 550: 84-88.

8. Brooks B & Karplus M (1983). *Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor.* **Prot. Sci.** 1:227-235.

9. Brown C, Madauss K, Lian W, Beck M, Tolbert W & Rodgers D (2001). *Structure of neurolysin reveals a deep channel that limits substrate access.* **Proc. Natl. Acad. Sci. USA.** 98: 3127-3132.

10. Butters T, Sparks L, Harlos K, Ikemizu S, Stuart D, Jones E & Davis S (1999). *Effects of* N*-butyldeoxynojirimycin and the Lec3.2.8.1 mutant phenotype on N-glycan processing in Chinese hamster ovary cells: Application to glycoprotein crystallization.* **Prot. Sci.** 8: 1696-1701.

11. Canutescu A, Shelenkov A, & Dunbrack R Jr. (2003). *A graph-theory algorithm for rapid protein side-chain prediction.* **Prot. Sci.** 12: 2001–2014.

12. Chayen N (1997). *The role of oil in macromolecular crystallization.* **Structure.** 5: 1269-1274.

13. Coates D, Isaac R, Cotton J, Siviter R, Williams T, Shirras A, Corvol P & Dive V (2000). *Functional conservation of the active sites of human and* Drosophila *angiotensin I-converting enzyme.* **Biochemistry.** 39: 8963-8969.

14. Cohen, G. E (1997). ALIGN: A program to superimpose protein coordinates, accounting for insertions and deletions. **J. Appl. Cryst.** 30: 1160-1161.

15. Collaborative Computational Project, Number 4 (1994). *The CCP4 suite: programs for protein crystallography*. **Acta Cryst.** D50: 760-763.

16. Cornell M, Williams T, Lamango N, Coates D, Corvol P, Soubrier F, Hoheisel J, Lehrach H & Isaac R (1995). *Cloning and expression of an evolutionary conserved single-domain angiotensin converting enzyme from Drosophila melanogaster.* **J Biol Chem.** 270: 13613-13619.

17. Cotton J, Hayashi M, Cuniasse P, Vazeux G, Ianzer D, Camargo A & Dive V (2002). *Selective inhibition of the C-domain of angiotensin I converting enzyme by bradykinin potentiating peptides.* **Biochemistry.** 41: 6065-6071.

18. Crackower M, Sarao R, Oudit G, Yagil C, Kozieradzki I, Scanga S, Oliveira-dos-Santos A, da Costa J, Zhang L, Pei Y, Scholey J, Ferrario C, Manoukian A, Chappell M, Backx P, Yagil Y & Penninger J (2002). *Angiotensin-converting enzyme 2 is an essential regulator of heart function.* **Nature.** 417: 822-828.

19. Danilczyk U, Eriksson U, Crackower M & Penninger J (2003). *A story of two ACEs.* **J. Mol. Med.** 81: 227-234.

20. Deddish P, Wang J, Michel B, Morris P, Davidson N, Skidgel R & Erdös E (1994). *Naturally occurring active N-domain of human angiotensin I-converting enzyme.* **Proc. Natl. Acad. Sci. USA.** 91: 7807-7811.

21. Deddish P, Marcic B, Jackman H, Wang H, Skidgel R & Erdös E (1998). *N-domain-specific substrate and C-domain inhibitors of angiotensin-converting enzyme: Angiotensin-(1-7) and keto-ACE.* **Hypertension.** 31: 912-917.

22. DeLano W. *The PyMOL molecular graphics system.* DeLano Scientific LLC, San Carlos, CA, USA. http://www.pymol.org

23. Donoghue M, Hsieh F, Baronas E, Godbout K, Gosselin M, Stagliano N, Donovan M, Woolf B, Robison K, Jeyaseelan R, Breitbart R & Acton S (2000). *A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9.* **Circ. Res.** 87: E1-9.

24. Ehlers M & Riordan J (1989). *Angiotensin-converting enzyme: New concepts concerning its biological role.* **Biochemistry.** 28: 5311-5319.

25. Ehlers M, Fox E, Strydom D & Riordan J (1989). *Molecular cloning of human testicular angiotensin-converting enzyme: The testis isozyme is identical to the C-terminal half of endothelial angiotensin-converting enzyme.* **Proc. Natl. Acad. Sci. USA.** 86: 7741-7745.

26. Ehlers M, Chen Y & Riordan J (1991a). *Purification and characterisation of recombinant human testis angiotensin-converting enzyme expressed in Chinese hamster ovary cells.* **Prot. Express. Pur.** 2: 1-9.

27. Ehlers M, Chen Y & Riordan J (1991b). *Spontaneous solubilization of membrane-bound human testis angiotensin-converting enzyme expressed in Chinese hamster ovary cells.* **Proc. Natl. Acad. Sci. USA.** 88: 1009-1013.

28. Ehlers M, Chen Y & Riordan J (1992). *The unique N-terminal sequence of testis angiotensin-converting enzyme is heavily O-glycosylated and unessential for activity or stability.* **Biochem. Biophys. Res. Commun.** 183: 199-205.

29. Ehlers M, Schwager S, Scholle R, Manji G, Brandt W & Riordan J (1996). *Proteolytic release of membrane-bound angiotensin-converting enzyme: Role of the juxtamembrane stalk sequence.* **Biochemistry.** 35: 9549-9559.

30. Fernandez J, Hayashi M, Camargo A & Neshich G (2003). *Structural basis of the lisinopril-binding specificity in N- and C-domains of human somatic ACE.* **Biochem. Biophys. Res. Commun.** 308: 219-226.

31. Fodje M & Al-Karadaghi S (2002). *Occurrence, conformational features and amino acid propensities for the pi-helix.* **Prot. Eng.** 15: 335-358.

32. Friedland J & Silverstein E (1976). *A sensitive fluorimetric assay for serum angiotensin-converting enzyme.* **Am. J. Clin. Pathol.** 66: 416-424.

33. Fuchs S, Xiao H, Cole J, Adams J, Frenzel K, Michaud A, Zhao H, Keshelava G, Capecchi M, Corvol P & Bernstein K (2004). *Role of the N-terminal catalytic domain of ACE investigated by targeted inactivation in mice.* **J. Biol. Chem.** 43: 854-859.

34. Georgiadis D, Beau F, Czarny B, Cotton J, Yiotakis A & Dive V (2003). *Roles of the two active sites of somatic angiotensin-converting enzyme in the cleavage of angiotensin I and bradykinin*. **Circ. Res.** 93: 148-154.

35. Gerstein M, Lesk A & Chothia C (1994). *Structural mechanisms for domain movements in proteins.* **Biochemistry.** 33: 6739-6749.

36. Gordon K, Redelinghuys P, Schwager S, Ehlers M, Papageorgiou A, Natesh R, Acharya K & Sturrock E (2003). *Deglycosylation, processing and crystallization of human testis angiotensin-converting enzyme.* **Biochem. J.** 371: 437-442.

37. Goulter A, Goddard M, Allen J & Clark K (2004). *ACE2 gene expression is up-regulated in the human failing heart.* **BMC Med.** 2:19.

38. Grueninger-Leitch F, D'Arcy A, D'Arcy B & Chene C (1996). *Deglycosylation of proteins for crystallization using recombinant fusion protein glycosidases.* **Prot. Sci.** 5: 2617-2622.

39. Guy J, Jackson R, Acharya K, Sturrock E, Hooper N & Turner A (2003). *Angiotensin-converting enzyme-2 (ACE2): Comparative modeling of the active site, specificity requirements, and chloride dependence.* **Biochemistry.** 42: 13185-13192.

40. Hagaman J, Moyer J, Bachman E, Sibony M, Magyar P, Welch J, Smithies O, Krege J & O'Brien D (1998). *Angiotensin-converting enzyme and male fertility.* **Proc. Natl. Acad. Sci. USA.** 95: 2552-2557.

41. Harris R, Ohlsson J & Wilson I (1981). *Purification of human serum angiotensin I-converting enzyme by affinity chromatography.* **Anal. Biochem.** 111: 227-234.

42. Hausrath A & Matthews B (2002). *Thermolysin in the absence of substrate has an open conformation.* **Acta. Cryst.** D58: 1002-1007.

43. Haznedaroglu I & Öztürk M (2003). *Towards the understanding of the local hematopoietic bone marrow renin-angiotensin system.* **Int. J. Biochem. Cell. Biol.** 35: 867-880.

44. Hens K, Vandingenen A, Macours N, Baggerman G, Karaoglanovic A, Schoofs L, de Loof A & Huybrechts R (2002). *Characterization of four substrates emphasizes kinetic similarity between insect and human C-domain angiotensin-converting enzyme.* **Eur. J. Biochem.** 269: 3522-3530.

45. Hogg T, Smatanova I, Bezouska K, Ulbrich N & Hilgenfeld R (2002). *Sugar-mediated lattice contacts in crystals of a plant glycoprotein*. **Acta Cryst.** D58: 1734-1739.

46. Hubert C, Houot A, Corvol P & Soubrier F (1991). *Structure of the angiotensin I-converting enzyme gene.* **J. Biol. Chem.** 266: 15377-15383.

47. Imperiali B & O'Connor S (1999). *Effect of N-linked glycosylation on glycopeptide and glycoprotein structure.* **Curr. Opin. Chem. Biol.** 3: 643-649.

48. Jääskeläinen S, Verma C, Hubbard R, Linko P & Caves L (1998). *Conformational change in the activation of lipase: An analysis in terms of low-frequency normal modes.* **Prot. Sci.** 7: 1359-1367.

49. Jaspard E, Wei L & Alhenc-Gelas F (1993). *Differences in the properties and enzymatic specificities of the two active sites of angiotensin I-converting enzyme (kininase II). Studies with bradykinin and other natural peptides.* **J. Biol. Chem.** 268: 9496-9503.

50. Jones A & Woods D (2003). *Skeletal muscle RAS and exercise performance.* **Int. J. Biochem. Cell. Biol.** 35: 855-866.

51. Junot C, Gonzales M, Ezan E, Cotton J, Vazeux G, Michaud A, Azizi M, Vassiliou S, Yiotakis A, Corvol P & Dive V (2001). *RXP407, a selective inhibitor of the N-domain of angiotensin I-converting enzyme, blocks in vivo the degradation of hemoregulatory peptide acetyl-Ser-Asp-Lys-Pro with no effect on angiotensin I hydrolysis.* **J. Pharmacol. Exp. Ther.** 297: 606-611.

52. Kasturi S, Jabbar M, Sen G & Sen I (1994). *Role of glycosylation in the biosynthesis and activity of rabbit testicular angiotensin-converting enzyme.* **Biochemistry.** 33: 6228-6234.

53. Kessler S, Gomos J, Scheidemantel T, Rowe T, Smith H & Sen G (2002). *The germinal isozyme of angiotensin-converting enzyme can substitute for the somatic isozyme in maintaining normal renal structure and functions.* **J. Biol. Chem.** 277: 4271-4276.

54. Kim H, Shin D, Yoo O, Lee H & Lee J (2003). *Crystal structure of* Drosophila *angiotensin I-converting enzyme bound to captopril and lisinopril.* **FEBS Lett.** 538: 65-70.

55. Kissinger C, Gehlhaar D & Fogel D (1999). *Rapid automated molecular replacement by evolutionary search.* **Acta Cryst.** D55: 484-491.

56. Kumar R, Kusari J, Roy S, Soffer R & Sen G (1989). *Structure of testicular angiotensin-converting enzyme. A segmental mosaic enzyme.* **J. Biol. Chem.** 264: 16754-16758.

57. Lipscomb W & Sträter N (1996). *Recent advances in zinc enzymology.* **Chem. Rev.** 96: 2375-2434.

58. Liu X, Fernandez M, Wouters M, Heyberger S & Husain A (2001). *Arg-1098 is critical fro the chloride dependence of human angiotensin I-converting enzyme C-domain catalytic activity.* **J. Biol. Chem.** 276: 33518-33525.

59. Ma J & Karplus M (1998). *The allosteric mechanism of the chaperonin GroEL:A dynamic analysis*. **Proc. Natl. Acad. Sci. USA.** 95: 8502-8507.

60. Magnusson U, Chaudhuri B, Ko J, Park C, Jones T & Mowbray S (2002). *Hinge-bending motion of D-allose-binding protein from Escherichia coli.* **J. Biol. Chem.** 277: 14077-14084.

61. Marcic B, Deddish P, Jackman H, Erdös E & Tan F (2000). *Effects of the N-terminal sequence of ACE on the properties of its C-domain.* **Hypertension**.36: 116-121.

62. Mchaourab H, Oh K, Fang C & Hubbel W (1997). *Conformation of T4 lysozyme in solution. Hinge-bending motion and the substrate-induced conformational transition studied by site-directed spin labeling.* **Biochemistry.** 36: 307-316.

63. Mehndiratta P, Walton W, Hare J, Pulido S, Parthasarathy G, Emmett M, Marshall A & Logan T (2004). *Expression, purification, and characterization of avian Thy-1 from Lec1 mammalian and Tn5 insect cells.* **Prot. Express. Pur.** 33: 274-287.

64. Mouawad L & Perahia D (1996). *Motions in hemoglobin studied by normal mode analysis and energy minimisation: Evidence for the existence of tertiary T-like, quaternary R-like intermediate structures.* **J. Mol. Biol.** 258: 393-410.

65. Murshudov G (1997). *Refinement of macromolecular structures by the maximum-likelihood method*. **Acta Cryst.** D53: 240-255.

66. Nachon F, Nicolet Y, Viguié N, Masson P, Fontecilla-Camps J & Lockridge O (2002). *Engineering of a monomeric and low-glycosylated form of human butyrylcholinesterase: Expression, purification, characterization and crystallization.* **Eur. J. Biochem.** 269: 630-637.

67. Natesh R, Schwager S, Sturrock E & Acharya K (2003). *Crystal structure of the human angiotensin-converting enzyme-lisinopril complex.* **Nature.** 421: 551-554.

68. Natesh R, Schwager S, Evans H, Sturrock E & Acharya K (2004). *Structural details on the binding of antihypertensive drugs captopril and enalaprilat to human testicular angiotensin I-converting enzyme.* **Biochemistry.** 43: 8718-8724.

69. Newcomer M, Lewis B & Quiocho F (1981). *The radius of gyration of L-arabinose-binding protein decreases upon binding of ligand.* **J. Biol. Chem.** 256: 13218-13222.

70. Otwinowski Z & Minor W (1997). *Processing of x-ray diffraction data collected in oscillation mode.* **Methods Enzymol.** 276: A307-326.

71. Pantoliano M, Holmquist B & Riordan J (1984). *Affinity chromatographic purification of angiotensin converting enzyme.* **Biochemistry.** 23: 1037-1042.

72. Patchett A, Harris E, Tristram E, Wyvratt M, Wu M, Taub D, Peterson E, Ikeler T, ten Broeke J, Payne L, Ondeyka D, Thorsett E, Greenlee W, Lohr N, Hoffsommer R, Joshua H, Ruyle W, Rothrock J, Aster S, Maycock A, Robinson F, Hirschmann R, Sweet C, Ulm E, Gross D, Vassil T & Stone C. (1980). *A new class of angiotensin-converting enzyme inhibitors.* **Nature**. 288: 280-283.

73. Pauls K, Metzger R, Steger K, Klonisch T, Danilov S & Franke F (2003). *Isoforms of angiotensin I-converting enzyme in the development and differentiation of human testis and epididymis.* **Andrologia.** 35: 32-43.

74. Pelmenschikov V, Blomberg M & Siegbahn P (2002). *A theoretical study of the mechanism for peptide hydrolysis by thermolysin.* **J. Biol. Inorg. Chem.** 7: 284-298.

75. Rousseau A, Michaud A, Chauvet M, Lenfant M & Corvol P (1995). *The hemoregulatory peptide* N*-acetyl-Ser-Asp-Lys-Pro is a natural and specific substrate of the N-terminal active site of angiotensin-converting enzyme.* **J. Biol. Chem.** 270: 3656-3661.

76. Sali A & Blundell T (1993). *Comparative protein modelling by satisfaction of spatial restraints.* **J. Mol. Biol.** 234: 779-815.

77. Santhamma K, Sadhukhan R, Kinter M, Chattopadhyay S, McCue B & Sen I (2004). *Role of tyrosine phosphorylation in the regulation of cleavage secretion of angiotensin-converting enzyme.* **J. Biol. Chem.** 279: 40227-40236.

78. Shai S, Fishel R, Martin B, Berk B & Bernstein K (1992). *Bovine angiotensin converting enzyme cDNA cloning and regulation. Increased expression during endothelial cell growth arrest.* **Circ. Res.** 70: 1274-1281.

79. Shapiro R & Riordan J (1984). *Inhibition of angiotensin converting enzyme: dependence on chloride.* **Biochemistry.** 23: 5234-5240.

80. Shen Y, Kong Y & Ma J (2002). *Intrinsic flexibility and gating mechanism of the potassium channel KcsA.* **Proc. Natl. Acad. Sci. USA.** 99: 1949-1953.

81. Sica D (2004). *Current concepts of pharmacotherapy in hypertension: Angiotensin-converting enzyme inhibitors side effects -- physiologic and non-physiologic considerations.* **J. Clin. Hypertens.** 6: 410-416.

82. Soubrier F, Alhenc-Gelas F, Hubert C, Allegrini J, John M, Tregear G & Corvol P (1988). *Two putative active centers in human angiotensin I-converting enzyme revealed by molecular cloning.* **Proc. Natl. Acad. Sci. USA.** 85: 9386-2390.

83. Suhre K & Sanejouand Y (2004). *ElNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement.* **Nucl. Ac. Res.** 32: W610-W614.

84. Tama F, Gadea F, Marques O & Sanejouand Y (2000). *Building-block approach for determining low-frequency normal modes of macromolecules.* **Proteins.** 41: 1-7.

85. Tama F & Sanejouand Y (2001). Conformational change of proteins arising from normal mode calculations. **Prot. Eng.** 14: 1-6.

86. Taylor C, Coates D & Shirras A (1996). *The Acer gene of Drosophila codes for an angiotensin-converting enzyme homologue.* **Gene.** 181: 191-197.

87. Terwilliger C (2001). *Map-likelihood phasing.* **Acta. Cryst.** D57: 1763-1775.

88. Thekkumkara T, Livingston W 3rd, Kumar R, and Sen G (1992). *Use of alternative polyadenylation sites for tissue-specific transcription of two angiotensin-converting enzyme mRNAs.* **Nucl. Ac. Res.** 20: 683-687.

89. Tipnis S, Hooper N, Hyde R, Karran E, Christie G & Turner A (2000). *A human homolog of angioensin-converting enzyme: Cloning and functional expression as a captopril-insensitive carboxypeptidase.* **J. Biol. Chem.** 275: 33238-33243.

90. Towler P, Staker B, Prasad S, Menon S, Tang J, Parsons T, Ryan D, Fisher M, Williams D, Dales N, Patane M & Pantoliano M (2004). *ACE2 X-ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis.* **J. Biol. Chem.** 279: 17996-18007.

91. Turner A & Hooper N (2002). *The angiotensin-converting enzyme gene family: genomics and pharmacology.* **Trends Pharm. Sci.** 23: 177-183.

92. Turner A, Tipnis S, Guy J, Rice G & Hooper N (2002). *ACEH/ACE2 is a novel mammalian metallocarboxypeptidase and a homologue of angiotensin-converting enzyme insensitive to ACE inhibitors.* **Can. J. Physiol. Pharmacol.** 80: 346-353.

93. Tzakos A, Galanis A, Spyroulias G, Cordopatis P, Manessi-Zoupa E & Gerothanassis I (2003). *Structure-function discrimination of the N- and C-catalytic domains of human angiotensin-converting enzyme: implications for Cl^- activation and peptide hydrolysis mechanisms.* **Prot. Eng.** 16: 993-1003.

94. Valadié H, Lacapčre J, Sanejouand Y & Etchebest C (2003). *Dynamical properties of the MscL of* Escherischia coli*: a normal mode analysis.* **J. Mol. Biol.** 332: 657-674.

95. Van Wynsberghe A, Li G & Cui Q (2004). *Normal-mode analysis suggests protein flexibility modulation throughout RNA polymerase's functional cycle*. **Biochemistry.** 43: 13083-13096.

96. Voronov S, Zueva N, Orlov V, Arutyunyan A & Kost O (2002). *Temperature-induced selective death of the C-domain within angiotensin-converting enzyme molecule.* **FEBS Lett**. 522: 77-82.

97. Wei L, Alhenc-Gelas F, Corvol P & Clauser E (1991). *The two homologous domains of human angiotensin I-converting enzyme are both catalytically active.* **J. Biol. Chem.** 266: 9002-9008.

98. Wei L, Clauser E, Alhenc-Gelas F & Corvol P (1992). *The two domains of human angiotensin I-converting enzyme interact differently with competitive inhibitors.* **J. Biol. Chem.** 267: 13398-13405.

99. Williams T, Corvol P & Soubrier F (1994). *Identification of two active site residues in human angiotensin I-converting enzyme.* **J. Biol. Chem.** 269: 29430-29434.

100. Williams T, Michaud A, Houard X, Chauvet M, Soubrier F & Corvol P (1996). Drosophila melanogaster *angiotensin I-converting enzyme expressed in* Pichia pastoris *resembles the C domain of the mammalian*

homologue and does not require glycosylation for secretion and enzymic activity. **Biochem J.** 318: 125-131.

101. Williams S, Addy J, Phillips J 3rd, Dai M, Kpodonu J, Afful J, Jackson H, Joseph K, Eason F, Murray M, Epperson P, Aduonum A, Wong L, Jose P & Felder R (2000). *Combinations of variations in multiple genes are associated with hypertension.* **Hypertension.** 36:2-6.

102. Woodman Z, Oppong S, Cook S, Hooper N, Schwager S, Brandt W, Ehlers M & Sturrock E (2000). *Shedding of somatic angiotensin-converting enzyme (ACE) is inefficient compared with testis ACE despite cleavage at identical stalk sites.* **Biochem J.** 347: 711-718.

103. Yu X, Sturrock E, Wu Z, Biemann K, Ehlers M & Riordan J (1997). *Identification of* N*-linked glycosylation sites in human testis angiotensin-converting enzyme and expression of an active deglycosylated form.* **J. Biol. Chem.** 272: 3511-3519.

104. Zaman M, Oparil S & Calhoun D (2002). *Drugs targeting the renin-angiotensin-aldosterone system*. **Nat. Rev. Drug. Discov.** 1: 621-636.

105. Zhang X, Wozniak J & Matthews B (1995). *Protein flexibility and adaptability as seen in 25 crystal forms of T4 lysozyme.* **J. Mol. Biol.** 250: 527-552.

106. Zhang H, Wada J, Hida K, Tsuchiyama Y, Hiragushi K, Shikata K, Wang H, Lin S, Kanwar Y & Makino H (2001). *Collectrin, a collecting duct-specific transmembrane glycoprotein, is a novel homolog of ACE2 and is developmentally regulated in embryonic kidneys.* **J Biol Chem.** 276: 17132-17139.