Structural determinants of the domain-selectivity of novel inhibitors of human testis angiotensin-converting enzyme

Jean Margaret Watermeyer

Thesis Presented for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Medical Biochemistry

UNIVERSITY OF CAPE TOWN

February 2008

Supervisors: E.D. Sturrock and B.T. Sewell

Abstract

Angiotensin-converting enzyme (ACE) is a zinc metallopeptidase and a key drug target for the treatment of hypertension, due to its action on vasoregulatory peptides angiotensin I and bradykinin. ACE consists of two active domains, the N and C domains, having similar structures and activities. Each domain is globular and largely α-helical, with the active zinc ion buried in a cleft that divides the domain in half. Some substrates are cleaved preferentially by one domain, with the cleavage of angiotensin I depending largely on the C domain. It has been suggested that C-domain-selective inhibitors might not cause the side effects arising from current ACE inhibitor therapies. Here, X-ray co-crystal structures have been solved of three novel C-domain-selective inhibitors in the active site of a glycosylation-deficient mutant of human testis ACE, tACE-G13, equivalent to the minimally-glycosylated C domain. This is the first successful application of this mutant in the drug design process. Inhibitors (5S)-5-[(N-benzoyl)amino]-4-oxo-6-phenyl-hexanoyl-L-tryptophan (kAW) and (5S)-5-[(N-benzoyl)amino]-4-oxo-6-phenyl-hexanoyl-L-phenylalanine (kAF) were derived from keto-ACE by P_2' substitution. Their co-crystal structures, to 2.18 and 2.17 Å, respectively, reveal the hydrated geminal diolate form of the ketone zinc-binding group, reminiscent of the tetrahedral transition state. Two enantiomers of inhibitor (SR)-1-(N(2)-(1-carboxy-3-phenylpropyl)-L-lysyl)-L-tryptophan (lisW) were derived from lisinopril by P_2' substitution, however structures solved to 2.40 and 2.30 Å revealed the selection of the L-amino-acid form, lisW-S, by the active site. Inter-domain substitutions have been identified among the side chains interacting with these three inhibitors which account for their domain-selectivity. Based on this work, several of these substitutions, E376D, V380T, Y391F and V518S, have subsequently been verified as important for the domain-selectivity of kAW and kAF. The P_2' Trp moieties of kAW and lisW-S take different conformations in the active site, indicating the large volume of the S_2' pocket and cooperativity of binding between the P_2' and P_1' moieties. Conserved active site water molecules were identified, including water 1165 of the kAW structure which may be important for the chloride activation of the C domain due to its proximity to the chloride ligand R522. These data provide a platform for further structure-based design of domain-selective inhibitors.

Acknowledgements

Many thanks to my supervisors, Prof. Edward Sturrock and Prof. Trevor Sewell for their enthusiastic support and input and for creating an environment in which it has been a pleasure to learn.

Thanks to the following people who gave welcome technical assistance: Riyad Domingo for producing the salts of purified inhibitor samples, and Dr. Hassan Belrhali (BM14(UK), ESRF) for his assistance with the data collection.

I am grateful for the help of Prof. Kelly Chibale in understanding reaction mechanisms, for the correspondence of Dr. David Christianson on the subject of his carboxypeptidase A structures, and for Dr. Muhammed Sayed's crystallography input.

Thanks to my colleagues and fellow students in the zinc metalloprotease and structural biology groups, Itai Chitapi, Wendy Kröger, Kerry Gordon, Ross Douglas, Dr. Trudi O'Neil, Serah Kimani, Jason van Rooyen, Nailah Conrad and Ayesha Parker, for their friendship and friendly collaboration.

Thanks to my parents, John and Fiona Watermeyer, to my sister-and-digsmate Kate, to Andrew, and to my friends who have supported and encouraged me in my studies.

This work could not have been undertaken without the generous support of the University of Cape Town (Harry Crossley Foundation), the South African Synchrotron Fund, the South African National Research Foundation and the South African Department of Labour.

Abbreviations

ACE	angiotensin-converting enzyme
ACE2	angiotensin-converting enzyme 2
ACEr	ACE homologue from *Drosophila melanogaster*
AcSDKP	N-acetyl-SDKP peptide
AI	angiotensin I decapeptide, DRVYIHPFHL
AII	angiotensin II octapeptide, DRVYIHPF
AMP	adenosine monophosphate
AnCE	ACE homologue from *Drosophila melanogaster*
Ang_{1-7}	angiotensin 1-7 peptide, DRVYIHP
AT_1	type 1 angiotensin receptor
AT_2	type 2 angiotensin receptor
B-factors	crystallographic atomic displacement parameters
BK	bradykinin nonapeptide, RPPGFSPFR
BOP	5-amino-(N-t-butoxycarbonyl)-2-benzyl-4-oxo-6-phenylhexanoic acid
BPP	bradykinin-potentiating peptide
C domain	C-terminal domain of human angiotensin-converting enzyme
CHO	Chinese hamster ovary
CPA	carboxypeptidase A
Dcp	dipeptidyl carboxypeptidase from *Escherichia coli*
DMEM	Dulbecco's modified eagle medium
DMSO	dimethyl sulphoxide
gem-diol	geminal diolate
GnRH	gonadotropin releasing hormone
Hepes	N-2-Hydroxyethylpiperazine-N'-2-Ethanesulfonic Acid
I/D	insertion/deletion polymorphism
kAF	keto-ACE derivative with P_2' Phe
kAP	keto-ACE, (5S)-5-[(N-benzoyl)amino]-4-oxo-6-phenyl-hexanoyl-L-Pro
kAW	keto-ACE derivative with P_2' Trp
KKS	kallikrein-kinin system
lisW	lisinopril derivative with P_2' Trp
lisW-R	R-enantiomer of lisW, at the P_1 alpha carbon

lisW-S	S-enantiomer of lisW, at the P_1 alpha carbon
N domain	N-terminal domain of human angiotensin-converting enzyme
NCBI	National Centre for Biotechnology Information
NEP	neutral endopeptidase
MR	molecular replacement
OligoF	oligoendopeptidase F from *Bacillus stearothermophilus*
PDB	Protein Data Bank (www.rcsb.org/pdb)
PDB ID	Protein Data Bank identity code
PfuCP	thermostable carboxypeptidase 1 from *Pyrococcus furiosus*
r.m.s.	root mean square
RAS	renin-angiotensin system
RXPA380	C-domain-selective phosphinic peptide inhibitor
RXP407	N-domain-selective phosphinic peptide inhibitor
sACE	somatic angiotensin-converting enzyme
SARS	severe acute respiratory syndrome
SDS-PAGE	sodium dodecylsulphate polyacrylamide gel electrophoresis
SNP	single nucleotide polymorphism
tACE	testis angiotensin-converting enzyme
tACE-G13	minimally glycosylated mutant of human testis angiotensin-converting enzyme
tACE-G1234	minimally glycosylated mutant of human testis angiotensin-converting enzyme
tACEΔ36	human testis angiotensin-converting enzyme construct lacking the N-terminal O-glycosylated region.
ThCP	carboxypeptidase 1 from *Thermus thermophilus*

Table of Contents

CHAPTER 1: Introduction

1.1. Hypertension and the discovery of angiotensin-converting enzyme

1.1.1. The discovery of angiotensin-converting enzyme

Arterial hypertension is a medical condition characterised by elevated blood pressure, which frequently contributes to the development of cardiac disease and chronic renal failure. Its causes and treatment have been the subjects of speculation and serious scientific investigation since as early as 2600 BC (reviewed briefly by Esunge (1991)). In the 1950's it became apparent that a substance found in the blood plasma of hypertensive humans and animals induced an increase in blood pressure when administered to test animals. The substance, then named hypertensin, was produced by the enzyme renin and was found to exist in inactive (hypertensin I) and active (hypertensin II) forms (Skeggs *et al.*, 1954). Since injection of either form into test animals resulted in an increase in blood pressure, a "hypertensin-converting enzyme" which catalysed the rapid conversion of hypertensin I to hypertensin II, was hypothesised and then isolated (Skeggs *et al.*, 1956). This enzyme, now known variously as angiotensin-converting enzyme (ACE; EC 3.4.15.1), peptidyl-dipetidase A or dipeptidyl carboxypeptidase I, was found to be a zinc-metallopeptidase expressed in most mammalian tissues, and a key component of what came to be known as the renin-angiotensin system (RAS).

1.1.2. The renin-angiotensin and kallikrein-kinin systems

In the circulating RAS (reviewed by Erdös (1990) and Zaman *et al.* (2002)), angiotensin I (AI, previously hypertensin I), a decapeptide, is derived from the action of the aspartic protease renin on angiotensinogen, a 118 amino-acid protein precursor synthesised in the liver. AI is in turn cleaved by ACE to produce the vasopressor octapeptide angiotensin II (AII, previously hypertensin II). The effects of AII are mediated by two G-protein-coupled receptors. Receptor AT_1 mediates the classic vasoconstrictor effects, while receptor AT_2 is expressed abundantly during foetal development and at sites of cellular stress in adults, where it mediates vasodilation, cell differentiation and tissue repair among other effects.

In addition to its RAS function, ACE was also found to be the same enzyme as "kininase II" of the kallikrein-kinin system (KKS, reviewed by Skidgel & Erdös (2004) and Moreau *et al.* (2005)). In the KKS, circulating bradykinin (BK), a nonapeptide derived from cleavage of

kininogens by a number of different proteases, binds to the B_2 kinin receptor, causing vasodilation and increased vascular permeability. The action of ACE on BK produces an inactive heptapeptide which is unable to bind to the B_2 receptor. ACE thus acts on two vasoregulatory peptides, causing an increase in blood pressure in both cases.

Based on this understanding of its dual hypertension-inducing activities in the RAS and KKS, ACE was identified as an inhibitor target. The history of ACE drug design has been reviewed by de Lima (1999) and Acharya *et al.* (2003). Today ACE inhibitors are commonly used in the treatment of hypertension, congestive heart failure, myocardial infarction and diabetic nephropathy (Table 1.1; Sica, 2004).

Table 1.1 United States food and drug administration indications for angiotensin-converting enzyme inhibitors (Sica, 2004).

Drug	Hypertension	Congestive Heart Failure	Diabetic Nephropathy	High Risk Patients without left ventricular dysfunction
Captopril	•	• (post-MI) [a]	•	
Benazepril	•			
Enalapril	•	• [b]		
Fosinopril	•	•		
Lisinopril	•	• (post-MI) [a]		
Moexipril	•			
Perindopril	•			
Quinapril	•	•		
Ramipril	•	• (post-MI)		•
Trandolapril	•	• (post-MI)		

MI = myocardial infarction; [a] captopril and lisinopril are indicated for congestive heart failure treatments both post-MI and as adjunctive therapy in general heart failure therapy; [b] enalapril is indicated for asymptomatic, left ventricular dysfunction.

1.2. ACE primary structure

1.2.1. The ACE gene encodes two homologous domains

When the ACE gene was sequenced in 1988, it was found to consist of two homologous domains (the N-terminal or N domain and C-terminal or C domain), each containing a zinc-binding motif of the sequence HEMGH + E (Soubrier *et al.*, 1988). This motif places ACE in the M2 family of the MA (zincins) clan of zinc metallopeptidases, which are characterised by the zinc binding motif HEXXH + E, containing two zinc-binding histidines, with a downstream glutamate as the third zinc ligand (reviewed by Turner & Hooper (2002) and Spyroulias *et al.* (2004)).

The N and C domains are approximately 55% identical to one another and both are active on peptide substrates (Wei *et al.*, 1991). They are preceded by a signal peptide (29 residues) targeting expression to the cell membrane, separated by a flexible linker peptide (about 11 residues) and have an additional 78 C-terminal residues comprising stalk, juxtamembrane, transmembrane and cytoplasmic tail regions (Appendix A; Hubert *et al.*, 1991). A second promoter in intron 12 of the ACE gene contains a cyclic AMP responsive element which results in the tissue-specific expression of a single domain isoform, testis ACE (tACE), in developing sperm cells (Hubert *et al.*, 1991; Langford *et al.*, 1993; Esther, Jr. *et al.*, 1997). tACE is identical to the C domain of ACE, except for an additional 56-residue N-terminal signal peptide and O-glycosylated region (Ehlers *et al.*, 1989).

The N-terminal signal peptide is cleaved off during post-translational processing, and both isoforms are shed from the cell membrane by the action of a putative ACE sheddase at a site in the stalk region (Ehlers *et al.*, 1996). This shedding, together with the evidence for high levels of ACE expression in lung endothelia (Ehlers & Riordan, 1989), explains how soluble ACE comes to form part of the circulating RAS and KKS.

1.2.2. Polymorphisms in the ACE gene

Since 1988, the human ACE gene has been studied in some detail, and a number of polymorphisms have been identified. Most significant among these is a common noncoding insertion/deletion (I/D) polymorphism of 287 bp in intron 16 of the ACE gene (Rigat *et al.*,

1990). Individuals with the D/D genotype have increased blood and tissue ACE activity relative to those having the I/I phenotype, and as could be expected this polymorphism has been implicated in susceptibility to a number of ACE-related disease states (reviewed by Sayed-Tabatabaei *et al.* (2006)). In addition to this polymorphism, 73 single nucleotide polymorphisms (SNPs) are listed by the SNP database (dbSNP) at NCBI (the National Centre for Biotechnology Information, http://www.ncbi.nlm.nih.gov/) as falling in the ACE gene locus. Thirty-one of these fall within coding regions of the gene, however only one coding SNP is associated with a known phenotype: a point mutation that results in the coding mutation P1199L near the sheddase site in the stalk region results in an increase in shedding and a non-pathogenic increase in serum ACE levels (Kramers *et al.*, 2001).

1.3. ACE in other species

The sequence of ACE is highly conserved (> 80% identical) among mammals, and rodent, porcine, rabbit and bovine ACE have been studied as models for human ACE activity. Such studies were initially carried out *in vitro* when the tissues of these animals were used as sources of purified ACE (for example Bunning *et al.* (1983) and Voronov *et al.* (2002)). More recently *in vivo* studies have been undertaken (for example Binevski *et al.* (2003), Georgiadis *et al.* (2003) and Ocaranza *et al.* (2006)), however it should be noted that the alterations in sequence between species have been shown to correlate with small changes in ACE substrate specificity and chloride dependence, so that conclusions drawn from studies with other mammalian ACEs should be extrapolated to human ACE only with caution (Ibarrarubio *et al.*, 1993; Jullien *et al.*, 2006).

ACE-related sequences also occur in a wide variety of organisms across the phylogenetic spectrum, and active ACE homologues have been identified in several non-mammalian species, including a number of insect species and annelids (Williams *et al.*, 1996; Macours *et al.*, 2004). Remarkably, an ACE homologue from the bacterium *Xanthomonas axonopodis* was recently found to be active, and to have a similar substrate and inhibitor profile to mammalian ACE (Riviere *et al.*, 2007). This ubiquity across phyla suggests a conserved fundamental role for ACE, which is supported by the observation that insect ACE plays a vital role in pupal development (Isaac *et al.*, 2007).

1.4. ACE substrates and inhibitors - domain-selectivity

1.4.1. Two domains result in dual activity

The presence of two active domains in ACE does not simply amount to a duplication of activity, but rather a dual activity, since the 45% of residues which are non-identical confer a certain degree of dissimilarity in kinetic properties between the N and C domains. For instance, the N domain is more thermally stable than the C domain (Voronov *et al.*, 2002) and less strongly activated by chloride ions (Wei *et al.*, 1991). Most notable in the context of inhibitor design, however, are differences in substrate and inhibitor specificity.

ACE can act on a range of peptide substrates. Its primary activity is as a C-terminal dipeptidase, but it is also capable of endopeptidase and tripeptidase activity on substrates with amidated C-termini (Ehlers & Riordan, 1991; Skidgel & Erdös, 2004; Naqvi *et al.*, 2005). Substrate binding is described using a subsite model, with S_1, S_2, S_3, S_4, etc. describing substrate side chain binding pockets increasingly distant from the scissile peptide bond to the N-terminal side, and S_1', S_2' (and occasionally S_3') describing binding pockets for side chains increasingly distant on the C-terminal side. The equivalent side chains in the peptide substrate are referred to as P_1, P_2, P_3, P_4 or P_1', P_2', P_3'.

1.4.2. Domain-selectivity of known substrates and inhibitors

The vasoregulatory peptides AI and BK have approximately equal affinity for both ACE domains, however AI cleavage is carried out more efficiently by the C domain than the N domain *in vitro* (Wei *et al.*, 1991). In addition to AI and BK, ACE cleaves a number of physiological peptides *in vitro*, including RAS and KKS peptides, neuropeptides, hormones and growth factors (see Table 1.2), and is also thought to be involved in the removal of basic C-terminal residues during prohormone processing (Isaac *et al.*, 1999). These additional activities illustrate the broad sequence specificity of ACE as well as implicating the enzyme in a more complex functionality than simply the regulation of blood pressure. Among these additional substrates, β-neoendorphin, AcSDKP, Ang_{1-7}, GnRH and amyloid β peptide are preferentially cleaved by the N domain (Jaspard *et al.*, 1993; Deddish *et al.*, 1998; Junot *et al.*, 2001; Oba *et al.*, 2005), with Ang_{1-7} actually having an inhibitory effect on the C domain (Deddish *et al.*, 1998). The domain-selectivity of biological substrates may have a physiological role.

A degree of domain-selectivity is also observed with several of the ACE inhibitors currently in clinical use (Table 1.3), despite the fact that these inhibitors were designed without knowledge of the two-domain structure of ACE. For example, the tripeptide analogue lisinopril has as much as 18-fold higher affinity for the C domain than for the N domain (Table 1.3; Wei *et al.*, 1992). Moreover, the short peptide substrate commonly used to assay for ACE activity, Hippuryl-His-Leu (Friedland & Silverstein, 1976), is cleaved approximately 10-fold more efficiently by the C domain than by the N domain (Jaspard *et al.*, 1993).

Bradykinin-potentiating peptides (BPPs) are naturally-occurring ACE inhibitors found in snake venom, which were instrumental in the development of captopril, the first ACE inhibitor to be used as an antihypertensive treatment (Ondetti *et al.*, 1977; de Lima, 1999; Hayashi & Camargo, 2005). They constitute a natural set of peptide inhibitors of varying sequence, many of which display strong domain-selectivity (Table 1.4; Cotton *et al.*, 2002; Hayashi & Camargo, 2005). Another inhibitor based on the BPP sequences and also designed early in the search for antihypertensive treatments, is keto-ACE, an analogue of benzoyl-FGP with a ketomethylene group replacing the scissile bond (Almquist *et al.*, 1980). This inhibitor was also found to be moderately selective for the C domain (Deddish *et al.*, 1998).

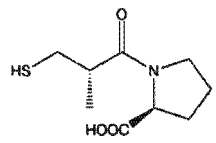
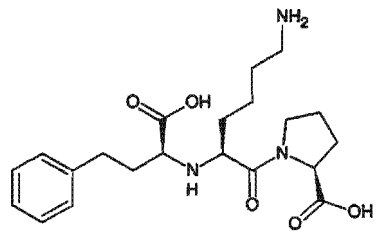
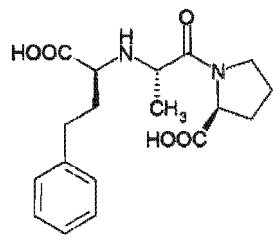
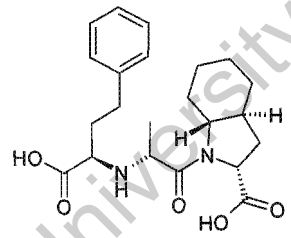
Examination of these compounds, together with studies using synthetic peptide libraries (for example, Michaud *et al.*, 1999; Cotton *et al.*, 2002; Bersanetti *et al.*, 2004; Jullien *et al.*, 2006), has yielded some insight into the determinants of domain-selectivity. C-domain-selectivity has been suggested to be caused by among other things, bulky P_1, P_1' or P_2' groups (Perich *et al.*, 1994; Acharya *et al.*, 2003), or a P_2 Lys (Cotton *et al.*, 2002). However, these determinants are complicated by cooperativity between side chains which can affect domain affinity (Michaud *et al.*, 1999).

Table 1.2. Sequence and functional diversity of biological peptide substrates of ACE.

Substrate	Sequence and cleavage site(s) *	Notes on cleavage and function.
Vasoregulatory peptides		
Angiotensin I (AI)	H_3N-DRVYIHPF / HL-COOH	First physiological substrate discovered. Cleaved by both domains (Jaspard *et al.*, 1993).
Bradykinin (BK)	H_3N-RPPGFSP / FR-COOH	Second physiological substrate discovered. Cleaved by both domains (Jaspard *et al.*, 1993).
$Angiotensin_{1-9}$	H_3N-DRVYI / HP / FH-COOH	Inactive product of AI cleavage by human ACE homologue ACE2 (Eriksson *et al.*, 2002).
$Angiotensin_{1-7}$	H_3N-DRVYI / HP-COOH	Product of the action of neutral endopeptidase on AI or of ACE2 on AII. Binds Ang_{1-7} receptor, with similar functions to those of AI via the AT_2 receptor (reviewed by Acharya *et al.* (2003)). Cleaved by the N domain, inhibits the C domain (Deddish *et al.*, 1998).
Des-Arg^9 bradykinin	H_3N-RPPGF / SPF-COOH	C-terminal tripeptidase activity (Inokuchi & Nagamatsu, 1981).
$Bradykinin_{1-7}$	H_3N-RPPG / F / SP-COOH	Produced from BK by action of ACE *in vitro;* C-terminal di- or tripeptidase activity. (Inokuchi & Nagamatsu, 1981; Jaspard *et al.*, 1993).
Other biological peptides		
Amyloid β peptide	H_3N-DAEFRHD / SGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA-COOH	Forms plaques in Alzheimer's disease. Cleavage by ACE is thought to reduce plaque formation (Hemming & Selkoe, 2005). Preferentially cleaved by the N domain acting as an endopeptidase (Oba *et al.*, 2005).
GnRH (or LH-RH)	pEHW / SYG / LRP-$CONH_2$	Reproductive hormone with a modified N- and C-termini. Tripeptidase activity on either end; cleaved preferentially by the N domain (Skidgel & Erdös, 1985; Jaspard *et al.*, 1993).
Substance P	H3N-RP / KPGGFF / G / LM-$CONH_2$	Neuropeptide cleaved in two orientations, with tripeptidase or dipeptidase activity (reviewed by Zaman *et al.* (2002), Naqvi *et al.* (2005) and Jaspard *et al.* (1993)).
β-neoendorphin	H3N-YGGFL / RK / YP-COOH	Neuropeptide and endogenous opioid, cleaved sequentially by ACE to yield Leu-enkephalin (β-$neoendorphin_{1-7}$) and β-$neoendorphin_{1-5}$ (Hayakari *et al.*, 1997).
Met-enkephalin	H3N-YGG / FM-COOH	Opioid neuropeptide. Reviewed by Erdös (1990).
Neurotensin	H3N-ELYENKPRRPY / IL-$CONH_2$	Neuropeptide (Skidgel *et al.*, 1984).
FMet-Leu-Phe	FMet-LF-COOH	Chemotactic peptide. Reviewed by Erdös (1990).
N-AcSDKP	Ac-SD / KP-COOH	Haematopoietic signalling peptide. Cleaved by the N domain (Georgiadis *et al.*, 2003)

* pE = pyroglutamic acid; $CONH_2$ = amidated C terminus; FMet = *N*-formyl methionine; Ac = *N*-acetyl

Table 1.3 Domain-selectivity of ACE inhibitors in clinical use (Wei *et al.*, 1992; Hayashi & Camargo, 2005).

Inhibitor	Structure	$Ki_{N\ domain}$ (nM)	$Ki_{C\ domain}$ (nM)
captopril		8.9	14.0
lisinopril		44	2.4
enalaprilat		26.0	6.3
trandolaprilat		3.1	0.3

1.4.3. Domain-selective inhibitors and domain cooperativity

In recent years, phosphinic peptide inhibitors RXP407 and RXPA380 have been identified as having strong selectivity for the N and C domains respectively, allowing for complete inhibition of one domain at a time (Dive *et al.*, 1999; Junot *et al.*, 2001; Georgiadis *et al.*, 2003; Georgiadis *et al.*, 2004). Studies with RXP407 and RXPA380 revealed that inhibition of either the N or C domain alone resulted in inhibition of AI hydrolysis in vivo, while BK hydrolysis occurred as long as one domain remained uninhibited (Georgiadis *et al.*, 2003).

Table 1.4 Sequences and domain affinity of bradykinin-potentiating peptides (Hayashi & Camargo, 2005).

Bradykinin-potentiating peptide	Sequence [a]	$Ki_{N\ domain}$ (nM)	$Ki_{C\ domain}$ (nM)	C-domain-selectivity ($Ki_{N\ domain}/Ki_{C\ domain}$)
BPP-5a	pEKWAP	400	1280	0.31
BPP-5b	pEWPRP	-	-	-
BPP-6a	pESWPGP	-	-	-
BPP-7a	pEDGPIPP	-	-	-
BPP-9a (BPPa)	pEWPRPQIPP	100	1	100.00
BPP-10a	pESWPGPNIPP	130	33	3.94
BPP-10b	pENWPRPQIPP	-	-	-
BPP-10c (BPP2)	pENWPHPQIPP	200	0.5	400.00
BPP-11a	pEWPRPTPQIPP	-	-	-
BPP-11b	pEGRAPGPPIPP	3000	57	52.63
BPP-11c	pEGRAPHPPIPP	2000	40	50.00
BPP-11d	pEGRPPGPPIPP	-	-	-
BPP-11e	pEARPPHPPIPP	100	300	0.33
BPPb	pEGLPPRPKIPP	10000	30	333.33
BPPc	pEGLPPGPPIPP	80	80	1
BPP-12a	pEGWAWPRPQIPP	-	-	-
BPP-12b	pEWGRPPGPPIPP	5	150	0.033
BPP-12c	pEWAQWPRPQIPP	-	-	-
BPP-13a	pEGGWPRPGPEIPP	50	50	1
BPP-13b	pEGGLPRPGPEIPP	-	-	-
BPP-14a	pEWAQWPRPTPQIPP	-	-	-

[a] pE = pyroglutamic acid

This finding suggests a degree of negative cooperativity between domains, which is supported by the active site titration of one mole of lisinopril per mole of ACE from various mammalian sources *in vitro* (Ehlers & Riordan, 1991; Binevski *et al.*, 2003; Andújar-Sánchez *et al.*, 2004).

However, the issue of domain cooperativity is not clearly understood, since the existence and degree of cooperativity seems to depend on the substrate and assay conditions, and on whether the study is carried out *in vivo* or *in vitro*. For example, studies in which RXP407 and RXPA380 have been used to inhibit one domain at a time demonstrated continued (non-cooperative) activity of the respective uninhibited domain towards two fluorogenic substrates as well as AI and BK *in vitro*, and negative domain cooperativity in the cleavage of AI but not BK *in vivo* (Georgiadis *et al.*, 2003; Jullien *et al.*, 2006). Moreover, administration of

RXP407 had no effect on circulating AII levels in mice (Junot *et al.*, 2001), and RXPA380 but not RXP407 inhibited a response to injected AI in porcine arteries (van Esch *et al.*, 2005), suggesting that the C domain may be largely responsible for conversion of AI to AII *in vivo*. These findings were confirmed by studies using ACE domain-inactivated mice, which revealed that an active C domain is both necessary and sufficient *in vivo* to carry out normal AI processing in mice, whereas BK processing proceeds normally whichever domain is inactivated (Fuchs *et al.*, 2004; Fuchs *et al.*, 2007).

1.5. The three-dimensional structure of ACE

1.5.1. The problem of glycosylation and three-dimensional structure solution

ACE is a glycoprotein with 16 potential N-linked glycosylation sites (Appendix A), nine in the N domain and six in the C domain, and it was this fact that delayed the solution of a crystal structure of ACE until recent years, since glycosylation is known to hinder the formation of a crystal lattice due to the flexibility and heterogeneity of glycoforms on the protein surface (Hogg *et al.*, 2002; Mehndiratta *et al.*, 2004). It was thus necessary to reduce glycosylation prior to crystallisation.

Structures have been solved of minimally glycosylated tACE (Protein Data Bank code (PDB ID), 1O8A; equivalent to the C domain) and of the N domain (PDB ID, 2C6N), expressed in the presence of glucosidase I inhibitor N-butyl-deoxynojirimycin to enforce uniform glycosylation (Natesh *et al.*, 2003; Gordon *et al.*, 2003; Corradi *et al.*, 2006). In addition, the technique of eliminating N-linked glycosylation sites by point mutation has yielded glycosylation-deficient tACE mutants having two (tACE-G13; PDB ID, 2IUL) and four (tACE-G1234; PDB ID, 2IUX) glycosylation sites, which are well folded and can be crystallised reproducibly in the same space group as the minimally glycosylated wild-type (Gordon *et al.*, 2003; Watermeyer *et al.*, 2006).

From these structures it can be seen that all of the potential N-linked glycosylation sites are on the surface, with the exception of the C-terminal site on the N domain, however it is unlikely that this site could be glycosylated. Glycan residues could be built into the structures at some, but not all, sites, and glycan chains were shown to make intermolecular and interhelical

contacts, possibly playing a role in stabilising the protein fold as well as the crystal lattice structure (Watermeyer *et al.*, 2006).

1.5.2. Structure overview

The crystal structures of the N and C domains of ACE share a high degree of structural homology. Each ACE domain is globular, consisting mostly of α-helices with a few β-sheets and 3_{10}-helices (Figures 1.1, 1.2 and 1.3). The CATH database (http://cathwww.biochem.ucl.ac.uk) divides an ACE domain into three subdomains (Figures 1.1 and 1.2).

Subdomain 1 is small and mainly α-helical, comprising the so-called "lid helices" α1 and α2, together with 3_{10}-helix H1 and α3 (tACE residues 40-121). α1 and α2 form an up-down bundle connected by a hairpin bend, while H1 and α3 pack against α1 and α2 in a helix-turn-helix motif (Figures 1.1 and 1.2).

Subdomain 2 is a large, complex, mainly α-helical orthogonal bundle with homology to domain 3 of neurolysin. It comprises 13 α-helices, five 3_{10}-helices and four β-strands in two regions of consecutive residues, 122-291 and 439-617 (numbered according to tACE). The core of the subdomain is formed by four long (up to 33 residues) α-helices; α8 and α20 on the protein surface are orthogonal to α15 and α17, which line the active cleft. This core is flanked on one side by an up-down bundle comprising α4, α5 and α6, which packs against the lid helices. Helices α4 and α5 together form one half of the bundle, broken in the middle by a β-hairpin formed by β1 and β2. Also surrounding the core are the shorter helices α7, α9, α16, α18, α19, H4, H6 and H7 and two short β-strands, β3 and β6. α17, α18, α19 and H7, together with helices from subdomain 3, form the floor of the active cleft. Two chloride ions are bound by subdomain 2. One chloride binding site (CL1) lies between the α4/α5/α6 bundle and the core of subdomain 2, and is formed by H6, α16, α6, and loops β6-H6 and β3-H4. The N-termini of α14 and α17, together with α7, form the second chloride binding site (CL2).

Subdomain 3 is a small mixed α and β domain comprising α10, α11, α12, α13, α14, H5, β4 and β5 (tACE residues 292-438). Together with subdomain 1, this domain forms a 3-layer αβα sandwich. The first layer is contributed by α1 of subdomain 1, the middle layer by the β-hairpin β4-β5 and the third by α13. Helix α13 forms a bundle with α14; the two together

making up the catalytic zinc-binding site; and with α11, which is on the surface of the molecule.

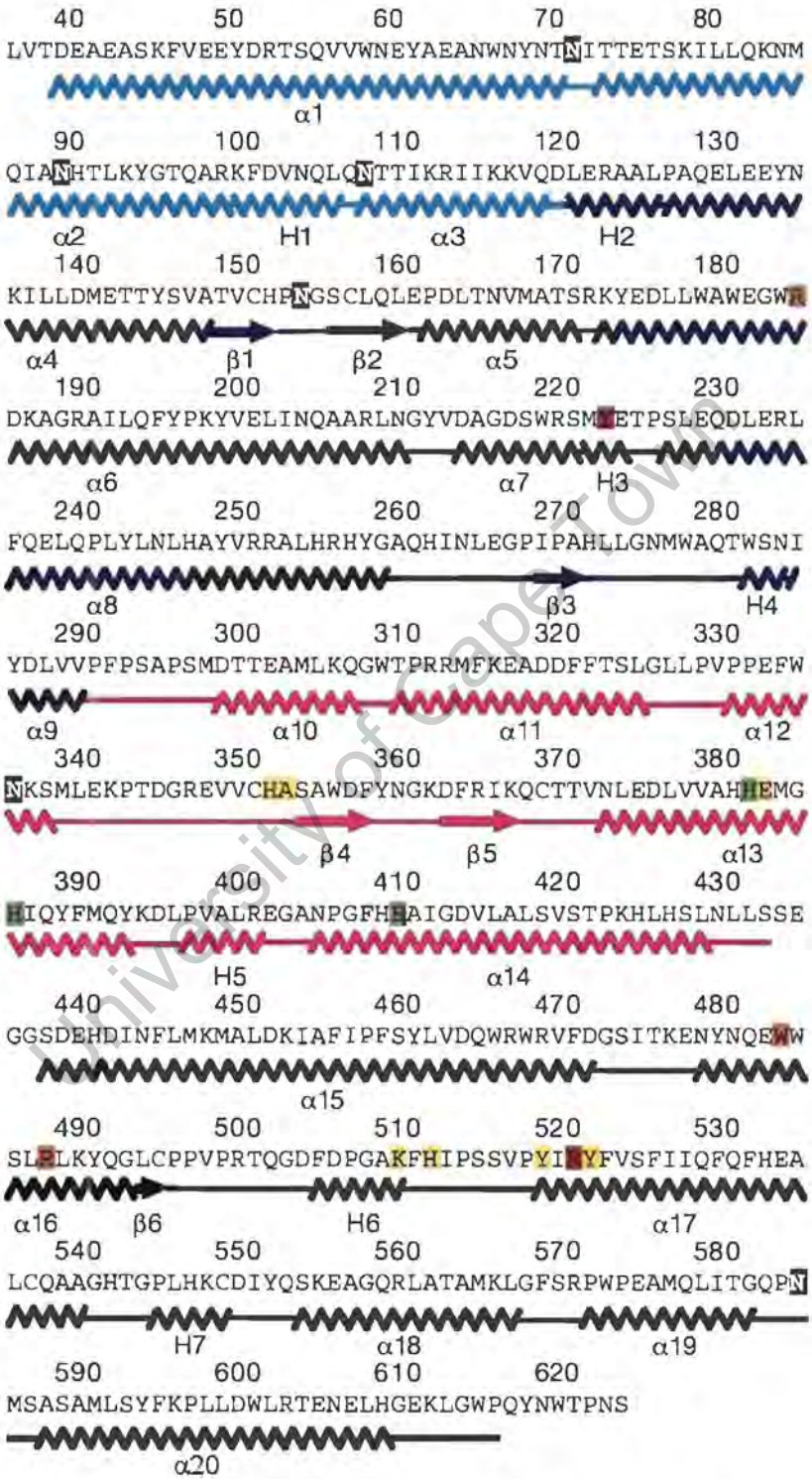


Figure 1.1. Sequence of human tACE, showing secondary structural elements. Residues in subdomains 1, 2, and 3 are coloured cyan, dark blue and pink, respectively. Secondary structural elements are labelled according to Natesh *et al.*, 2003; α = α-helices; β = β-strands, H = 3_{10}-helices. Zinc ligands are highlighted green, chloride-binding residues are highlighted orange (CL1) and red (CLl2), lisinopril-binding residues are highlighted yellow, and N-linked glycosylation sites are highlighted black.

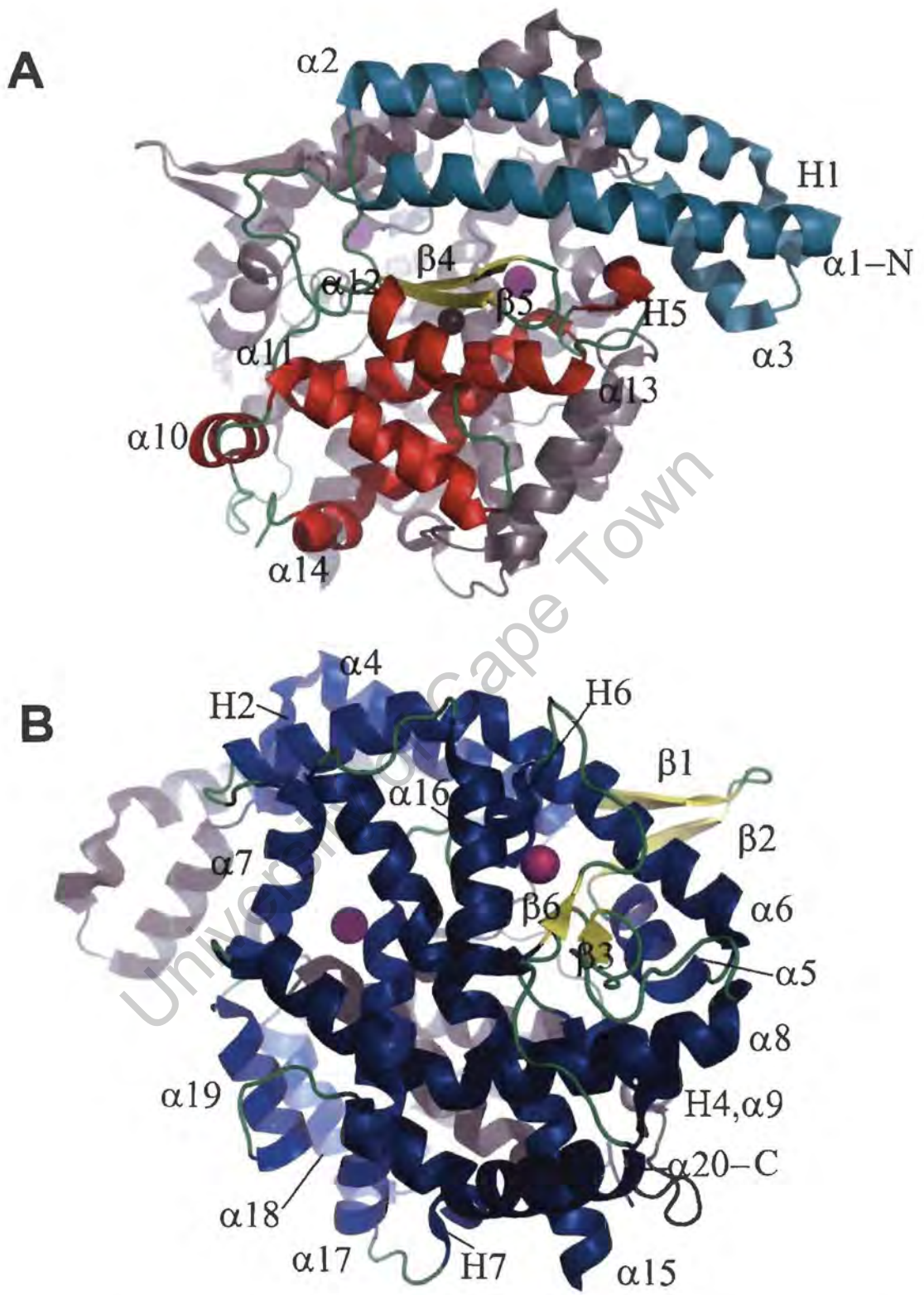


Figure 1.2. Cartoon representation of the three-dimensional structure of tACE (PDB-ID 1O8A), showing CATH domain classification. A and B are related by a 180° rotation about the vertical, with A showing subdomains 1 and 3 in colour, and B showing subdomain 2 in colour. Loop regions are coloured green, β-strands are coloured yellow, and helices are coloured according to subdomain: the lid helices (subdomain 1) are coloured cyan, the large α-helical orthogonal bundle (subdomain 2) is dark blue and the β-sandwich (subdomain 3) is red. The catalytic zinc ion and chloride ions are shown as black and magenta spheres, respectively. Secondary structure elements are labelled according to figure 1.1, and the N- and C-termini are marked.

The zinc-binding site is buried at the bottom of a deep cleft between subdomain 3 and the large subdomain 2, which effectively divides the globule in half (Figure 1.3). The cleft is closed off further by the lid helices which form subdomain 1. This domain organisation explains the specificity of ACE for short peptide substrates, since larger folded proteins would not be able to access the buried active site. Openings connecting the solvent to the active cleft are present in the structures, but these are narrow and while the active site is solvent-accessible, it is unclear how a bulky peptide substrate could gain access (Sturrock *et al.*, 2004).

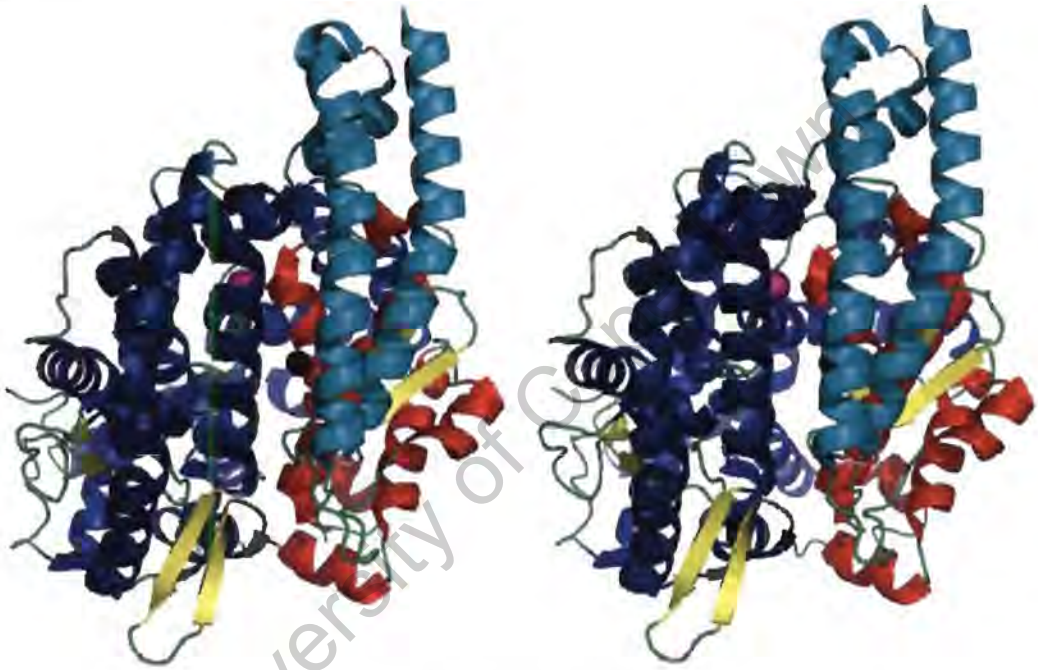


Figure 1.3. Cartoon representation of human tACE (PDB ID 1O8A), in stereo, showing the active zinc ion buried in a cleft between subdomains 2 and 3. The secondary structure elements are coloured according to subdomain, as in figure 1.2. The zinc ion is shown as a black sphere, and the chloride ions as magenta spheres.

The inclusion of much of the linker region in the N domain crystal structure allows for conjecture as to the position of the C domain relative to the N domain (Corradi *et al.*, 2006), although these suggestions rely on the assumption that the linker retains its native conformation in the crystal structure despite the absence of the C domain. Observations that a degree of interdomain cooperativity is involved in the cleavage of some substrates (Marcic *et al.*, 2000; Binevski *et al.*, 2003; Georgiadis *et al.*, 2003; Woodman *et al.*, 2005) support a close association between the active sites rather than a more extended configuration. Based on this model, the flexible loop region C128-C136 in the N domain with close proximity to the linker, has been proposed to form part of the domain interface, interacting either with

glycan chains on the lid helices of the C domain, or with a portion of the C domain surface (Figure 1.4; Corradi *et al.*, 2006).

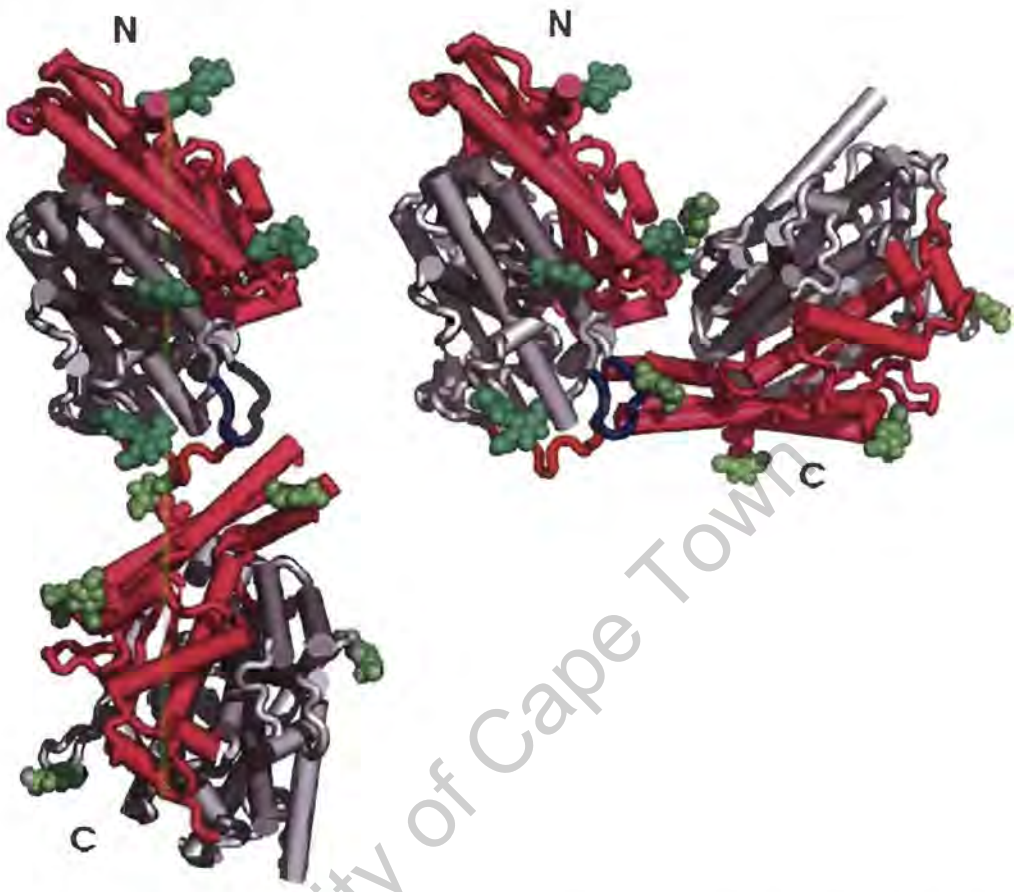


Figure 1.4 Possible arrangements of the N and C domains, based on the N domain and tACE crystal structures. Figure taken from Corradi *et al.* (2006). The N domain is in a similar orientation to Figure 1.3, with subdomains 1 and 3 coloured pink and subdomain 2 grey. The linker region is coloured red and the flexible loop C128-C136 dark blue. Glycan residues that were detected in the crystal structures are shown in dark green and light green sphere representation for the N and C domains, respectively.

1.5.3. Chloride ions in the structure of ACE

Two chloride ions are present in the structures of the C domain, which has been shown to be activated by chloride, although the mechanism of activation is not clearly understood (Wei *et al.*, 1991). One of the chloride ions (CL1) is coordinated by R186 (NE1)[1], W485 (NE), R489 (NH1) and a water molecule, and the other (CL2) is coordinated by R522 (NE), Y224 (OH) and a water molecule. Both sites are surrounded by a "shell" of hydrophobic residues. The CL1 site is 20 Å from the zinc ion, separated from the substrate C terminus by the side chains of W279, K511 and Q281, which play a crucial role in substrate docking. It has been

[1] tACE numbering is used for all references to C domain residues in this section since this numbering is used in the C domain crystal structures. To convert to full-length ACE numbering, add 576.

proposed that the function of this chloride ion is to position K511 correctly for substrate binding (Tzakos *et al.*, 2003). The CL2 site is closer to the zinc ion (10 Å) than the CL1 site, and separated from the active cleft only by R522, which has been shown to be crucial for chloride dependence (Liu *et al.*, 2001). This chloride site has been suggested to hold the active site in the correct conformation for substrate binding, possibly by pulling Y520 and Y523 into position to interact with the substrate, or by preventing R522 from interfering with substrate binding by holding it in place (Tzakos *et al.*, 2003; Sturrock *et al.*, 2004).

In the N domain, which is not activated by chloride ions to as great a degree as the C domain (Wei *et al.*, 1991), R186 of the CL1 binding site is replaced by H164 and no chloride ion was detected at this site in the N domain crystal structures. The CL2 site is conserved in the N domain and contains a bound chloride ion in the N domain crystal structures.

Tzakos and co-workers (2003) identified a putative conserved chloride ion channel for CL2 with homology to the *Salmonella enterica* serovar *typhimurium* CLC chloride channel, and proposed that these residues perform a similar function in ACE, creating an electrostatically favourable environment for chloride binding. They also proposed that E403 serves as an ion gate in the C domain, and that its replacement with R381 in the N domain is the reason for the lower chloride requirement of this domain. However in contrast to CLC, which is a membrane channel protein, the CL2 ion site of ACE is partially exposed to the external milieu; moreover the proposed channel residues are arranged in different orientations to those of the CLC ion channel, suggesting only the loosest functional conservation.

1.5.4. The structure of the active site

The catalytic zinc ion is located approximately in the centre of the molecule, at the bottom of the active cleft, forming a bottle-neck between a smaller substrate C-terminus binding chamber and a larger N-terminus binding chamber. Co-crystal structures with lisinopril (N and C domains), captopril and enalaprilat (C domain only) allowed the identification of the residues involved in substrate binding (Natesh *et al.*, 2003; Natesh *et al.*, 2004; Sturrock *et al.*, 2004; Corradi *et al.*, 2006). In addition, a co-crystal structure of tACE with the C-domain-selective inhibitor RXPA380 has shed light onto the residues important for domain-selectivity (Corradi *et al.*, 2007).

Active site interactions in the co-crystal structures are presented in Appendix B, and residues within 6 Å of any of the inhibitors (captopril, lisinopril, enalaprilat and RXPA380) in the active sites of the N and C domains are compared in Table 1.5. All the inhibitors take on similar extended conformations in the active site, and the conformation of lisinopril does not change significantly between the N and C domains.

Table 1.5 Residues within 6 Å of bound inhibitor in the crystal structures of ACE, showing their participation in inhibitor binding subsites. Highlights indicate residues that differ between the domains.

N domain	tACE (C domain)	Subsite
Q 259	W 279 (855)	S2': C terminus binding site
Q 259	Q 281 (857)	S2'
S 260	T 282 (858)	S2'
D 354	E 376 (952)	S2'
S 357	V379 (955)	S2'
H 361	H 383 (959)	S2' and zinc ligand
D 393	D 415 (991)	S2'
E 431	D 453 (1029)	S2'
K 432	K454 (1030)	S2'
F 435	F 457 (1033)	S2'
F 438	F 460 (1036)	S2': C-terminus binding site.
K 489	K 511 (1087)	S2': C-terminus ligand
Y 498	Y 520 (1096)	S2': C-terminus ligand
Y 501	Y 523 (1099)	S2'
F 505	F 527 (1103)	S2'
T 358	V 380 (956)	S2'/S1'
D 140	E 162 (738)	S1'
H 331	H 353 (929)	S1'
A 332	A 354 (930)	S1'
Q 355	D 377 (953)	S1'
H 491	H 513 (1089)	S1'
E 362	E 384 (960)	S1': ligand of zinc-binding oxygen atom
Y 501	Y 523 (1099)	S1': ligand of zinc-binding oxygen atom
D 43	N 70 (646)	S1
S 119	E 143 (719)	S1
S333	S 355 (931)	S1
E 389	E 411 (987)	S1 and zinc ligand
F 490	F 512 (1088)	S1
N 494	S 516 (1092)	S1
T 496	V 518 (1094)	S1
W 335	W 357 (933)	S1/S2
R 500	R 522 (1098)	S1/S2
Y 338	Y 360 (936)	S2
A 334	A 356 (932)	S2
D 336	D 358 (934)	S2
H 365	H 387 (963)	S2 and zinc ligand
Y 369	F 391 (967)	S2
Y 372	Y 394 (970)	S2
R 380	R 402 (978)	S2
R 381	E 403 (979)	S2
G 382	G 404 (980)	S2
H 388	H 410 (986)	S2.

The inhibitors bind to the active site zinc ion via their respective zinc binding groups; carboxyalkyl carboxylates in the case of lisinopril and enalaprilat, a sulphydryl group in captopril and a phosphinate in RXPA380. Comparison between the co-crystal structures of tACE with lisinopril, captopril and enalaprilat revealed a correlation between inhibitor potency and the number of polar protein-inhibitor interactions observed (Natesh *et al.*, 2004). Binding affinity is also strongly influenced by the strength of the zinc-binding groups, with the sulphydryl of captopril being stronger than the carboxylate of lisinopril and enalaprilat (Redelinghuys *et al.*, 2005).

The inhibitor C-terminal carboxylate interacts with a positively-charged patch formed by Y520, Q281 and K511, interactions which are thought to be important in determining the correct positioning of the substrate scissile bond at the active site (Kim *et al.*, 2003; Sturrock *et al.*, 2004; Naqvi *et al.*, 2005). This would explain why ACE can act as a tripeptidase or endopeptidase on substrates with amidated C-termini (Naqvi *et al.*, 2005).

1.6. ACE reaction mechanism

A general base mechanism of catalysis is thought to be conserved in zinc metalloproteases having a HEXXH zinc-binding motif, based on extensive catalytic and structural studies of related enzymes thermolysin and carboxypeptidase A, among others (Lipscomb & Sträter, 1996; Pelmenschikov *et al.*, 2002; Tzakos *et al.*, 2003). In ACE, the following details of the mechanism have been proposed (Figure 1.5; tACE numbering; Tzakos *et al.*, 2003; Sturrock *et al.*, 2004):

When no ligand is bound, the active site zinc ion is tetrahedrally coordinated by the two His side chains (H383 and H387) and downstream Glu (E411) of the HEXXH + E motif, with a water molecule or hydroxide ion as the fourth ligand, in an arrangement such as that seen in the unliganded structure of the N domain (PDB ID 2C6F). As a peptide substrate binds, its scissile carbonyl oxygen interacts with the zinc as a fifth ligand, forming a Michaelis complex. This arrangement displaces the zinc-bound water molecule towards E384 of the HEXXH motif, where it becomes strongly polarised between the positively charged zinc ion and the negatively charged glutamate, and hence more nucleophilic to attack the substrate carbonyl carbon.

Nucleophilic attack occurs, with the water being deprotonated by E384, and results in the formation of a tetrahedral transition state. In this state, the scissile carbonyl group has been converted to a negatively-charged geminal diolate (*gem*-diol) intermediate with two covalently bound oxygens. The proton from the attacking water is transferred by E384 onto the scissile amide nitrogen, and the intermediate quickly decomposes with peptide bond breakage, into carboxyl and amino leaving groups. These leaving groups are likely to be similar to those observed in the crystal structure of *Escherischia coli* ACE homologue dipeptidyl carboxypeptidase (Dcp), which was co-crystallised with the products of slow cleavage of a peptide inhibitor still bound in the active site (Comellas-Bigler *et al.*, 2005). The transition state is stabilised by interactions with conserved residues H331/353, A332/354 and Y501/523 (N domain/tACE numbering).

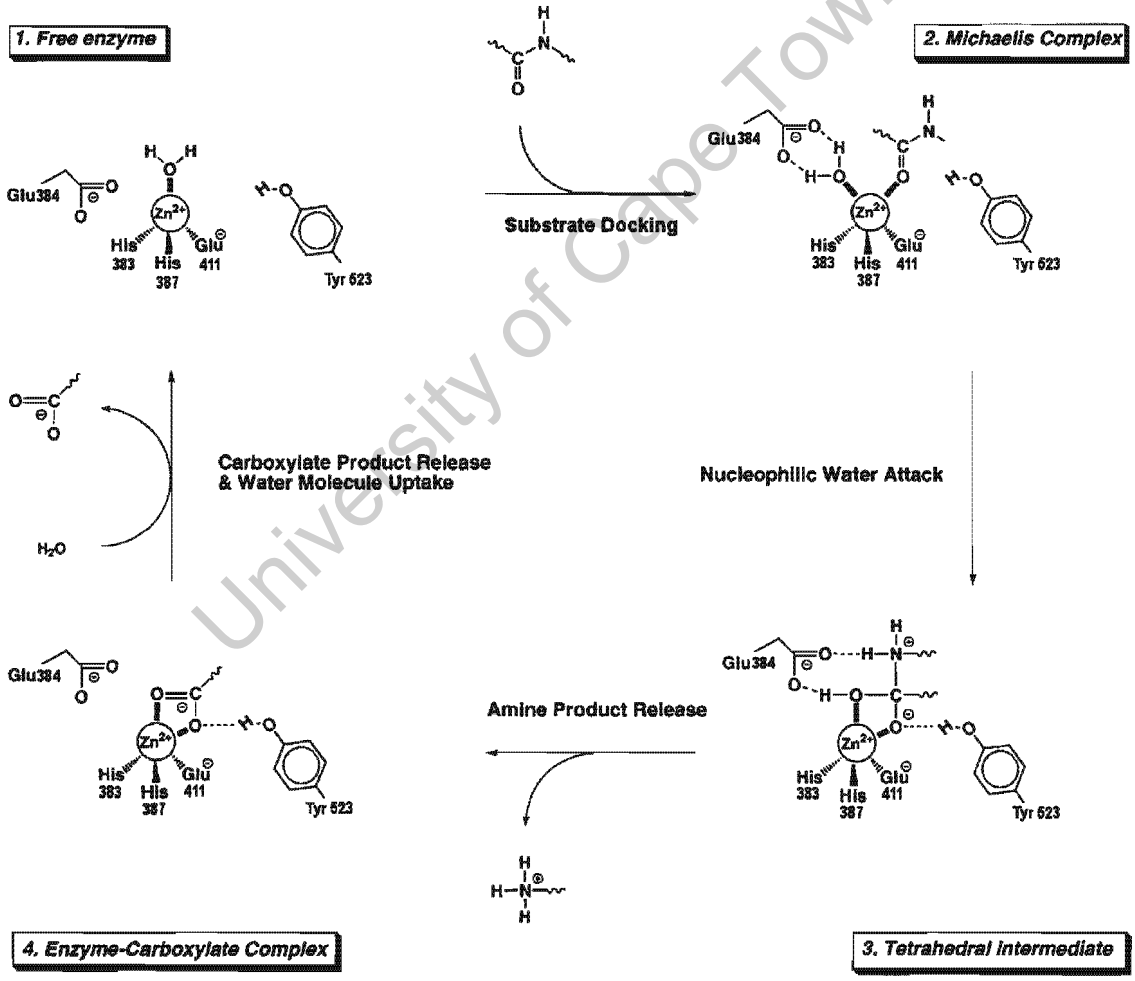


Figure 1.5 Proposed general base catalysis mechanism for ACE. Figure adapted from Pelmenschikov *et al.* (2002), with numbering for tACE. In the free enzyme (1) the zinc ion is tetrahedrally coordinated by H383, H387, E411 and a water molecule. A substrate docks into the active site forming a Michaelis complex (2) in which the zinc-bound water molecule is polarised between E384 and the zinc ion. The Michaelis complex is converted to a tetrahedral transition state (3) upon nucleophilic attack of the carbonyl carbon by the zinc bound water. The intermediate decomposes with bond breakage and product release, returning via an enzyme-carboxylate intermediate (4) to the unliganded state.

1.7. Structures of sequence homologues of ACE

1.7.1. Angiotensin-converting enzyme 2

Angiotensin-converting enzyme 2 (ACE2) is a recently-discovered mammalian homologue of ACE with its own distinct genetic locus (Tipnis *et al.*, 2000; Donoghue *et al.*, 2000). It is expressed in the heart, kidneys, testes and gastrointestinal tract and is in itself an interesting potential drug target, as it has been implicated in the regulation of cardiac function and the RAS (reviewed by Warner *et al.* (2004)). Moreover ACE2 has recently been shown to be the membrane receptor for the SARS coronavirus. This two-domain enzyme has an ACE-like domain with 42% sequence identity to the N domain of human ACE and a C-terminal domain with 48% identity to human collectrin (Warner *et al.*, 2004). The ACE domain of ACE2 acts as a carboxypeptidase, removing a single amino acid from the C-terminus of a number of peptide substrates, including neuropeptides and RAS and KKS-related peptides, some of which are also cleaved by ACE (Turner *et al.*, 2002).

Structures of unliganded and inhibitor-bound ACE2 have been solved recently (PDB ID 1R42 and 1R4L; Towler *et al.*, 2004), revealing close structural homology between ACE and the ACE domain of ACE2, however only fragments of the collectrin-like domain were defined in the crystal structures. The monopeptidase activity of ACE2 is explained by the observation that the P_2' pocket of this enzyme is occluded by an arginine side chain (Guy *et al.*, 2003; Towler *et al.*, 2004). Interestingly, in the unliganded ACE2 structure, the smaller subdomains 1 and 3 (the lid helices and adjacent β-sandwich) are hinged away from subdomain 2 (the larger zinc-binding orthogonal bundle), with residues at the outer edges moving apart by 20 Å relative to the closed state (Figure 1.6 (B); Towler *et al.*, 2004).

1.7.2. Insect ACE structures

Six ACE homologues occur in the fruit fly, *Drosophila melanogaster*, of which two, the single-domain enzymes AnCE and ACEr, have thus far been shown to encode active metallopeptidases with similar substrate specificity and inhibitor sensitivity to ACE (Williams *et al.*, 1996; Bingham *et al.*, 2006). A crystal structure of AnCE has been solved (PDB ID 1J38; Kim *et al.*, 2003), revealing close structural identity with the human enzyme, as expected from the high degree of sequence identity (41% and 42% to the N and C domains of human ACE, respectively). AnCE structures were also solved with captopril and lisinopril in

the active site (PDB ID 1J36, 1J37), revealing almost identical modes of substrate binding to those seen in the active site of tACE (Kim *et al.*, 2003).

1.8. Structural homologues of ACE

1.8.1. Structural classification of ACE and ACE homologues

The ACE fold is classified by the SCOP database (Structural Classification of Proteins; http://scop.mrc-lmb.cam.ac.uk/scop/) as being of the class "α and β" (having mainly antiparallel β-sheets with segregated α and β regions), with a "zincin-like" fold (containing mixed β-sheet with connections over the free side of the sheet), belonging to the "metalloproteases" ("zincins") catalytic domain superfamily and "neurolysin-like" family. This structural family includes the M2, M3 and M32 families of metalloproteases having a multihelical N-terminal half and a C-terminal thermolysin-like catalytic domain.

The classification of zinc metalloproteases has been reviewed extensively elsewhere (Hooper, 1994; Rawlings & Barrett, 1995; Spyroulias *et al.*, 2004), however it should be noted that the grouping of families and superfamilies has changed considerably in recent years as more structural information has become available, and is unlikely to remain static in the future. Briefly, according to the most recent review (Spyroulias *et al.*, 2004), the "zincins" include all metalloproteases containing the HEXXH zinc-binding motif, and this group can be sub-divided into "gluzincins", in which and third zinc ligand is a downstream glutamate, and "metzincins", which contain a conserved active site methionine. Within these major subdivisions, metalloproteases are grouped according to sequence conservation and active site structure (a complete list of families is available in the MEROPS database, http://merops.sanger.ac.uk), with the M2 family containing ACE-related enzymes, the M3 family comprising metalloproteases with homology to thimet oligopeptidase, and the M32 family having homology to carboxypeptidase Taq from *Thermus aquaticus*.

According to the SCOP database, other members of the "neurolysin-like" family for which structures have been determined are rat neurolysin and thermostable carboxypeptidase 1 from *Pyrococcus furiosus* (PfuCP) (Table 1.6; Figure 1.6 (C,G)). NCBI's macromolecular modelling database (MMDB) lists several additional structure neighbours of ACE, namely dipeptidyl carboxypeptidase from *Escherichia coli* (Dcp), oligoendopeptidase F from *Bacillus*

stearothermophilus (OligoF), and carboxypeptidase 1 from *Thermus thermophilus* (ThCP) (Table 1.6; Figure 1.6 (E,F,H)). Although these enzymes come from diverse sources and have negligible sequence identity, they share a high degree of structural homology (Appendix C).

Table 1.6 Structural homologues of ACE in the Protein Data Bank, showing sequence identity.

ACE homologue; source	Enzyme family	PDB ID	Identity (%) [a]
ACE2; *Homo sapiens*	M2	1R4F, 1R4L	42
AnCE; *Drosophila melanogaster*	M2	1J36, 1J37, 1J38	41
neurolysin; *Rattus norvegicus*	M3	1ILI	N/D
thimet oligopeptidase; *Homo sapiens*	M3	1S4B	N/D
dipeptidyl carboxypeptidase (Dcp); *Escherichia coli*	M3	1Y79	N/D
oligoendopeptidase F (OligoF); *Bacillus stearothermophilus*	M3	2HLN, 2HLJ	N/D
carboxypeptidase I (ThCP); *Thermus thermophilus*	M32	1WGZ	N/D
thermostable carboxypeptidase I (PfuCP); *Pyrococcus furiosus*	M32	1KA2, 1KA4, 1K9X	N/D

[a] % sequence identity with the N domain of human ACE, determined by pair-wise BLAST using the BLOSUM62 matrix (http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi). N/D: no identity detected.

1.8.2. Thimet oligopeptidase, neurolysin and Dcp

Neurolysin and its close homologue thimet oligopeptidase (60% identity) are well-characterised mammalian oligopeptidases of the M3 family. Their activity profiles are similar, both having endopeptidase activity on a broad range of peptide substrates. Both are active as monomers, and rat thimet oligopeptidase forms inactive oligomers by disulphide bond formation under oxidising conditions (Ray *et al.*, 2004). The best-understood role of neurolysin is the inactivation of the 13-residue neuropeptide neurotensin, by removal of a C-terminal tripeptide (Brown *et al.*, 2001). Thimet oligopeptidase degrades peptides released by the proteasome (Ray *et al.*, 2004). Differences in electrostatic conditions on the floor of the active site, which are contributed largely by a single α-helix, correlate to the differences in activity between the two enzymes (Ray *et al.*, 2004; Lim *et al.*, 2007). The presence of flexible loop regions in the active cleft is thought to allow the active site to adapt to accommodate a variety of sequences in the active site (Brown *et al.*, 2001).

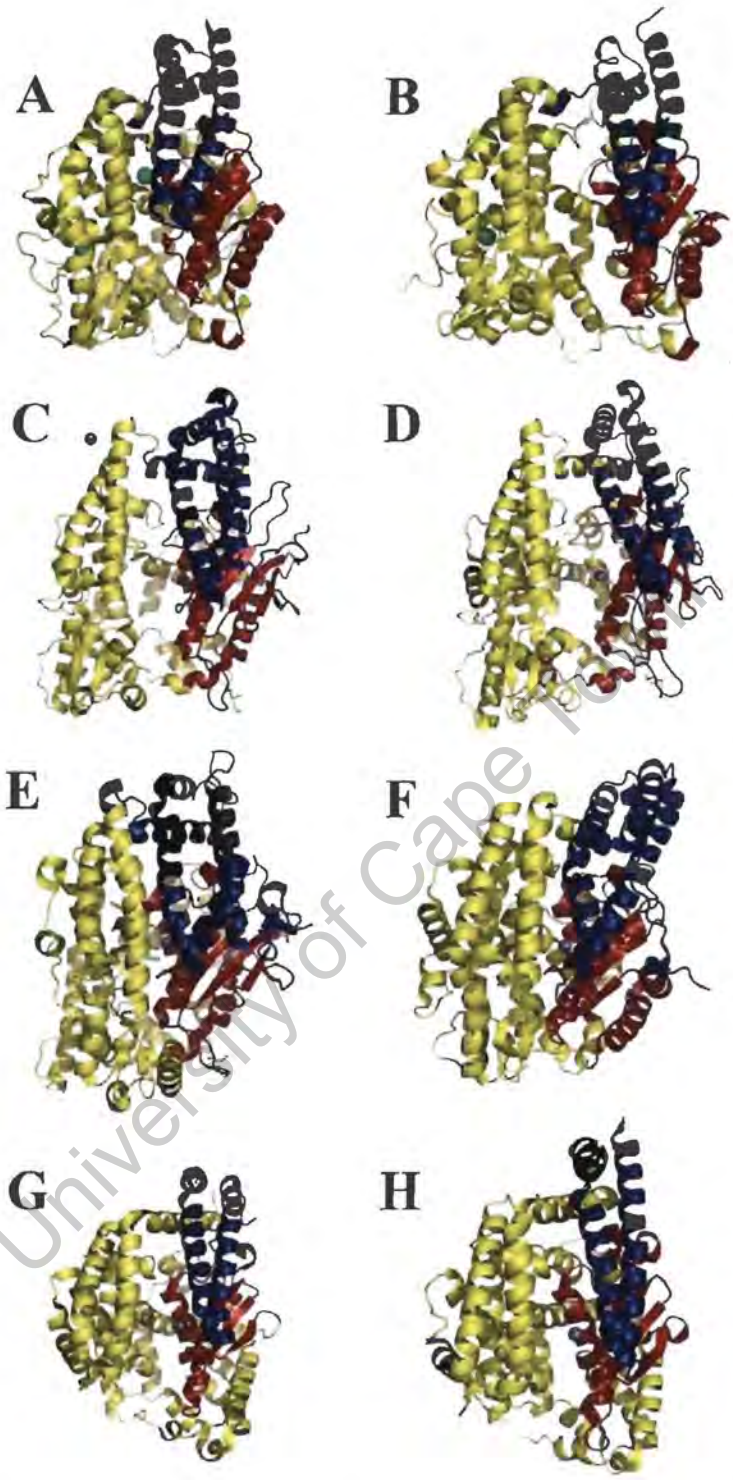


Figure 1.6 Structural homologues of ACE; cartoon representation showing subdomain structure and hinging about the active cleft. (A) human tACE (PDB ID 1O86), (B) ACE domain of unliganded human ACE2 (PDB ID 1R42), (C) rat neurolysin (PDB ID 1ILI), (D) human thimet oligopeptidase (PDB ID 1S4B), (E) *E. coli* dipeptidyl carboxypeptidase (PDB ID 1Y79), (F) *B. stearothermophilus* oligoendopeptidase F (PDB ID 2HLN), (G) *Pyrococcus furiosus* carboxypeptidase 1 (PDB ID 1KA2) and (H) *Thermus thermophilus* carboxypeptidase 1 (PDB ID 1WGZ). Subdomains 1, 2 and 3, identified by alignment with tACE, are coloured blue, yellow and red respectively. Zinc (magnesium in the case of (G)) and chloride ions are shown as grey and green spheres, respectively. Figure prepared using Pymol 0.99 (DeLano Scientific, Palo Alto, CA, U.S.A).

Dcp from *Escherichia coli* is another M3 zinc metallopeptidase, having 25% identity to neurolysin. Despite its low sequence identity to ACE, it has similar activity, acting as a dipeptidyl carboxypeptidase on ACE substrates and binding to ACE inhibitors (Comellas-Bigler *et al.*, 2005). Dcp is active as monomer in the cytoplasm where it catalyses the degradation of intracellular peptides (Comellas-Bigler *et al.*, 2005).

1.8.3. Thermostable homologues

OligoF, PfuCP and ThCP are all thermostable metallopeptidases, of interest due to their potential as industrial enzymes because they are active at temperatures close to 100 °C. PfuCP and ThCP are both members of the M32 family and share 35% identity, originating from archaea and bacteria respectively (Nagata *et al.*, 2004; Gerdts *et al.*, 2006), while OligoF is a bacterial M3 endopeptidase. PfuCP differs from other zincins in requiring a non-zinc metal cation such as cobalt for optimal activity, and in the formation of an active homodimer joined by a four helix bundle made between the lid helices of two monomers (Arndt *et al.*, 2002).

1.8.4. Conserved activity among structural homologues

These structurally similar enzymes from diverse sequence groups also share a selectivity for unfolded oligopeptide substrates, probably a result of the buried nature of their active sites (Brown *et al.*, 2001), as well as the ability to cleave a broad range of sequences. Peptide substrates are predicted to associate with the active cleft by formation of an twisted antiparallel β-sheet with the "edge-strand" of a conserved active site β-sheet in subdomain 3 (Comellas-Bigler *et al.*, 2005). This mode of substrate binding has been observed to be common to almost all proteases (Tyndall *et al.*, 2005). The registration of peptides at the active zinc ion is determined by the size and charge of the substrate N- and C-terminus binding pockets (Brown *et al.*, 2001; Ray *et al.*, 2004; Comellas-Bigler *et al.*, 2005).

1.8.5. Hinging about the active cleft

Although the active site is buried in a deep cleft in all cases, the degree of closure differs from structure to structure and from one crystal form to another (Figure 1.6; Ray *et al.*, 2004). For example, the unliganded structures of PfuCP and neurolysin align closely (Arndt *et al.*, 2002),

but the liganded structure of Dcp is considerably (~ 12 ° rotation) more closed about a similar hinge axis to that observed in ACE2 (Comellas-Bigler *et al.*, 2005). Thimet oligopeptidase also exhibits a slight hinge displacement relative to neurolysin (Ray *et al.*, 2004). Similar hinging could be predicted for unliganded ACE2 and a modelled open form of ACE by normal mode analysis and is thus probably an intrinsic feature of this type of fold (Watermeyer *et al.*, 2006).

Crystallographic atomic displacement parameters (temperature factors; Watermeyer *et al.*, 2006), the preferential binding of monoclonal antibodies to liganded, over unliganded, N domain (Skirgello *et al.*, 2006), and calorimetric studies that revealed a large entropic component to substrate binding (Andújar-Sánchez *et al.*, 2004), are also consistent with a closing active site model for ACE.

1.9. Structure-based design of ACE inhibitors

1.9.1. Structural determinants of domain-selectivity

The 55% identity of the N and C domains of ACE is not evenly distributed throughout the structure. For example, the lid helices (N domain 22-73 and C domain 613-675) are only 30% identical, while the residues between the zinc-binding HEMGH motif and downstream E (N domain 361-389 and C domain 904-932) are 86% identical (Appendix A).

Comparison of the crystal structures of tACE (C domain) and the N domain reveals 69% identity for the 42 residues within 6 Å of the inhibitors captopril, lisinopril, enalapril and RXPA380 (Table 1.5). If one excludes from this comparison zinc ligands, residues that bind the substrate C-terminal carboxylate, and residues that interact with substrate via backbone atoms, the overall sequence identity is only 51%, with nine of the thirteen variations occurring in the S_2' and S_1 subsites (Table 1.5).

This variation is reflected in the difference in affinity (*Ki*) of lisinopril for each domain in isolation (Table 1.3). For example, the slight C-domain-selectivity of lisinopril is thought to be the result of the substitution of tACE residues E162 and D377 with D140 and Q355 in the S_1' pocket of the N domain, which reduces the number of protein-inhibitor hydrogen bonds (Natesh *et al.*, 2004; Corradi *et al.*, 2006). The replacement of F391, V379 and V380 in the

S_2 and S_2' pockets of tACE with the more polar Y369, S357 and T358 in the N domain was proposed to account for the marked domain-selectivity of RXPA380, based on the tACE co-crystal structure (Corradi *et al.*, 2007).

1.9.2. Design of novel domain-specific ACE inhibitors

The domain-selectivity of ACE suggests a new direction for structure-based drug design. Although numerous ACE inhibitors are available for clinical use, these are not without undesirable side effects, which can include persistent cough, anaemia, impaired renal function, and potentially life-threatening angioedema (Sica, 2004). Elevated levels of BK and other peptides have been associated with ACE inhibitor-related cough and angioedema, implicating the inhibition of cleavage of non-AI peptides in the onset of these symptoms (reviewed by Acharya *et al.* (2003)).

The situation is made more complex by the observed domain cooperativity in the cleavage of some substrates *in vitro* (Binevski *et al.*, 2003; Georgiadis *et al.*, 2003; Andújar-Sánchez *et al.*, 2004), as well as by the presence of local tissue RAS and the differential effects of AI signalling via two receptors (reviewed by Zaman *et al.* (2002), Acharya *et al.* (2003) and Fleming *et al.* (2006)). However, the recent observations that the C domain of ACE is largely responsible for the cleavage of AI in mice, *in vivo*, while BK cleavage can be carried out by either domain (Fuchs *et al.*, 2004; Fuchs *et al.*, 2007), strengthens the idea that a C domain-specific ACE inhibitor might have a different clinical profile from the currently available treatments, possibly eliminating some of the side effects common to ACE inhibitor therapy today.

1.9.3. Aims of this study

Two sets of inhibitors have recently been designed and synthesised based on known domain-specific inhibitors keto-ACE (hereafter referred to as kAP) and lisinopril (Redelinghuys *et al.*, 2006; Nchinda *et al.*, 2006a; Nchinda *et al.*, 2006b). Of the sets of derivatives, two of the ketone inhibitors, kAW and kAF, and a lisinopril-derivative, lisW, all derived from their parent compounds by P_2' substitution, were found to be strongly selective for the C domain (Redelinghuys *et al.*, 2006; Nchinda *et al.*, 2006a). The aim of this study is to determine X-ray crystal structures of the C domain of human ACE in complex with these novel domain-

selective inhibitors, in order to determine their mode of binding, thereby gaining an understanding of the active site components involved in domain-selectivity.

CHAPTER 2: Materials and Methods

2.1. Preparation of the protein sample

2.1.1. Minimally glycosylated tACE construct

A minimally glycosylated human tACE mutant, tACE-G13, was used as a C domain equivalent for X-ray crystallographic studies. This mutant was generated previously by deletion of the N-terminal O-glycosylated 36 residues of tACE, and elimination of all but the first and third N-glycosylation sites by Asn-Gln point mutation (Figure 2.1; Gordon *et al.*, 2003). tACE-G13 has previously been shown to be reproducibly crystallisable and to have an identical conformation to minimally glycosylated wild-type tACE (Watermeyer *et al.*, 2006).

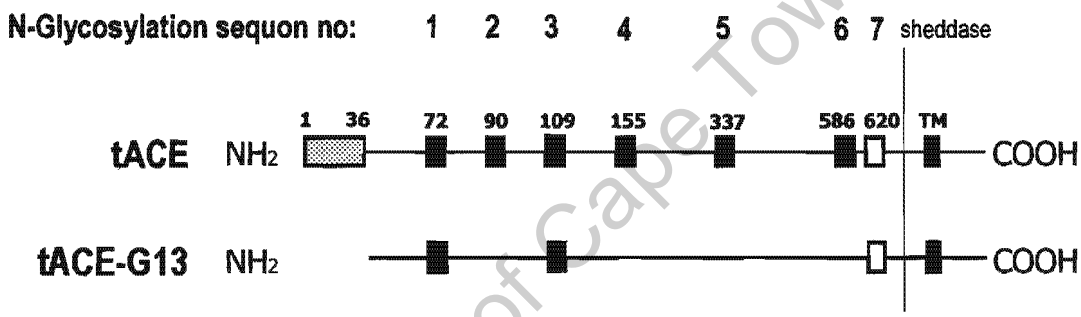


Figure 2.1 Schematic diagram of tACE constructs, showing wild-type tACE and tACE-G13, the minimally-glycosylated construct used for crystallisation. Full-length wild-type tACE has an N-terminal O-glycosylated region (grey box), six N-glycosylation sites (filled squares), one non-glycosylated N-glycosylation sequon (open square), and a C-terminal transmembrane domain (TM). tACE-G13 lacks the O-glycosylated region and has only two intact N-glycosylation sites. Expressed protein is shed into the medium by a sheddase that cleaves between Arg 627 and Ser 628.

2.1.2. Expression

A stably transfected Chinese hamster ovary cell line expressing the tACE-G13 construct from the vector pLEN-tACEΔ36-g13 (Gordon *et al.*, 2003) was used to produce protein for crystallisation. Cells were grown to confluence in 10% FCS medium (50% DMEM (Highveld Biological, South Africa), 50% Ham's medium (Highveld Biological, South Africa), supplemented with 10% heat-inactivated foetal calf serum (Invitrogen Life Technologies, United Kingdom)). Expression was induced with 40 μM $ZnCl_2$ in 2% FCS medium (50% DMEM, 50% Ham's medium supplemented with 2% heat-inactivated foetal calf serum). Expressed tACE-G13 was shed from the cell membrane into the 2% FCS

medium by the action of an endogenous sheddase between R627 and S628 (Figure 2.1; Ehlers *et al.*, 1996).

2.1.3. Purification

Soluble tACE-G13 was purified from culture medium by lisinopril affinity chromatography, as described previously (Ehlers *et al.*, 1989; Ehlers *et al.*, 1991; Gordon *et al.*, 2003). Briefly, cell culture medium containing shed tACE-G13 was applied to a Sepharose-28-lisinopril affinity column and unbound protein was eluted with 0.5 M NaCl, 20 mM Hepes at pH 7.5. Bound tACE-G13 was eluted as a single peak (detected by measuring the absorption at 280 nm) using 50 mM borate at pH 9.5. Peak fractions were assayed for ACE activity and fractions containing ACE were pooled, dialysed against 2 L of 0.5 mM Hepes (pH 7.5) overnight and concentrated to between 1 and 5 mg/ml using an Amicon Ultra-4 spin column (Millipore).

The protein concentration in purified samples was determined by a Bradford protein assay (Bio-Rad protein micro-assay). Concentrated samples were shown to contain a single purified species of the correct molecular weight (~ 70 kDa) by 10% denaturing SDS-PAGE with Coomassie staining.

2.1.4. ACE activity assay

ACE activity in the harvested medium, elution peak fractions, column flow-through and dialysate was determined by measuring the hydrolysis of Hippuryl-His-Leu (Sigma) in a fluorimetric assay (Friedland & Silverstein, 1976; Schwager *et al.*, 2006) as follows. Samples were incubated for 15 minutes at 37 °C in the presence of 5.65 mM Hippuryl-His-Leu, 5.2 mM NaOH, 0.052 M Hepes (pH 7.5) and 0.31 M NaCl. The hydrolysis reaction was stopped by the addition of 750 µl of 0.28 M NaOH. This mixture was then incubated at room temperature for ten minutes in the presence of 8 mM o-phthaldialdehyde, allowing the formation of the fluorescent adduct of o-phthaldialdehyde and L-His-L-Leu. 3 N HCl (100 µl) 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. 1 mU of ACE activity is defined as the production of 1 nmole L-His-L-Leu per minute.

2.2. *Preparation of inhibitor samples*

kAP-derived inhibitors (5S)-5-[(N-benzoyl)amino]-4-oxo-6-phenyl-hexanoyl-L-tryptophan (kAW) and (5S)-5-[(N-benzoyl)amino]-4-oxo-6-phenyl-hexanoyl-L-phenylalanine (kAF), as well as lisinopril-derivative (SR)-1-(N(2)-(1-carboxy-3-phenylpropyl)-L-lysyl)-L-tryptophan (lisW), were synthesised previously (Nchinda *et al.*, 2006a; Nchinda *et al.*, 2006b).

Stocks for crystallisation trials were made by dissolving the lyophilised inhibitors in water, or in a small volume of DMSO and then diluting in water, where necessary. Trials were carried out using solutions of the two separated enantiomeric forms of lisW, lisW-S and lisW-R, differing in the chirality of the P_1 alpha carbon.

2.3. *Crystallisation*

Co-crystallisation trials were carried out at a range of conditions based on the crystallising conditions for tACE-G13 (Watermeyer *et al.*, 2006).

Various concentrations of inhibitors were pre-mixed with 2-5 mg/ml tACE-G13 (in 5 mM Hepes, pH7.5), and 2 µl of these mixtures were added to 2 µl reservoir buffer in crystallisation drops. Hanging drop experiments were carried out at 16 °C, over 1 ml reservoirs comprising 5-20 % PEG 4000, 10 µM $ZnSO_4$, 10-200 mM sodium acetate, pH 4.0-6.5. To prevent condensation on the cover slides, buffers were chilled to 4 °C before setting up trials, and wells were covered with 500 µl of oil (3 parts paraffin, 2 parts silicon oil, Hampton; Chayen, 1997). All buffers were filtered at 0.22 µM.

For seeding experiments, inhibitor-free microcrystal seeds were grown in drops containing 1.45 mg/ml tACE-G13 over a reservoir containing 10 mM sodium acetate pH 4.7 (Merck), 15 % PEG 4000 (Fluka), 10 µM $ZnSO_4$. An eyelash was used to streak seeds into fresh drops containing tACE-G13 mixed with inhibitors.

2.4. Data collection and processing

2.4.1. Data collection

Crystals were transported in their hanging drops to beamline BM14(UK) of the European Synchrotron Radiation Facility, Grenoble, France. From the drops, suitable crystals were mounted in loops with magnetic bases, dipped briefly in cryo-protectant (20% glycerol, 80% reservoir buffer), plunge-frozen in liquid nitrogen and stored in labelled vials containing liquid nitrogen. These vials were inserted into a cartridge for handling by the robotic data collection apparatus.

Data were collected at 100 K from a single crystal per inhibitor, using a single wavelength of 1.033 Å with a maximum intensity of ~7.5×10^{11} photons. The beam had a focal area of 0.25 mm^2. Diffraction images were recorded on a MARMOSAIC225 charge-coupled detector, and had an area of 225 mm × 225 mm. Data collection statistics are presented in Chapter 3, Table 3.2.

2.4.2. Data processing and phase determination

Data were processed using *HKL2000* (Otwinowski & Minor, 1997). Phasing was carried out by molecular replacement (MR) with *EPMR 2.5* (Kissinger *et al.*, 1999) using protein atoms from the crystal structure of tACE-G13 (PDB ID 2IUL) as a model.

Model building was carried out using *O v9.0.7* (Jones, 1978). Model building involved identification, insertion and adjustment of the inhibitor molecules, minor adjustment of some protein side chains and surface loops, as well as addition of ions and water molecules. In order to minimize model bias, σ_A-weighted composite omit maps, calculated with simulated annealing (5% omission) using *CNS* (Brünger *et al.*, 1998) were used for the initial manual building and for much of the subsequent rebuilding, in addition to σ_A-weighted $2F_{obs}$-F_{calc} and F_{obs}-F_{calc} maps. The *CNS* water_pick protocol (Brünger *et al.*, 1998) was used to add water molecules, followed by visual inspection. Glycan residues were added to the structures by aligning the structure of tACE-G13 (PDB ID 2IUL) with the structure concerned and copying the coordinates of the glycans where clear density was present in all maps. Glycan residues were then adjusted and refined along with the rest of the structures. Ligand molecules were

built using *PRODRG* (Schüttelkopf & van Aalten, 2004) and *ARP/wARP* (Collaborative Computational Project, 1994) or by hand where *ARP/wARP* did not succeed in recognising ligand density.

Maximum likelihood minimization was carried out against the working set of crystallographic data, using *CNS* (Brünger *et al.*, 1998). In order to minimize model bias, the test set for cross-validation was selected to consist of the same set of reflection indices as were used during the solution of the structure of the MR model (PDB ID 2IUL). This was achieved by using *CNS* to merge the "TEST" column (containing the randomly-generated numerical flag used for the selection of test reflections) from the structure factor file that was used to solve the MR model structure, into the new co-crystal structure factor files. Since the MR model was of higher resolution than the present data, no extension of the test set was required.

Simulated annealing refinement was used to adjust selected regions to fit the data better. Individual isotropic atomic displacement parameters (B-factors) were refined in later rounds of refinement. Where surface side chains were disordered, and no discreet conformations could be detected in composite omit maps, a plausible conformer was selected and B-factors were allowed to refine to high values. This method was chosen due to the difficulty with which *CNS* handles missing atoms in residues.

Structure validation was carried out using *PROCHECK* and *SFCHECK* from the *CCP4* suite (Collaborative Computational Project, 1994), as well as *Molprobity* (Davis *et al.*, 2004). Hydrogen bonds and close contacts were identified using *HBPLUS* (McDonald & Thornton, 1994). Figures were generated using *Pymol v0.99* (DeLano Scientific, Palo Alto, CA, U.S.A).

CHAPTER 3: Results and Discussion

3.1. Introduction

The novel inhibitors kAW, kAF and lisW have been shown to have a high degree of selectivity for the C domain of human ACE (Redelinghuys *et al.*, 2006; Nchinda *et al.*, 2006a). kAW and kAF were derived from kAP (Redelinghuys *et al.*, 2006), and lisW from lisinopril (Nchinda *et al.*, 2006a), by P_2' substitution (Figure 3.1). Two enantiomers of lisW, lisW-S and lisW-R, were synthesised, differing in the chirality of the P_1 Phe alpha carbon (Figure 3.1). In order to elucidate the mechanism of domain-selectivity, the inhibitors were co-crystallised with human tACE, which has a sequence identical to that of the C domain of sACE.

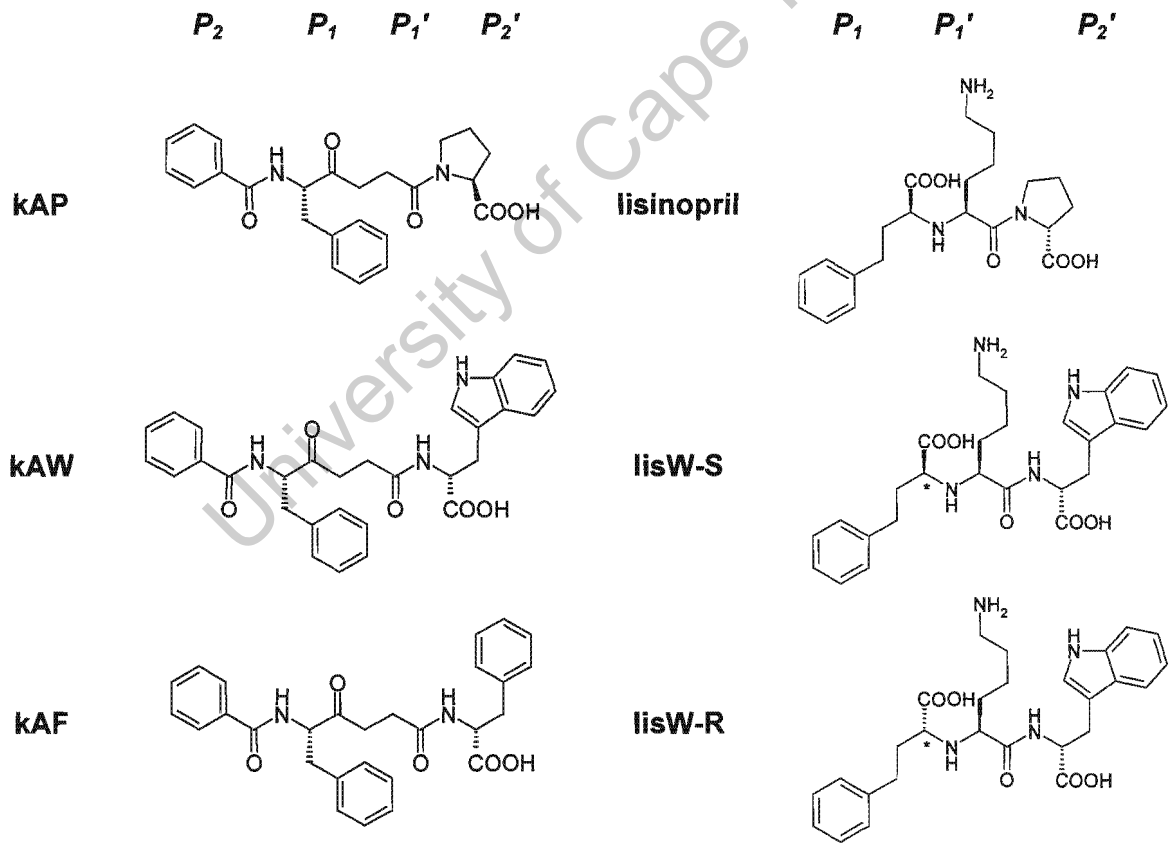


Figure 3.1 Chemical structures of keto-ACE (kAP) and its derivatives, kAW and kAF, and lisinopril and its derivative, lisW, showing residue positions relative to the zinc-binding group (P_2, P_1, P_1', P_2'). Both enantiomers of lisW are shown, and the chiral centre is indicated with a star (*).

3.2. Co-crystallisation of minimally glycosylated tACE with inhibitors

Crystallographic studies were carried out using a minimally glycosylated mutant of tACE, tACE-G13, lacking the N-terminal O-glycosylated region and having all but the first and third glycosylation sites (N72 and N109) knocked out by Asn-Gln point mutations (Gordon *et al.*, 2003). This mutant has been crystallised previously under similar conditions to those used for the crystallisation of minimally glycosylated wild-type tACE, in the same crystal form, and found to have essentially the same structure as the wild-type (Watermeyer *et al.*, 2006). Since tACE differs from the C domain of ACE only in its O-glycosylated region, tACE-G13 is equivalent to a minimally glycosylated C domain.

Initial co-crystallisation trials tested conditions based on those used for crystallisation of tACE-G13 in the absence of inhibitor, but including various concentrations of inhibitor pre-mixed with the protein before setting up the drop. Crystals of similar morphology to the tACE-G13 crystals obtained previously were grown in drops containing kAW (40 μM, with 1.25 mg/ml tACE-G13) and the S enantiomer of lisW (51 μM and 109 μM, 1.5 mg/ml tACE-G13), however the latter were non-diffracting. Diffraction data were collected from the kAW co-crystals grown using this method, processed to 3.0 Å, and phased using molecular replacement, however no continuous inhibitor density was present in the active site, indicating that these were not true co-crystals. Since these conditions had thus failed to produce suitable co-crystals, progressively higher concentrations of inhibitor were included in co-crystallisation experiments. However, higher concentrations of inhibitor, with and without a small amount of DMSO (0-10 %), were found to inhibit nucleation.

During the course of these crystallisation trials, the rate of crystal nucleation in inhibitor-free drops was found to depend strongly on the concentration of sodium acetate, with showers of microcrystals forming in drops containing lower concentrations of acetate ions (Figure 3.2 (B)). Some of these microcrystals were streaked as seeds into drops containing tACE-G13 and inhibitor, and co-crystals were grown within 14 days after seeding (Figure 3.2 (C)). The inhibition of nucleation by the presence of ACE inhibitors in the crystallisation drops was further demonstrated by the formation of showers of tiny microcrystals when seeds were streaked into inhibitor-free drops (Figure 3.2 (D)), versus the growth of fewer, larger crystals in the presence of inhibitors (Figure 3.2 (C)). The conditions for growth of the tACE-G13-inhibitor co-crystals used for data collection are presented in Table 3.1.

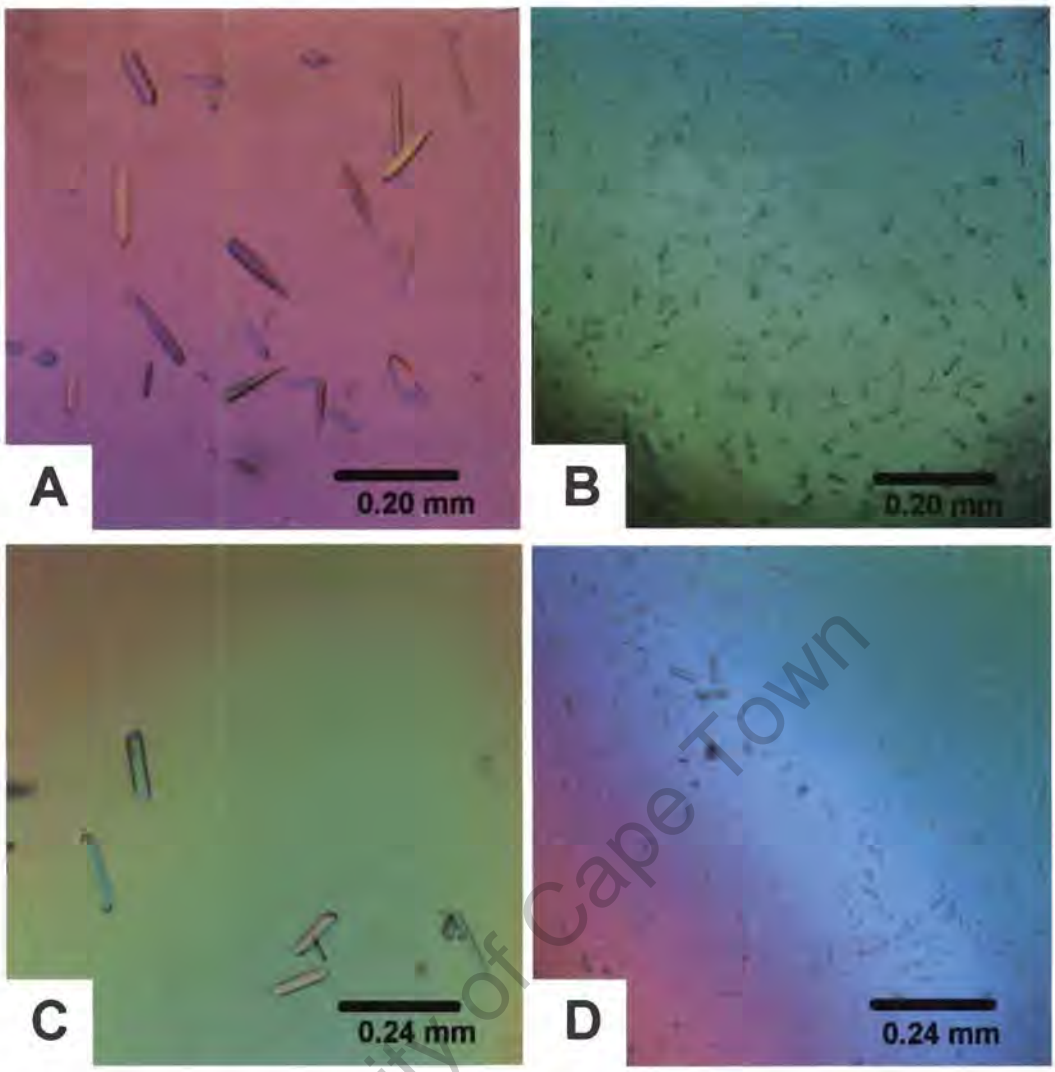


Figure 3.2 Co-crystallisation of tACE-G13 with inhibitors by streak-seeding. All experiments shown were 4 µl hanging drops incubated at 16 °C, with a starting protein concentration of 1.45 mg/ml and reservoirs containing PEG 4000 (15%), ZnSO4 (10 µM) and sodium acetate (various concentrations), pH 4.7. (A) tACE-G13 crystals grown in the absence of inhibitor, with 50 mM sodium acetate. (B) Microcrystals grown under the same conditions as in (A), but using 10 mM sodium acetate. (C) Microcrystals such as those shown in (B) were used as seeds to grow co-crystals such as these, grown in the presence of 250 mM lisW-S. (D) If microcrystals were streaked into drops containing no inhibitor, excessive nucleation resulted in the formation of showers of tiny microcrystals.

Table 3.1 Conditions for growth of tACE-G13-inhibitor co-crystals from which data were collected. All crystals used for data collection were grown over reservoirs containing 15 % PEG 4000, 50 mM sodium acetate and 10 µM $ZnSO_4$, pH 4.7.

Inhibitor	Protein concentration (mg/ml) [a]	Inhibitor concentration (µM) [a]	K_i C-domain (µM) [b]
lisW-S	1.45	250	7.0
lisW-R	1.45	50	26.3
kAW	1.00	1000	0.68
kAF	1.45	1250	0.83

[a] Starting conditions in drops; [b] (Nchinda *et al.*, 2006b; Watermeyer *et al.*, Appendix D).

3.3. Data collection and processing

Diffraction data were collected from single crystals of tACE-G13 grown in the presence of kAW, kAF and both enantiomers of lisW, having approximate dimensions of 0.2 × 0.004 mm. Data collection statistics are presented in Table 3.2. The datasets obtained from crystals grown in the presence of lisW-S and lisW-R will be referred to as lisW-A and lisW-B respectively, for reasons which will become clear during the course of the discussion.

Table 3.2: Diffraction data statistics.

	kAW	**kAF**	**lisW-A**	**lisW-B**
Resolution range (Å)	43-2.18	43-2.17	43-2.40	43-2.30
Highest resolution shell (Å)	2.26-2.18	2.25-2.17	2.49-2.40	2.38-2.30
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell dimensions (Å)				
a	59.85	60.38	60.07	56.38
b	85.04	85.07	85.04	84.79
c	134.92	135.51	135.09	133.96
Unit cell volume (A^3)	686667	696078	690161	640407
No. reflections	130768	158385	186516	84906
No. unique reflections	34837	33306	27439	26578
Redundancy [a]	3.8 (3.2)	4.8 (4.1)	6.8 (4.9)	3.2 (2.9)
Completeness (%) [a]	95.0 (95.2)	88.3 (80.2)	99.1 (95.1)	89.9 (86.4)
Refined detector distance (mm)	180.8	193.2	193.2	193.5
Refined mosaicity (°) [b]	0.57	0.81	0.68	1.05
$\langle I/\sigma(I)\rangle$ [a]	10.9 (2.4)	14.4 (2.6)	13.86 (2.4)	12.08 (2.2)
R_{merge} (%) [a, c]	10.9 (42.1)	8.7 (39.2)	12.1 (44.6)	8.7 (46.5)

[a] Values in parentheses refer to the highest resolution shell; [b] (Otwinowski & Minor, 1997); [c] $R_{merge} = [\Sigma_h\Sigma_j |I_j(h) - \langle I(h)\rangle|] / \Sigma_h\Sigma_j |I_j(h)|$, where $I_i(h)$ and $\langle I(h)\rangle$ are the ith and the mean measurements of the intensity of reflection h, respectively.

The resolution cut-off for structure determination was determined by the high R_{merge} values at higher resolutions, and by a drop in the signal-to-noise ratio below 2.0. Despite the small size of the crystals, the diffraction data were of a resolution suitable for co-crystal structure determination. The kAW and kAF datasets fall within the resolution range of other datasets collected previously from tACE crystals at synchrotron sources (2.25 to 1.82 Å), while the

lisW datasets are of slightly lower resolution. The space group and unit cell dimensions for all four datasets are the same as those of the other tACE structures in the PDB.

3.4. Co-crystal structures of kAP-derived inhibitors

3.4.1. Structure solution, model building and refinement

Initial phases were determined by molecular replacement using protein atoms from the structure of tACE-G13 (PDB ID 2IUL) as a model. There was little change in unit cell dimensions of the co-crystals in comparison with the model structure, the molecular replacement solutions had high scores, and visual inspection of the initial electron density maps showed that the phases required little further refinement, indicating a conservation of the structure of the protein component between the liganded an unliganded structures.

Structures of kAW and kAF in the active site of tACE-G13 were determined to 2.18 and 2.17 Å, respectively. The progress of model building was monitored by the changes in R_{cryst} and R_{free} after each cycle of model-building and refinement (Figure 3.3). Structure refinement was considered complete when R_{free} ceased to decrease, and when any errors of stereochemistry introduced during the model building process had been addressed. Model refinement and validation statistics are presented in Table 3.3.

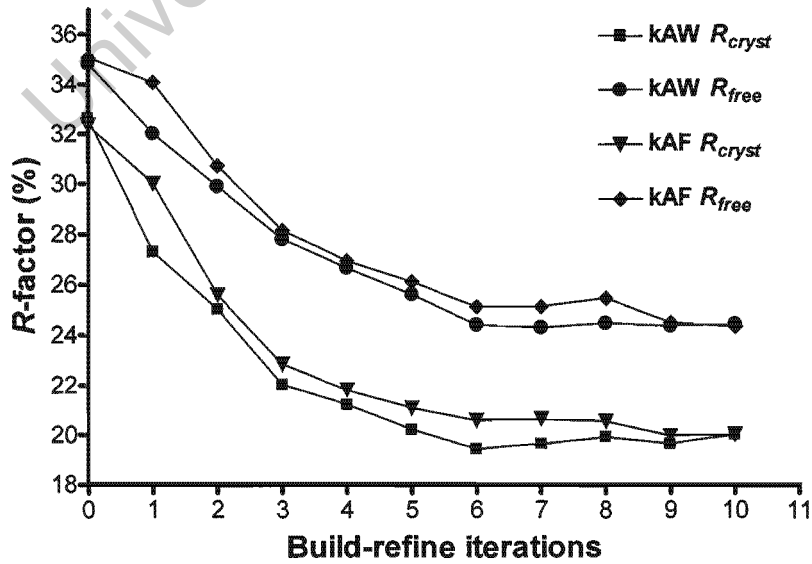


Figure 3.3 Progress of model building and refinement for the kAW and kAF datasets, as demonstrated by changes in R_{cryst} and R_{free}. $R_{cryst} = \sum_h |F_o - F_c| / \sum_h F_o$, where F_o and F_c are the observed and calculated structure factor amplitudes of reflection h of the working set, respectively; R_{free} is equal to R_{cryst} for h belonging to the test set of reflections.

Table 3.3: Model refinement statistics for the co-crystal structures of kAW and kAF in the active site of tACE-G13.

	kAW	**kAF**
No. reflections, working set [% of total measured]	31173 [94.7]	30340 [95.4]
No. reflections, test set [% of total measured]	1753 [5.3]	1459 [4.6]
R_{cryst} (%) [a,b]	20.0 (26.6)	20.0 (25.0)
R_{free} (%) [a,c]	23.9 (29.7)	24.3 (26.4)
Luzzati error (Å) [test set]	0.24 [0.30]	0.25 [0.31]
SigmaA error (Å) [test set]	0.22 [0.21]	0.20 [0.21]
R.m.s. deviations from ideality		
bonds (Å)	0.01	0.01
angles (°)	1.4	1.3
dihedral angles (°)	22.2	21.2
improper angles (°)	1.2	1.0
Mean B-factor (Å^2) [min-max]		
all atoms	21.6 [2.8-69.0]	22.7 [5.9-72.5]
ligand atoms	17.3 [11.5-24.4]	21.8 [14.4-35.2]
B-factor r.m.s. deviation (Å^2)		
bonded mainchain atoms	1.26	1.19
bonded sidechain atoms	2.05	1.91
angle mainchain atoms	2.00	1.91
angle sidechain atoms	2.95	2.79
Number of atoms		
solvent atoms	254	230
protein atoms [d]	4750	4777
ions	3	3
inhibitor atoms	39	36
glycan atoms	67	67
% unit cell volume occupied	49.51	49.49
Ramachandran plot, % residues in favoured regions	97.40	96.23

[a] Values in parentheses represent the highest resolution shell: 2.18-2.26 Å for kAW and 2.17-2.25 Å for kAF;
[b] $R_{cryst} = \sum_h |F_o - F_c| / \sum_h F_o$, where F_o and F_c are the observed and calculated structure factor amplitudes of reflection h of the working set, respectively; [c] R_{free} is equal to R_{cryst} for h belonging to the test set of reflections;
[d] excluding duplicated atoms in alternate conformations.

The final *R*-factors are what could be expected for data of this resolution and containing flexible regions. The use of the same test set of reflections as was selected for solution of the original structure of tACE-G13 used for molecular replacement (PDB ID 2IUL) resulted in the selection of test sets comprising 5.3% and 4.6% of reflections, close to the recommended 5% (Brünger, 1992).

The high upper limit of B-factors (individual isotropic atomic displacement parameters) for all atoms is the result of the inclusion of side chain atoms having ambiguous density by selecting a plausible conformation and allowing the B-factors to refine to high values as an indication of disorder. The average B-factors are thus a better indication of the quality of the models in this case. Atoms with very low B-factors include both protein and water atoms. The lowest B-factor in the kAW structure is that of R489 (Nη1), which is a chloride ligand and is thus probably stabilised by the chloride ion. The lowest B-factor in the kAF structure is that of K454 (Nζ), which lines the active site; there is no clear reason for the high degree of order for this atom. Low B-factors on water molecules (such as waters 1002 (B = 4.4), 1007 (B = 5.23) and 1016 (B = 5.44) of the kAW structure) could indicate partial occupancy of the sites by chloride ions, however since the B-factors of these water molecules are comparable to the B-factors some protein atoms, this evidence is not conclusive. The ligand B-factors indicate a high degree of order for most atoms, with a few high B-factors arising from poor density for the P_2' Phe of kAF.

The electron density around the N and C termini and some surface loops (103-109, 295-300, 433-440) was weak, indicating disorder in the crystal lattice. Depending on the quality of the density maps, these residues were included with high B-factors or deleted, resulting in differing numbers of protein atoms between the four models. In particular, residues 297, 435 and 436 could not be built in the kAW structure, while other residues in these loops have high B-factors. Loops 295-300 and 433-440 have been observed previously to be partially disordered in all of the ACE crystal structures solved to date and their location across one end of the active cleft may play a role in allowing the cleft to open during substrate binding (Watermeyer *et al.*, 2006). Loop region 103-109 also has high B-factors and contains an N-linked glycosylation site. Apparent disorder in this region may thus be real structural variation or may result from heterogeneity in glycosylation state. None of these loops participates directly in inhibitor binding.

The backbone torsion angles of all of the residues built fell in the allowed or favoured regions of the Ramachandran plot, as calculated using *Molprobity* (Davis *et al.*, 2004). Ramachandran plots are shown in Figure 3.4.

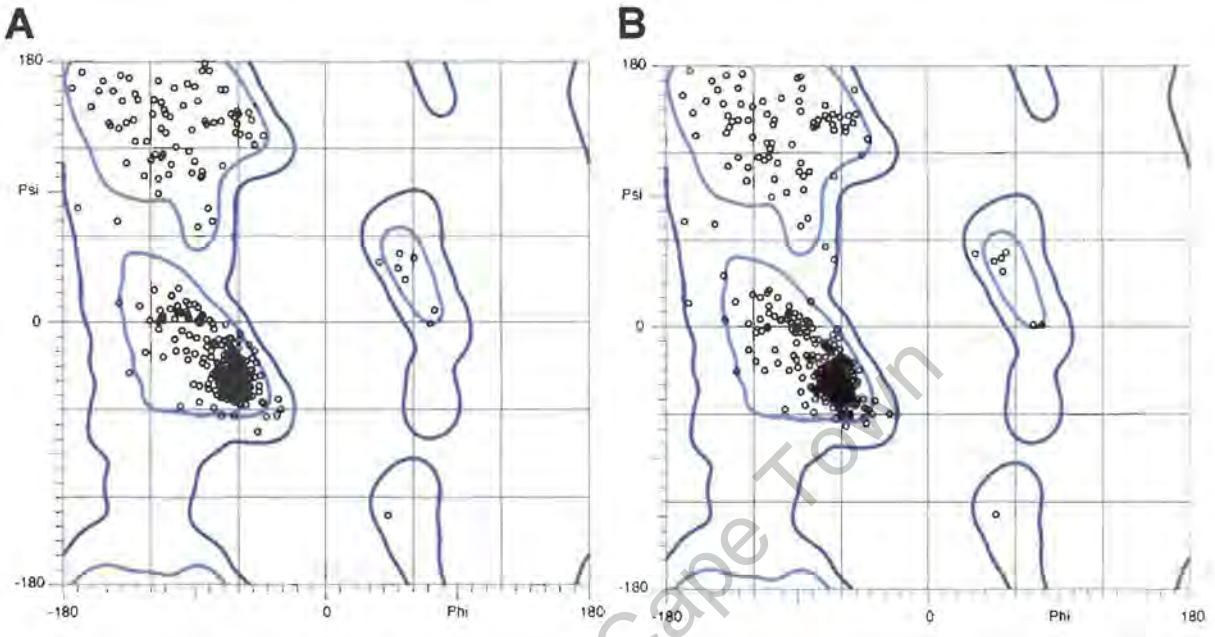


Figure 3.4 Ramachandran plots for the co-crystal structures of kAW (A) and kAF (B) in the active site of tACE-G13. Favoured regions are enclosed by the light blue contours and additionally allowed regions by the dark blue contours. Plots were generated using *Molprobity* (Davis *et al.*, 2004).

3.4.2. Structure overview

Co-crystal structures of kAW and kAF with tACE-G13 were solved by molecular replacement and refined to final crystallographic *R*-factors of 20.0 % in both cases (see Table 3.3). In both the kAW and kAF co-crystal structures, the protein component is in an identical conformation to the inhibitor-free tACE-G13 structure (PDB ID 2IUL), and to the minimally glycosylated wild-type tACE structure (PDB ID 1O86), as evidenced by low all-atom r.m.s. deviations (Table 3.4). Overall, this structure is mostly alpha helical and ellipsoid in shape, containing two chloride ions, with the zinc ion buried deep in the active site cleft (Figure 3.5). The lowest r.m.s. deviations in atomic positions between structures were seen in the zinc-binding residues and at the second chloride-binding site (CL2), which is close to the active site and is conserved in the N and C domains.

Table 3.4. Root mean square deviations between co-crystal structures of tACE-G13 with kAW (kAW) and kAF (kAF), unliganded tACE-G13 (PDB ID 2IUL) and minimally-glycosylated wild-type tACE bound to lisinopril (tACE-LPR; PDB ID 1O86).

Structures aligned		R.m.s. deviations (Å)						
Reference	Comparison	overall [a]	side [b]	main [c]	zinc [d]	CL1 [e]	CL2 [f]	active site [g]
kAW	kAF	0.60	0.79	0.27	0.09	0.52	0.17	0.17
kAW	tACE-G13	0.61	0.82	0.26	0.18	0.59	0.19	0.33
kAF	tACE-G13	0.67	0.87	0.37	0.19	0.47	0.19	0.35
kAW	tACE-LPR	0.69	0.91	0.32	0.17	0.57	0.16	0.35
kAF	tACE-LPR	0.76	0.98	0.40	0.19	0.49	0.17	0.38
tACE-G13	tACE-LPR	0.50	0.65	0.27	0.14	0.33	0.18	0.19

[a] all atoms; [b] side chain atoms; [c] main chain atoms; [d] zinc-binding residues; [e] residues within 5 Å of chloride CL1 (W182, R186, D507, P508, R489, W485, W486); [f] residues within 5 Å of chloride CL2 (M223, Y224, P407, P519, Y520, I521, R522); [g] residues within 6 Å of kAW and kAF in the active site (N70, E143, Q281-T282, H353-W357, E376, V379-V380, H383-E384, H387, F391, Y394, H410-E411, D415, D453-K454, F457, F460, K511-H513, S516, V518, Y520, R522-Y523, F527).

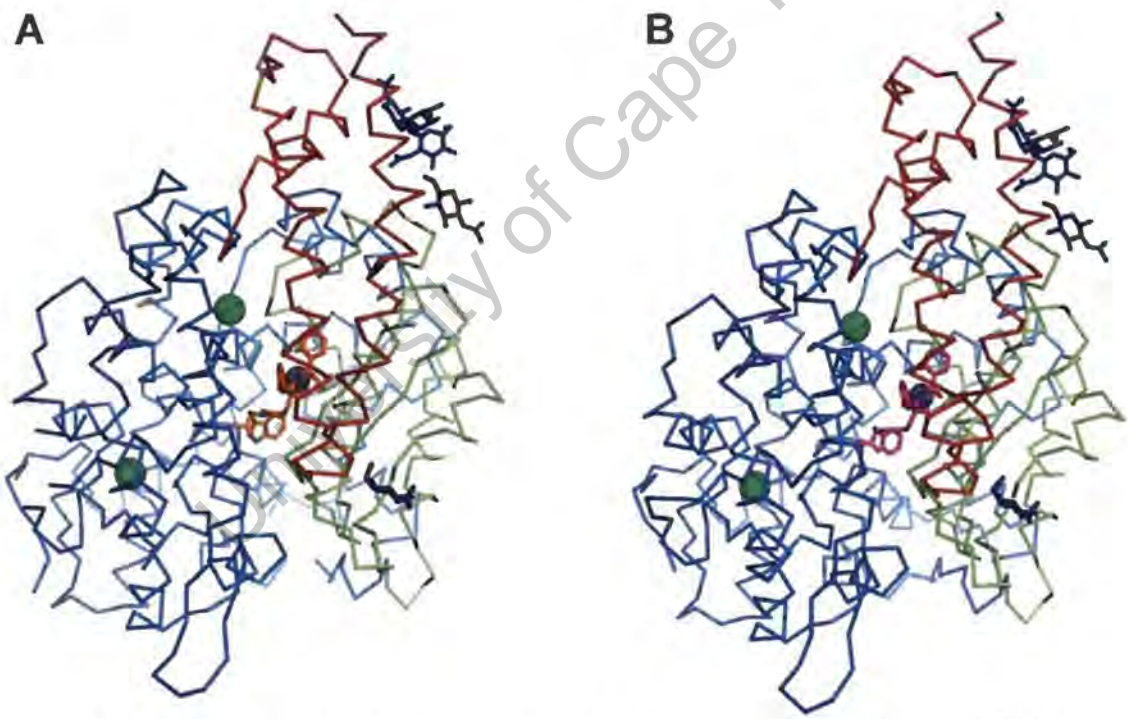


Figure 3.5 Cartoon representation showing overviews of the structures of kAW (A) and kAF (B) in the active site of tACE-G13. The protein backbone is shown in ribbon representation, with subdomains 1, 2 and 3, as described in chapter 1, coloured red, blue and light green, respectively. The zinc and chloride ions are shown as grey and green spheres, respectively, glycan residues are shown as dark blue sticks, kAW as orange sticks and kAF as magenta sticks.

The inhibitors were built into clear density present in the active site in σ_A-weighed F_{obs}–F_{calc} and composite omit maps (Figure 3.6). Weak density around the P_2' side chain of kAF indicates a degree of flexibility, and is also expressed in the higher B-factors of these atoms.

Nine residues that interact with kAW and kAF directly, or indirectly via solvent atoms, are substituted in the N domain and may thus contribute to their domain-selectivity (Figure 3.6).

3.4.3. Trapping of a tetrahedral transition state in the active site

These inhibitors have a ketone zinc-binding moiety analogous to the carbonyl oxygen of the scissile peptide bond, however the density maps surprisingly indicated the presence of a hydrated geminal diolate (*gem*-diol) at the active site, in which two oxygens are attached to the carbonyl carbon (Figures 3.6 and 3.7).

In zinc metalloproteases, peptide bond hydrolysis is thought to occur via a tetrahedral transition state, which forms upon nucleophilic attack of the carbonyl carbon by an activated zinc-bound water molecule at the active site (Lipscomb & Sträter, 1996; Pelmenschikov *et al.*, 2002). This is followed by protonation of the scissile amide nitrogen and subsequent peptide bond cleavage. The *gem*-diol observed here is reminiscent of the tetrahedral transition state, indicating that the replacement of the scissile peptide bond nitrogen with a carbon atom in these inhibitors (Figure 3.1) led to the trapping of the hydrated transition state at the active site (Figure 3.7).

Similar transition states have been observed previously in the structures of carboxypeptidase A (CPA) co-crystallised with ketomethylene inhibitors 5-benzamido-2-benzyl-4-oxopentanoic acid (BOP) and 5-amino-(N-t-butoxycarbonyl)-2-benzyl-4-oxo-6-phenylhexanoic acid (Christianson *et al.*, 1987; Shoham *et al.*, 1988; structures not available in the PDB). In these structures, the hydrated form of the ketone was observed at the enzyme active site despite the fact that this form does not occur (< 1%) in aqueous solution. Active site stabilisation of *gem*-diols is commonly observed for aldehyde inhibitors (for example, PDB IDs 1LAN (Sträter & Lipscomb, 1995) and 1CBX (Mangani *et al.*, 1992)) and inhibitors containing activated electrophilic ketones (for example, PDB IDs 1OD1 (Coates *et al.*, 2003) and 1NJs (Zhang *et al.*, 2003)). However, the hydration of a ketone in the absence of an activating group, as seen here, has to the best of our knowledge only been reported in two other crystal structures, those of CPA mentioned above.

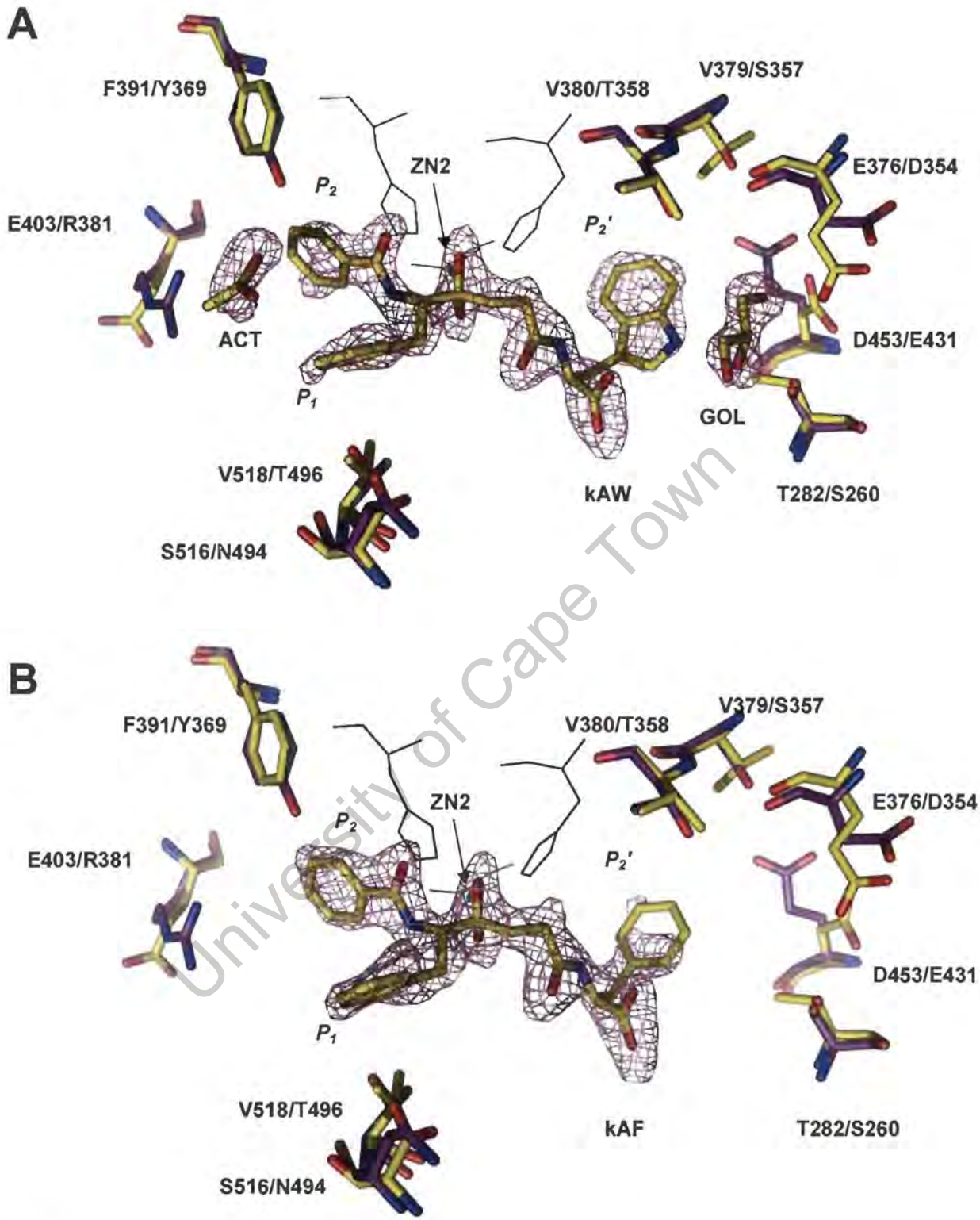


Figure 3.6 Stick representation of kAW (A) and kAF (B) in the active site of tACE-G13. The inhibitor molecules are shown in yellow sticks in density from the first σ_A-weighted F_{obs}-F_{calc} difference density map (purple mesh) into which the inhibitor model was built, contoured at 3σ. In A, the active site-associated glycerol (GOL) and acetate (ACT) molecules are also shown in σ_A-weighted F_{obs}-F_{calc} difference density. The position of the active site zinc ion (ZN2) is shown as a green non-bonded sphere, indicated with an arrow, and zinc-binding residues are shown in black lines in the background. Inhibitor-contacting residues that differ between domains are shown in yellow (tACE residues from the co-crystal structure) and purple (N domain residues from the aligned N domain structure, PDB ID 2C6F) sticks. Residue labels are given as tACE/N domain, and the P_1, P_2 and P_2' moieties are labelled.

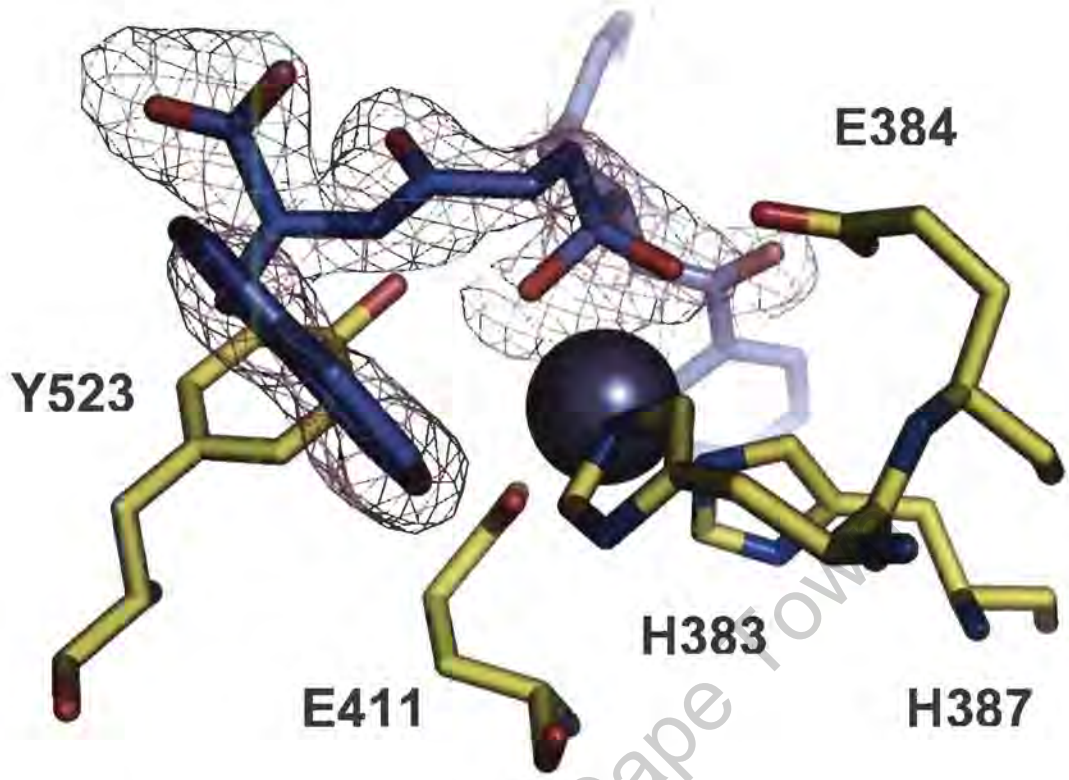


Figure 3.7 Stabilisation of the hydrated gem-diol transition state in the active site. The ketone inhibitor kAW (blue sticks) is shown in σ_A-weighted F_{obs}-F_{calc} density (purple mesh), as in Figure 3.6. The view is orientated to show the *gem*-diol moiety coordinating the active zinc ion (grey sphere). Zinc-coordinating residues and those involved in stabilising the intermediate are shown as yellow sticks and labelled.

The *gem*-diol oxygens in the structures of kAW and kAF have partial negative charges and interact with the active site zinc ion (Table 3.5), in similar positions to the carboxyl oxygens of tACE-bound lisinopril and enalaprilat (PDB IDs 1O86 and 1UZE). Y523 and E384, which has been proposed to shuttle a proton between the attacking nucleophilic water and the leaving amide, also make hydrogen bonds with the *gem*-diol (Figure 3.7). In the structure of CPA bound to BOP the *gem*-diol has longer zinc-coordinating distances (2.51 and 2.91 Å; Christianson *et al.*, 1987) than those observed here, however it is clear from the overlap of the *gem*-diol oxygens of kAW and kAF with the zinc-binding oxygens of the other inhibitors that these positions are characteristic of the ACE active site.

Thus, the visualisation of the *gem*-diol in the present structures demonstrates that in the absence of a leaving amide nitrogen, the catalytic residues of ACE are able to stabilise the tetrahedral transition state. This stabilisation probably plays a transient role in overcoming the energy barrier between substrate and products during catalysis.

3.4.4. Active site interactions of kAW and kAF

kAW and kAF take on a similar conformation in the active site to that seen in previous inhibitor co-crystal structures. The P_2' Trp and P_1 Phe align closely with the equivalent groups of the C-domain-selective inhibitor RXPA380 in the active site of tACE, although the backbone of RXPA380 is constrained to a different path around the P_1' Pro moiety (Figure 3.8). Docking studies of kAP predicted similar binding sites to those observed here for the P_1 and P_2 moieties, although the ketone group was not modelled in the hydrated state (Tzakos & Gerothanassis, 2005).

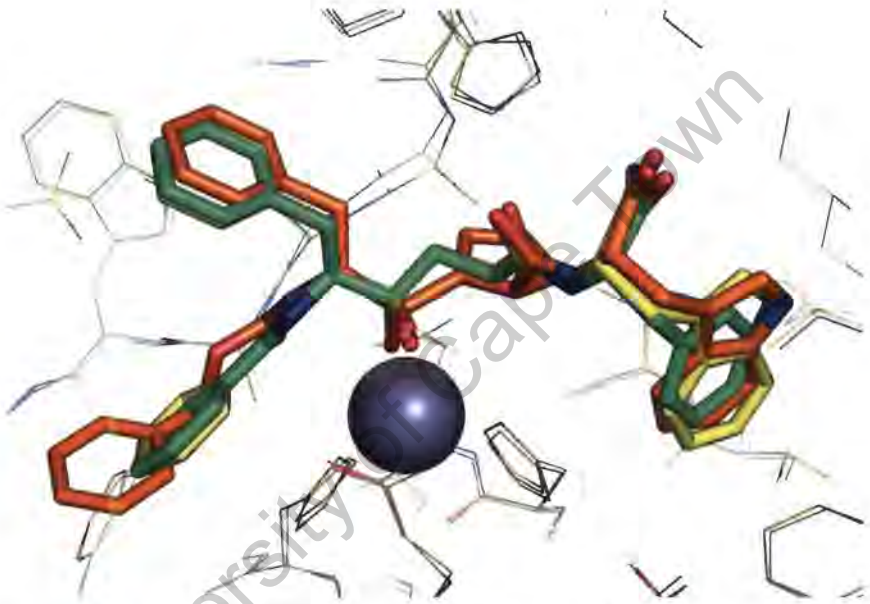


Figure 3.8 Alignment of the co-crystal structures of tACE with kAW (yellow sticks), kAF (green sticks) and RXPA380 (PDB ID 2OC2; orange sticks) in the active site. Structures were aligned by pair-wise alignment of the alpha carbons of the zinc-binding residues. The zinc ion is shown as a grey sphere, and adjacent protein components of all three structures are shown in lines of the same colour as the inhibitors to indicate the degree of alignment.

There is very little variation in the active site of tACE in any of the crystal structures solved to date. Notable exceptions are adjacent residues D453 and S284, which appear to be flexible, taking on different conformations in the minimally glycosylated wild-type tACE structures from the tACE-G13 structures, and Q281, which appears to swivel slightly about χ_1 and χ_3 to make room for the bulky Trp moiety of kAW. The details of the polar contacts and non-polar proximities involved in inhibitor binding are presented in Tables 3.5 and 3.6, respectively.

kAW makes nine polar contacts (potential hydrogen bonds and ionic interactions) with seven protein side chains, while kAF makes nine polar contacts with six protein side chains (Table 3.5). The difference are due to small changes in the orientation of Q281, as mentioned above,

and to negligible changes in E384 which nonetheless affect hydrogen-bond detection. Additional hydrogen bonds are made with four water molecules, whose positions are conserved between the structures, as well as with a glycerol molecule derived from the cryo-protectant in the kAW structure. These solvent molecules form part of a hydrogen bonding network connecting the inhibitor to other active site residues and to the bulk solvent (Figure 3.9).

Table 3.5 Polar contacts and close proximities between the inhibitors kAW and kAF, and protein or solvent atoms.

Inhibitor moiety	Inhibitor atom [a]	Interacting residue, atom [a]	N domain equivalent	Distance (Å) [b]		Polar and *nonpolar* contacts via solvent atom
				kAW	kAF	
P_2'	OXT	LYS 511, Nζ	LYS 489	2.73	2.76	-
	OXT	GLN 281, Nε2	GLN 259	(3.76)	2.89	-
	OXT	TYR 520, OH	TYR 498	2.63	2.76	-
	OXT	HOH 1006, O (1009)	-	(3.47)	3.20	***W279***, Q281, K511, GOL, water
	O	HOH 1006, O (1009)	-	2.79	2.78	***W279***, Q281, K511, GOL, water
	O	HOH 1186, O (1175)	-	2.67	2.59	***V380, GOL,*** water
	NE1	GOL 704, O2	-	2.79	-	N277, Q281, T282
	CE3 (CE1)	HOH 1162, O (1199)	-	2.93	2.95	***V380***, ***H383***, water
	N	HOH 1162, O (1199)	-	3.18	3.22	***V380***, ***H383***, water
P_1'	OAA	HIS 513, Nε2	HIS 491	2.78	2.84	-
	OAA	HIS 353, Nε2	HIS 331	2.80	2.66	-
Gem-diol	OAE	ZN2	ZN2	2.23	2.07	-
	OAE	TYR 523, OH	TYR 501	2.34	2.41	-
	OAE	HOH 1165, O (1081)	-	3.38	3.32	E411, ***V518,*** R522, Y523
	OAF	ZN2	ZN2	2.42	2.52	-
	OAF	GLU 384, Oε2	GLU 362	2.58	2.80	-
	OAF	GLU 384, Oε1	GLU 362	3.08	3.11	-
P_1	NBA	HOH 1165, O (1081)	-	3.10	3.08	E411, ***V518,*** R522, Y523
P_2	OAC	ALA 356, N	ALA 334	2.73	2.83	-
	OAC	GLU 384, Oε1	GLU 362	3.32	(3.53)	-

[a] kAF atom names and water numbers are given in parentheses where these differ from those of kAW.
[b] Values in parentheses are greater than the maximum hydrogen-bonding distance (3.4 Å).

The active site glycerol molecule in the kAW structure is highly ordered and was clearly present in σ_A-weighted F_{obs}-F_{calc} and composite omit maps calculated early in the refinement process (Figures 3.6 (A) and 3.9 (A)). It is linked via several hydrogen bonds to the water network in the S_2' pocket, and to polar and charged residues T282, Q281, N277 and E376 (Figure 3.9 (A)). The conformation of the Trp moiety of kAW does not appear to be affected by the glycerol however, since it takes the same conformation as the P_2' Trp of RXPA380 (PDB ID 2OC2; Corradi *et al.*, 2007), which does not have a glycerol at this position. Moreover, no glycerol-like density was present at this position in σ_A-weighted F_{obs}-F_{calc} and composite omit maps generated using the RXPA380 structure factors from the PDB (Figure 3.10).

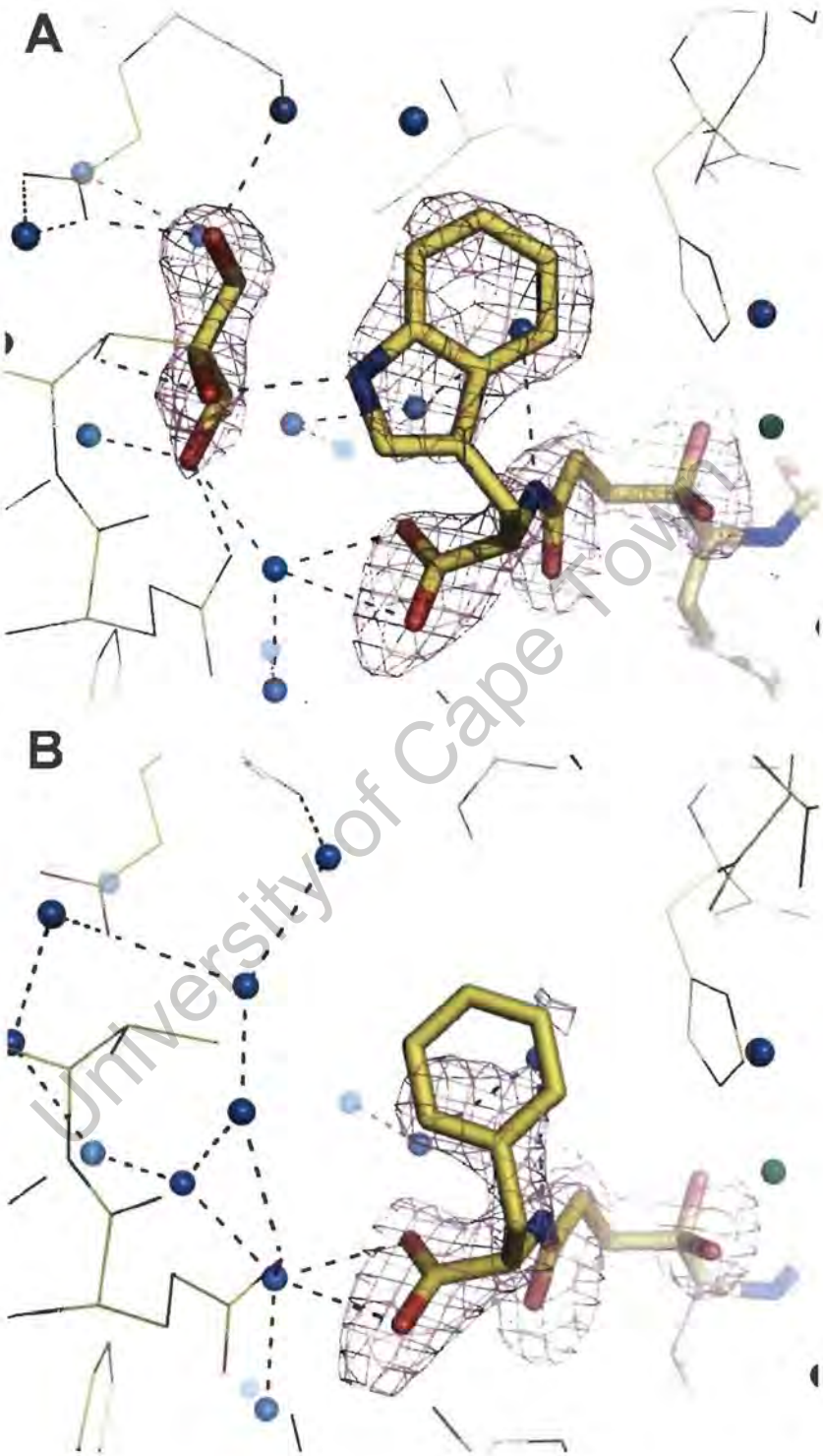


Figure 3.9 Close-up of the S_2' pocket with kAW (A) and kAF (B), showing the water network and glycerol molecule. Glycerol and inhibitors are shown as yellow sticks, water oxygens and the active site zinc ion as blue and green nonbonded spheres, respectively, and water hydrogen bonds as blue dashed lines. Protein residues are shown in line representation. As in Figures 3.6 and 3.7, the inhibitors and glycerol are shown in the first σ_A-weighted F_{obs}-F_{calc} difference density map (purple mesh) into which the inhibitor model was built, contoured at 3σ.

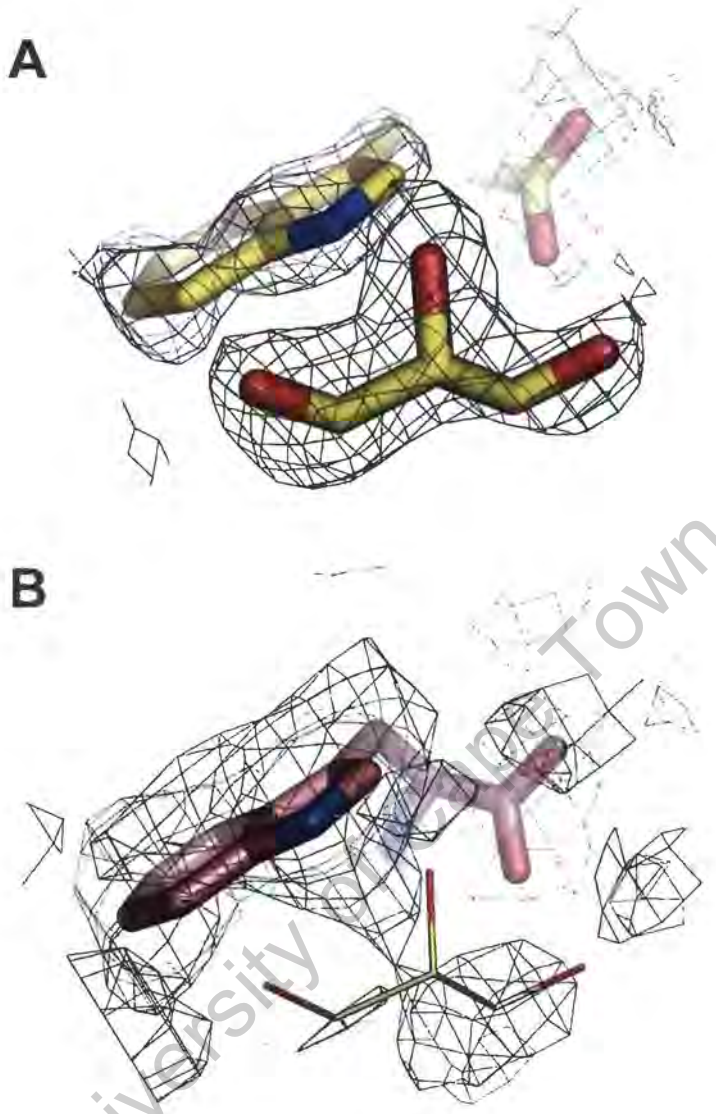


Figure 3.10 Evidence for the presence of glycerol in the active site of tACE-G13 with kAW, versus the data from the structure of tACE with RXPA380 (PDB ID 2OC2). (A) The P_2' Trp of kAW and associated glycerol molecule (yellow sticks) in the active site of tACE-G13, in density (blue mesh) calculated using the (σ_A-weighted) composite omit map protocol of *CNS* (Brünger *et al.*, 1998) with simulated annealing, and contoured at 1.0 σ. (B) No glycerol-like density was observed in an annealed σ_A-weighted composite omit map (blue mesh) calculated using the RXPA380 structure (pink sticks) with a glycerol (yellow lines) added in at the same position, and refined against the deposited structure factors.

Close proximities between nonpolar atoms are also observed between the inhibitors and 17 active site residues as well as a number of water molecules (11 and 13 for kAW and kAF, respectively), and glycerol and acetate ions in the case of kAW (Table 3.6).

Table 3.6 Nonpolar contacts and close proximities of kAW and kAF, including hydrophobic contacts, unfavourable contacts and waters.

Inhibitor Moiety	Inhibitor Atom [a]	Adjacent residue, atom [a]	N domain equivalent	Distance (Å)		Polar and *nonpolar* contacts via solvent atom
				kAW	kAF	
P_2'	CA*	TYR 523, Cε2	TYR 501	3.61	3.54	-
	CB	TYR 523, Cδ2	TYR 501	3.52	3.43	-
	CB	PHE 457, CZ	PHE 435	3.65	3.68	-
	CD1 (CD2)	GLN 281, Nε2	GLN 259	3.52	3.74	-
	CD1 (CD2)	PHE 457, Cε1	PHE 435	3.78	4.61	-
	NE1 (CE2)	THR 282, Cγ2	SER 260	4.10	4.94	-
	(CE2) *	HOH 1096, O	-	-	3.39	Q281, T282, water
	(CE2)	HOH 1083, O	-	-	3.56	E376, ***V380***, water
	CZ2	GOL 704, C3	-	3.82	-	V380
	CZ2 (CZ)	HOH 1161, O (1105)	-	3.72	3.93	E376, ***V379***, D453, GOL, water
	CH2 (CZ)	VAL 379, Cγ1	SER 357	4.05	5.17	-
	CH2 (CZ)	VAL 380, Cγ2	THR 358	4.87	4.99	-
	CH2 *	HOH 1191, O	-	3.20	-	K454, ***F527***
	Cz3 (CE1)*	HIS 383, Nδ1	HIS 361	3.65	4.20	-
	Cz3 (CE1)	ASP 415, Oδ1	ASP 393	3.64	4.23	-
	Cz3 (CE1)	ASP 415, Oδ2	ASP 393	3.47	3.92	-
	Cz3 (CD1)	PHE 527, Cε1	PHE 505	4.04	4.11	-
	CE3 (CD1)	HOH 1016, O (1008)	-	4.05	3.86	E411, D415, S526
	CE3 (CD1)	HOH 1162, O (1199)	-	2.93	3.05	***V380***, ***H383***, water
	N	TYR 523, Cε2	TYR 501	3.60	3.37	-
P_1'	CAV	HOH 1162, O (1199)	-	3.34	3.41	***V380***, ***H383***, water
	CAW	HIS 353, Nε2	HIS 331	3.61	3.77	-
	CAW	ALA 354, O	ALA 332	2.92	3.00	-
P_1	CBL	ALA 354, O	ALA 332	3.56	3.67	-
	CBL	SER 355, Cα	SER 333	4.11	4.18	-
	CAO *	PHE 512, Cε2	PHE 490	3.63	3.68	-
	CAO	SER 355, Cβ	SER 333	3.79	3.83	-
	CAO*	HOH 1020, O (1056)	-	3.57	3.88	H353, S355, K368, ***F512***, water
	CAI	SER 516, Cβ	ASN 494	5.70	5.69	-
	CAI	HOH 1169, O (1077)	-	3.53	3.66	N70, ***V351***, ***F512***, water
	CAG	ACT 706, C	-	3.91	-	N66, ***W357***, water
	CAJ *	VAL 518, Cγ2	THR 496	3.59	3.35	-
	CAP *	HOH 1165, O (1081)	-	3.66	3.68	E411, ***V518***, R522, Y523
P_2	CBE	HIS 387, Nε2	HIS 365	3.59	3.55	-
	CAR	GLU 411, Oε2	GLE 389	3.46	3.54	-
	CAR *	HIS 387, Cε1	HIS 365	3.50	3.77	-
	CAR	HOH 1165, O (1081)	-	3.21	3.20	E411, R522, Y523
	CAL	HOH 1058, O (1090)	-	3.48	3.44	P407, ***H410***, E411, R522, water
	CAL	HOH 1163, O (1086)	-	3.68	3.64	***P407***, water
	CAH *	HIS 410, Cδ2	HIS 388	3.98	3.80	-
	CAH	HOH 1228, O	-	-	3.48	***E403***, water
	CAK *	PHE 391, Cζ	TYR 369	3.35	3.45	-
	CAK	HOH 1232, O (1229)	-	3.82	3.98	Y360, water
	CAQ	ALA 356, O	ALA 334	3.32	3.50	-

[a] kAF atom names and numbers are given in parentheses when these differ from those of kAW. * More than one close proximity between the side chain moieties concerned (e.g. aromatic ring stacking); shortest interatomic distance quoted.

Some of these residues (T282, S355, V379, V380, H387, F457, F512, V518 and F527) could contribute favourably to binding entropy by the hydrophobic effect, while others (Q281,

H353, A354, E384, D415 and the water molecules) are polar or charged groups, and some (H383 and Y523) make both favourable and unfavourable contacts.

3.5. Determinants of domain-selectivity from ketone inhibitor interactions

3.5.1. kAP, kAW and kAF as tools for identifying determinants of domain-selectivity

The ketone zinc-binding moiety of kAP and its derivative inhibitors has lower affinity for the active site zinc ion than the carboxyl, sulphydryl and phosphinyl groups of other previously-studied inhibitors (reviewed by Redelinghuys *et al.* (2005)). This results in a situation similar to the binding of a natural substrate, with the affinity of the inhibitor depending to a great extent on the contributions of other groups. These inhibitors are thus suitable tools for investigating the contributions of moieties other than the zinc ion to inhibitor binding.

The chemical structures of kAW and kAF are very similar to their parent molecule, kAP, differing only at the P_2' position (Figure 3.1), however they show markedly lower affinity for the N domain, leading to enhanced C-domain-selectivity (Table 3.7). The dramatic reduction in affinity for the N domain can be explained by their modes of interaction with the active site, summarised in Tables 3.5 and 3.6. The following discussion uses the numbering of tACE when referring to the C domain. tACE residue numbers can be converted to C domain numbering by adding 576.

Table 3.7 Domain-selectivity of novel inhibitors kAW and kAF, in comparison with their parent molecule kAP (Watermeyer *et al.*, Appendix D).

Inhibitor	$K_{iC\ domain}$ (μM) [a]	$K_{iN\ domain}$ (μM) [a]	C-selectivity ($K_{iN\ domain}$ /$K_{iC\ domain}$)
kAP	0.05	1.5	30
kAW	0.68	854.2	1256
kAF	0.83	> 500 [b]	> 600

[a] K_i data were determined using the fluorogenic peptide substrate Abz-FRK(Dnp)P-OH. [b] No inhibition detected up to a concentration of 500 μM.

3.5.2. Selectivity of the parent molecule kAP – the S_1 and S_2 pockets

kAP has 30-fold higher affinity for the C domain than for the N domain (Table 3.7). This domain-selectivity does not reside in the P_2' Pro moiety, since it has been shown that Pro at this position results in strong binding to both domains (Michaud *et al.*, 1999). Moreover,

from the co-crystal structures of tACE and the N domain with lisinopril (PDB IDs 1O86 and 2C6N, respectively), one can see that a P_2' Pro stacks against Y523 and its N domain equivalent, Y501, making no interactions with residues that differ between domains. The selectivity of kAP must thus be attributed to its P_1 and P_2 moieties, which are retained in kAW and kAF.

The S_1 and S_2 pockets of the N and C domains differ in four residues that lie within 6 Å of kAW and kAF, of which S516, V518 (S_1 pocket) and F391 (S_2 pocket) make hydrophobic contact with the inhibitor P_1 and P_2 side chains. These are replaced by the charged and polar residues N494, T496 and Y369, respectively, in the N domain (Figure 3.6; Tables 3.5 and 3.6).

The substitution of F391 with Tyr in the N domain introduces a hydroxyl group into the S_2 pocket, which would result in steric hindrance of the binding of the P_2 benzoyl moiety of kAP (Figure 3.6). The importance of this residue for selectivity has been confirmed by mutation of F391 to Tyr in the C domain context, which resulted in an 8.6-fold reduction in affinity for kAP (Watermeyer *et al.*, Appendix D). This substitution has also been identified as contributing to the C-domain-selectivity of the phosphinic peptide inhibitor RXPA380, which also has a benzyl group at this position (Corradi *et al.*, 2007) and may explain the observation that a P_2 Ser residue, which in contrast to a benzyl group is small and polar, confers N-domain-selectivity (Michaud *et al.*, 1999).

Also in the S_2 pocket, E403 of the C domain is replaced by R381 in the N domain, which is 4.6 Å from the P_2 benzoyl ring when the structure of the N domain (PDB ID 2C6F) is aligned with the co-crystal structure of kAW (Figure 3.6). The bulky, positively charged side chain of R381 would thus approach the P_2 moiety, replacing the smaller, negatively charged E403, which does not interact with the inhibitor (Figure 3.6). This could account for the observation that a positively-charged P_2 Lys is unfavourable for N domain binding (Cotton *et al.*, 2002), however no effect on kAW affinity was observed upon mutation of E403 to Arg in the C domain (Watermeyer *et al.*, Appendix D), suggesting that this residue is not important for the binding of compounds with a P_2 benzyl ring.

In the S_1 pocket, the substitution of S516 with N494 in the N domain replaces a small polar side chain with a the bulky polar group, while the replacement of V518 with T496 introduces

a hydroxyl moiety in close proximity to the P_1 Phe ring of kAW and kAF (Figure 3.6). N494 of the N domain is a potential N-glycosylation sequon, although no glycan was observed at this site in the N domain crystal structure. The presence of Ser instead on Asn at this position in the N domain of rat ACE has been suggested to account for its C domain-like selectivity (Jullien *et al.*, 2006).

V518 has been confirmed by mutagenesis to be important for the C-domain-selectivity of kAP, with replacement of V518 with Thr resulting in a low affinity for kAP similar to that of the N domain, while mutation of S516 had little effect (Watermeyer *et al.*, Appendix D). The effect of the V518T substitution is not steric, however the association of the P_1 Phe side chain with T496 of the N domain is probably made unfavourable by the displacement of any water molecules hydrogen-bonded to the hydroxyl moiety, which would be necessary for binding of the bulky phenyl group. Displacement of waters around V518, in contrast, would result in a favourable increase in entropy. No waters are present at this site in the N domain crystal structures (PDB IDs 2C6N and 2C6F), however few water molecules were added to these structures due to the low resolution of the data (Corradi *et al.*, 2006). This S_1 subsite alteration may explain the moderate C-domain-selectivity of lisinopril and enalaprilat, whose P_1 benzyl groups overlap with the Phe ring of the present structures, as well as the observation that a bulky P_1 moiety confers C-domain-selectivity (Acharya *et al.*, 2003).

The C-domain-selectivity of kAP is thus due to sterically and entropically favourable hydrophobic interactions between the bulky hydrophobic P_1 and P_2 benzyl moieties and F391 and V518, which are replaced with polar groups in the N domain. Mutation of these residues to their N domain counterparts further confirmed that these substitutions contribute to the selectivity of kAW and kAF, which differ from kAP only in the P_2' position (Watermeyer *et al.*, Appendix D).

3.5.3. kAW selectivity in the S_2' pocket

The single residue replacement of Trp for the P_2' Pro of kAP reduces the affinity of kAW for the N domain dramatically, making it approximately 40-fold more C-domain-selective than kAP. This effect is the net result of a 570-fold decrease in affinity for the N domain, together with a 13-fold decrease in affinity for the C domain, relative to kAP (Table 3.7). The loss in affinity for the C domain is probably due to the loss of the favourable backbone rigidity and

hydrophobic stacking interactions of the Pro moiety of kAP, however the greater loss of affinity for the N domain can be attributed to interactions in the S_2' pocket, described below.

In the crystal structure, the P_2' Trp indole nitrogen of kAW makes a hydrogen bond to a highly ordered glycerol molecule, which is linked via numerous hydrogen bonds to the water network and to polar and charged residues T282, Q281, N277 and E376 (Figures 3.6 (A) and 3.9 (A)). Since this glycerol is derived from the cryo-protectant used during data collection, crystal growth must have occurred in its absence, suggesting that a water molecule was the original acceptor for the Trp H-bond. In addition, the Trp interacts closely with a hydrophobic patch created by T282, F457, F527 and Y523 on one side, and with H383, V379, V380 and a water molecule on the other side. Four of these ten residues, T282, E376, V379 and V380, are substituted in the N domain (Figure 3.6 (A); Table 3.6).

The substitution of S284 and D453, located deep in the S_2' pocket in the C domain, with the bulkier charged residues E262 and E431 in the N domain, may also contribute to C-domain-selectivity. The side chains of S284 and D453 have been built in different conformations in tACE crystal structures, and appear to be able to swing away from the active site to allow for the binding of bulkier side chains (see for comparison PDB IDs 1O86 and 2IUL). This mobility may not be possible for their bulkier N domain equivalents. Interestingly, the N domains of rat and mouse ACE have substrate selectivity similar to that of the human C domain, thought to result from the presence of human ACE C domain-like sequence at 494, 453 and 357 (Jullien *et al.*, 2006).

All these substitutions together amount to an N domain S_2' pocket that is considerably more polar and constrained in nature than the C domain pocket, and hence less conducive to the binding of bulky hydrophobic residues such as Trp. This property of the S_2' subsite has been predicted from the moderate C-selectivity of known inhibitors trandolaprilat, quinaprilat and perindoprilat, which also have bicyclic P_2' moieties (Perich *et al.*, 1994; Acharya *et al.*, 2003).

Mutagenesis studies have confirmed the importance of some of these substitutions for the C-domain-selectivity of kAW, with mutation of E376 and V380 to their N domain equivalents resulting in 3.4- and 3.9-fold decreases in affinity for kAW (Watermeyer *et al.*, Appendix D). Similar mutation of T282 had no effect on affinity and mutation of V379 caused an increase in affinity for kAW (Watermeyer *et al.*, Appendix D). The replacement of V380 with the

more polar T358 in the N domain removes the entropic contribution of this hydrophobic interaction to kAW binding, similar to the V518T substitution in the S_1 pocket. Since E376 interacts with the Trp indole N indirectly via the glycerol (or water network), the effect of the E376D substitution is probably a result of the weakening of this interaction due to the lengthening of the distance between the negatively charged group and the indole nitrogen. In the N domain structure (PDB ID 2C6F), this side chain is turned away from the active site. The increase in affinity for kAW observed when V379 is mutated to Ser is not readily explained by the crystal structure but may be due to a rearrangement of side chains in the S2 pocket, possibly leading to the formation of a water-mediated hydrogen bond between the Ser and the Trp indole N. This is supported by the observation that this mutant did not have increased affinity for kAF, which lacks a P_2' hydrogen-bonding moiety (Watermeyer *et al.*, Appendix D).

The overall low affinity of the N domain for kAW is thus probably due to the combined effect of the S_2' substitutions E376D and V380T, possibly with a contribution from the S284E and D453E substitutions, together with the S_1 and S_2 pocket substitutions described above.

3.5.4. kAF selectivity in the S_2' pocket

Like the Trp of kAW, the P_2' sidechain of kAF is in close proximity to V379 and V380, as well as the conserved hydrophobic residues H383, Y523, F527 and F457, but many of these residues are more distant than in the case of kAW because of the smaller size of the Phe side chain (Table 3.6). Moreover, the aromatic ring is unable to make a hydrogen bond to the water network, and in place of the glycerol moiety seen in the kAW structure, a ring of ordered waters is present, interacting with the polar residues beyond (Figure 3.9 (B)). The effect of the smaller size of this side chain, together with the lack of a hydrogen-bonding moiety, is evident in the poorly-defined density and higher atomic displacement parameters of this group in the crystal structure, which indicate a degree of disorder (Figures 3.6 (B) and 3.9 (B)).

kAF has similar affinity for the C domain to kAW, and exhibits no inhibition of the N domain at up to 500μM inhibitor, implying a *Ki* considerably greater than 500 μM, similar to what was observed for kAW (Table 3.7). Mutagenesis studies confirmed that the same P_2' interactions that contribute to the selectivity of kAW are also active in kAF binding, despite

the smaller size and disorder of the Phe group (Watermeyer *et al.*, Appendix D). From this investigation then, it can be concluded that the C-domain-selectivity of the kAP P_2' derivatives kAW and kAF results from interactions between their bulky hydrophobic side chain moieties and the C domain-specific residues F391 (S_2), V518 (S_1), E376 (S_2') and V380 (S_2'). Interactions with their N domain equivalents are less favourable.

3.6. Co-crystal structures of tACE-G13 with a lisinopril-derivative inhibitor

3.6.1. Structure solution, model building and refinement

Two enantiomers of lisW were derived from lisinopril by P_2' substitution (Figure 3.1), one of which, lisW-S, was shown to be strongly C-domain-selective (Redelinghuys *et al.*, 2006). X-ray diffraction datasets were collected for crystals grown from solutions containing both enantiomers of lisW (see Table 3.2), and phases were determined by molecular replacement using protein atoms from the structure of tACE-G13 (PDB ID 2IUL) as a model. As was the case for the kAW and kAP datasets, the molecular replacement solutions had high scores and yielded initial phases that required little further refinement.

During the model building process, it became clear that both datasets contained the same enantiomer, lisW-S, in the active site. Attempts to build the R enantiomer into the active site produced negative difference density near the chiral carbon, and failed to position the P_1 Phe side chain properly in the density (Figure 3.11). This suggests that the purified lisW-R used for co-crystallisation contained traces of lisW-S, which were selected by the tACE active site. Models containing the S enantiomer were thus built for both datasets, to 2.3 and 2.4 Å, and refined to final crystallographic *R*-factors of 21.6% and 22.6% (Table 3.8). The models are referred to as lisW-A and lisW-B, according to the dataset used.

As was the case for the kAW and kAP structures, the *R*-factors are good for structures of this resolution that contain flexible regions. The method used for test set selection was the same as for the ketone inhibitor structures, so that the reflections used were the same set that were used for the molecular replacement model, amounting to 4.4% of reflections for both datasets.

The B-factors for the ligand in the lisW-A model were higher than those in the lisW-B model, suggesting that the lisW-B model is more reliable. This was probably due to the quality of the

ligand density in the active site, which was better for the lisW-B dataset. The lisW-B dataset is of slightly higher resolution (2.30 Å) than the lisW-A dataset (2.40 Å), however the redundancy in the outer shell of the lisW-B dataset is slightly lower than that of the lisW-A dataset (Table 3.2), which suggests that resolution is not the reason for the difference in quality of the maps in the active site. The poor quality of the ligand density is thus likely due to lower occupancy of the ligand in the lisW-A dataset, so that the lisW-B model should be regarded as more reliable than the lisW-A model.

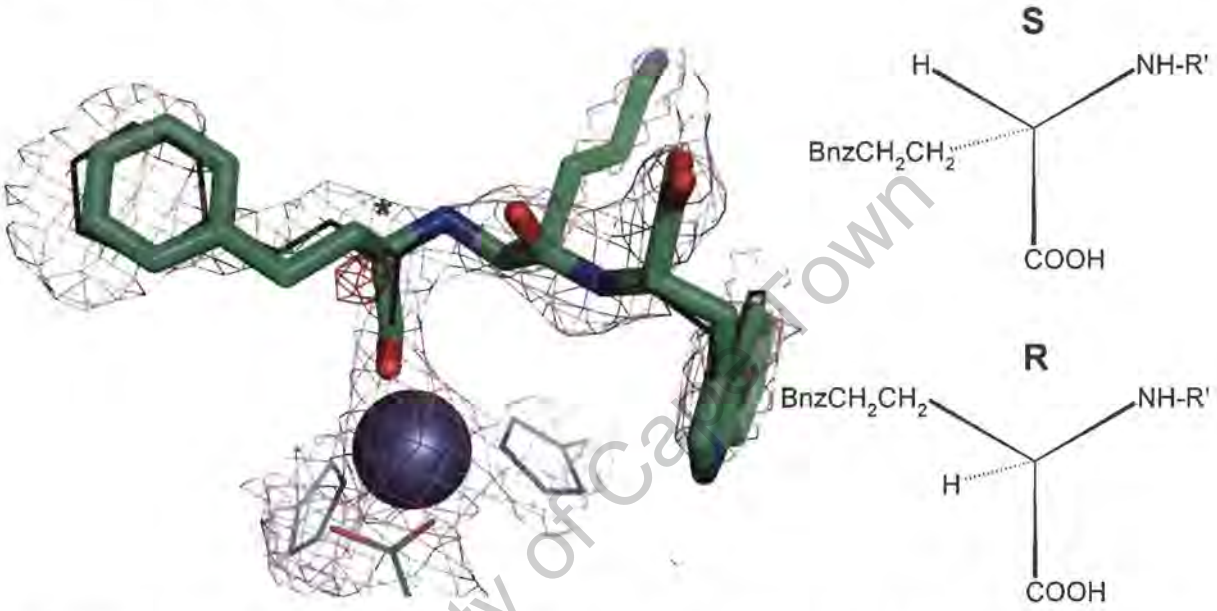


Figure 3.11 Left panel: The lisW-R enantiomer (green sticks) is shown in σ_A-weighted $2F_{obs}$-F_{calc} (purple mesh, contoured at 1.0σ) and negative F_{obs}-F_{calc} density (red mesh, contoured at -3.0σ), calculated after attempting to build this enantiomer into the active site using the lisW-B dataset. lisW-S is shown in black lines for comparison, and the chiral centre is indicated with an asterisk. The active zinc ion (grey sphere) and zinc-binding residues (green lines) are also shown in σ_A-weighted $2F_{obs}$-F_{calc} density. Right panel: The arrangement of groups about the indicated chiral centre for the S and R enantiomers. R′ = the remainder of the molecule (P_1′ and P_2′ groups).

The progress of model building was monitored by the changes in R_{cryst} and R_{free} after each cycle (Figure 3.12). Increases in R-factors towards the end of the process are the result of the deletion of disordered surface loop residues, which had produced a misleading lowering of R_{free}, as well as the re-refinement of individual B-factors. Structure refinement was considered complete when R_{free} ceased to decrease, and when any errors of stereochemistry introduced during the model building process had been addressed. Ramachandran plots for the lisW-A and lisW-B models are shown in Figure 3.13, where it can be seen that all of the residues built fell within the allowed and favoured regions Table 3.8).

Table 3.8: Refinement statistics for two models of lisW-S in the active site of tACE-G13, built using datasets lisW-A and lisW-B.

	lisW-A	**lisW-B**
No. reflections, working set [%]	21638 [88.7]	23814 [81.3]
No. reflections, test set [%]	1228 [4.4]	1294 [4.4]
R_{cryst} (%) [a,b]	21.6 (33.2)	22.6 (36.8)
R_{free} (%) [a,c]	25.4 (29.4)	26.7 (33.3)
Luzzati error (Å) [test set]	0.31 [0.33]	0.30 [0.35]
SigmaA error (Å) [test set]	0.35 [0.34]	0.33 [0.33]
R.m.s. deviations from ideality		
bonds (Å)	0.01	0.01
angles (°)	1.4	1.4
dihedrals (°)	21.4	21.6
improper angles (°)	1.1	1.1
Mean B-factor (Å^2) [min-max]		
all atoms	27.7 [4.2-76.8]	27.7 [10.7-58.8]
ligand atoms	39.4 [31.4-49.8]	26.0 [20.6-32.0]
B-factor r.m.s. deviation (Å^2)		
bonded mainchain atoms	1.29	0.95
bonded sidechain atoms	1.66	1.49
angle mainchain atoms	2.12	1.51
angle sidechain atoms	2.60	2.16
Number of atoms		
solvent atoms	200	181
protein atoms [d]	4729	4689
ions	3	3
inhibitor atoms	36	36
glycan atoms	38	38
% unit cell volume occupied	51.82	53.15
Ramachandran plot, % residues in favoured regions	97.56	98.40

[a] Values in parentheses represent the highest resolution shell: 2.40-2.49 Å for lisW-A and 2.30-2.38 Å for lisW-B; [b] $R_{cryst} = \sum_h |F_o - F_c| / \sum_h F_o$, where F_o and F_c are the observed and calculated structure factor amplitudes of reflection h of the working set, respectively; [c] R_{free} is equal to R_{cryst} for h belonging to the test set of reflections; [d] excluding duplicated atoms in alternate conformations.

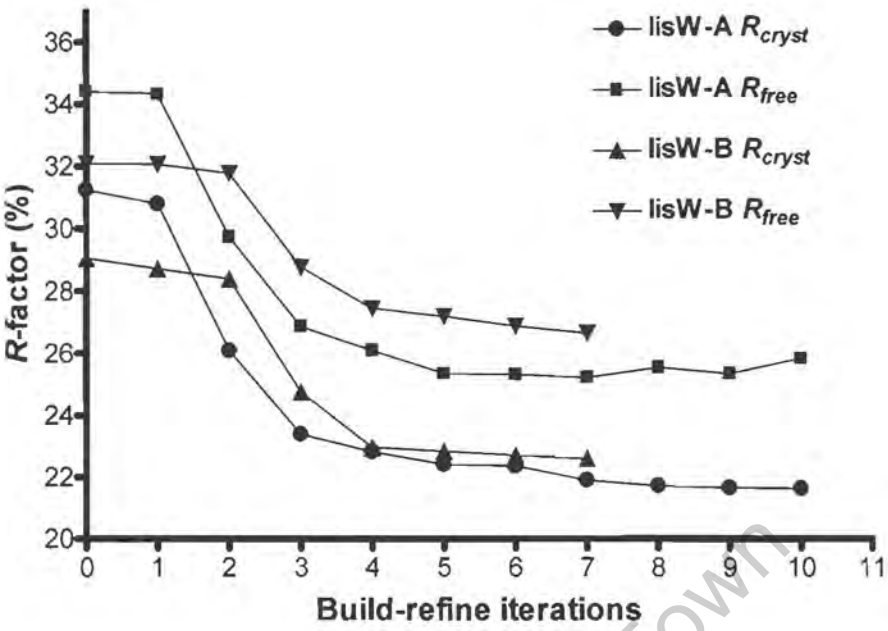


Figure 3.12 Progress of model building and refinement for the lisW-A and lisW-B datasets, as indicated by R_{cryst} and R_{free}. $R_{cryst} = \sum_h |F_o - F_c| / \sum_h F_o$, where F_o and F_c are the observed and calculated structure factor amplitudes of reflection h of the working set, respectively; R_{free} is equal to R_{cryst} for h belonging to the test set of reflections.

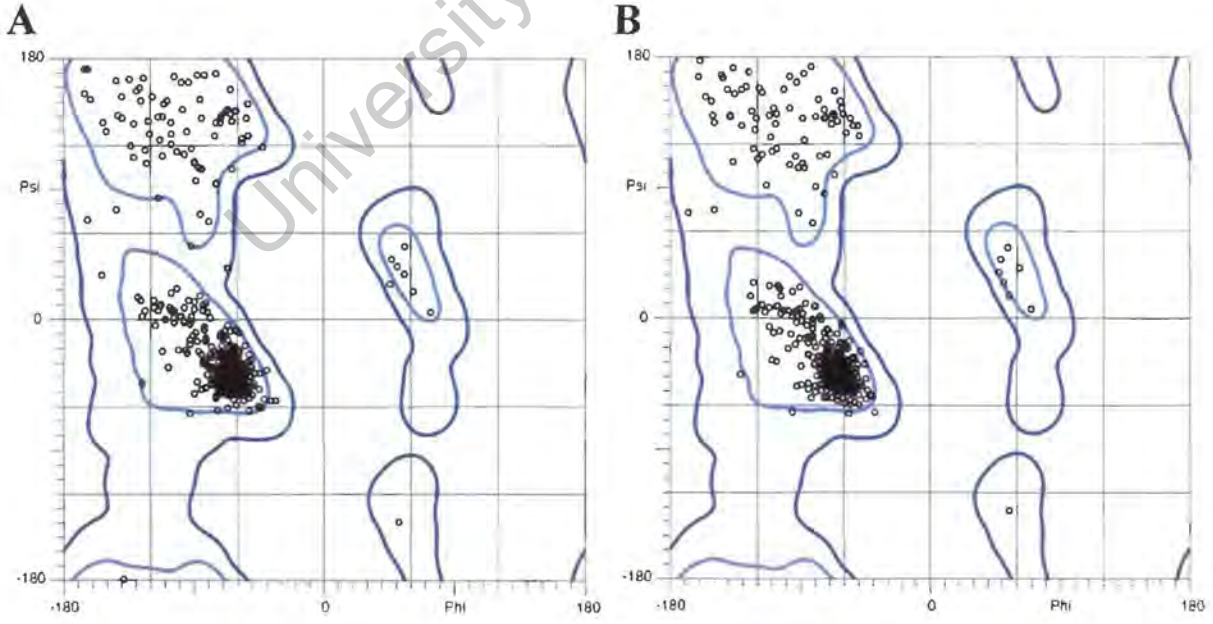


Figure 3.13 Ramachandran plots for the co-crystal structures of lisW-A (A) and lisW-B (B) in the active site of tACE-G13. Favoured regions are enclosed by the light blue contours and additionally allowed regions by the dark blue contours. Plots were generated using Molprobity (Davis *et al.*, 2004).

3.6.2. Structure overview

As was the case for the ketone inhibitors, the protein component is in an identical conformation to the unliganded structure (PDB ID 2IUL), and there is very little variation between the two structures (Table 3.9). Fewer glycan and protein residues were built than for the higher resolution kAW and kAF structures, however the zinc and chloride ions were clearly visible (Table 3.8; Figure 3.14).

Table 3.9. Root mean square deviations in Å between the two co-crystal structures of tACE-G13 with lisW (A and B), unliganded tACE-G13 (PDB ID 2IUL) and minimally glycosylated wild-type tACE bound to lisinopril (tACE-LPR; PDB ID 1O86).

Structure 1	Structure 2	overall[a]	side[b]	main[c]	zinc[d]	CL1[e]	CL2[f]	active site[g]
A	B	0.52	0.69	0.21	0.15	0.53	0.15	0.23
A	tACE-G13	0.48	0.65	0.19	0.04	0.48	0.12	0.15
B	tACE-G13	0.36	0.49	0.13	0.14	0.29	0.10	0.16
A	tACE-LPR	0.62	0.83	0.27	0.12	0.56	0.15	0.22
B	tACE-LPR	0.56	0.74	0.25	0.13	0.33	0.16	0.17
tACE-G13	tACE-LPR	0.54	0.70	0.27	0.14	0.40	0.18	0.22

[a]all atoms; [b]side chain atoms; [c] main chain atoms; [d] zinc-binding residues; [e] residues within 5 Å of CL1 (W182, R186, D507, P508, R489, W485, W486); [f] residues within 5 Å of CL2 (M223, Y224, P407, P519, Y520, I521, R522); [g] residues within 6 Å of lisW-S in the active site (N70, E143, E162, Q281, H353-A356, Q369, E376-D377, V379-V380, H383-E384, H387, E411, D415, K454, F457, F460, K511-H513, S516, V518, Y520, R522-Y523, S526-F527, Q530)

The r.m.s. deviation between the ligand atoms in the two structures, calculated in the context of an all-atom alignment, is 0.55 Å. The r.m.s. deviation calculated using ligand atoms along is 0.46 Å. Deviation between the ligand structures is probably largely due to the poor quality of the density for the ligand in the lisW-A dataset (Figure 3.15 (A)), also evident in high ligand B-factors (Table 3.8). These data may be indicative of partial occupancy of the ligand in this dataset, however the resolution limit of the data did not allow refinement of occupancies. Due to the high B-factors of the ligand in the lisW-A model, the lisW-B model will be used to draw conclusions about lisW-S binding.

As was seen in the ketone inhibitor structures, and indeed in all ACE structures solved to date (Watermeyer *et al.*, 2006), surface loops 102-112, 296-299 and 434-439 showed a degree of structural disorder, and residues 436-438 were omitted from the lisW-S model. High B-factors and weak density indicative of disorder were also seen in the surface loop 154-157.

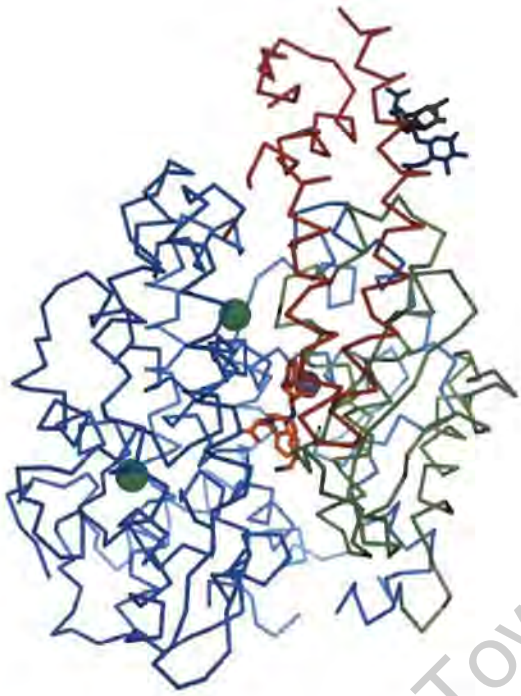


Figure 3.14 Cartoon representation showing an overview of the structure of lisW-S in the active site of tACE-G13, determined from the lisW-B dataset. The protein backbone is shown in ribbon representation, with subdomains 1, 2 and 3, as described in chapter 1, coloured red, blue and light green, respectively. The zinc and chloride ions are shown as grey and green spheres, respectively, glycan residues are shown as dark blue sticks, and lisW-S as orange sticks.

3.6.3. Variation in the orientation of the P_2' Trp between inhibitors

Interestingly, the P_2' Trp of lisW-S takes a different conformation to that observed in kAW and RXPA380, being flipped by approximately 70° and 180° about χ_1 and χ_2 respectively, relative to kAW (Figure 3.16). Despite the relatively low resolution of the lisW-B dataset, the change in orientation was demonstrated unambiguously by the electron density (Figure 3.15 (B)). In this orientation the Trp ring makes a new hydrogen bond with D415 (Figures 3.15 and 3.16).

This conformation is probably favoured in lisW-S due to hydrophobic contacts between the P_2' Trp ring, the P_1' Lys sidechain and V380, which is sandwiched between the two (Figure 3.17). kAW and RXPA380 lack a P_1' Lys, so that the conformation observed here is probably less stable in these structures. The potential for a P_2' Trp moiety to take two different conformations highlights the large volume of the S_2' pocket, which allows this side chain to sample different orientations during binding.

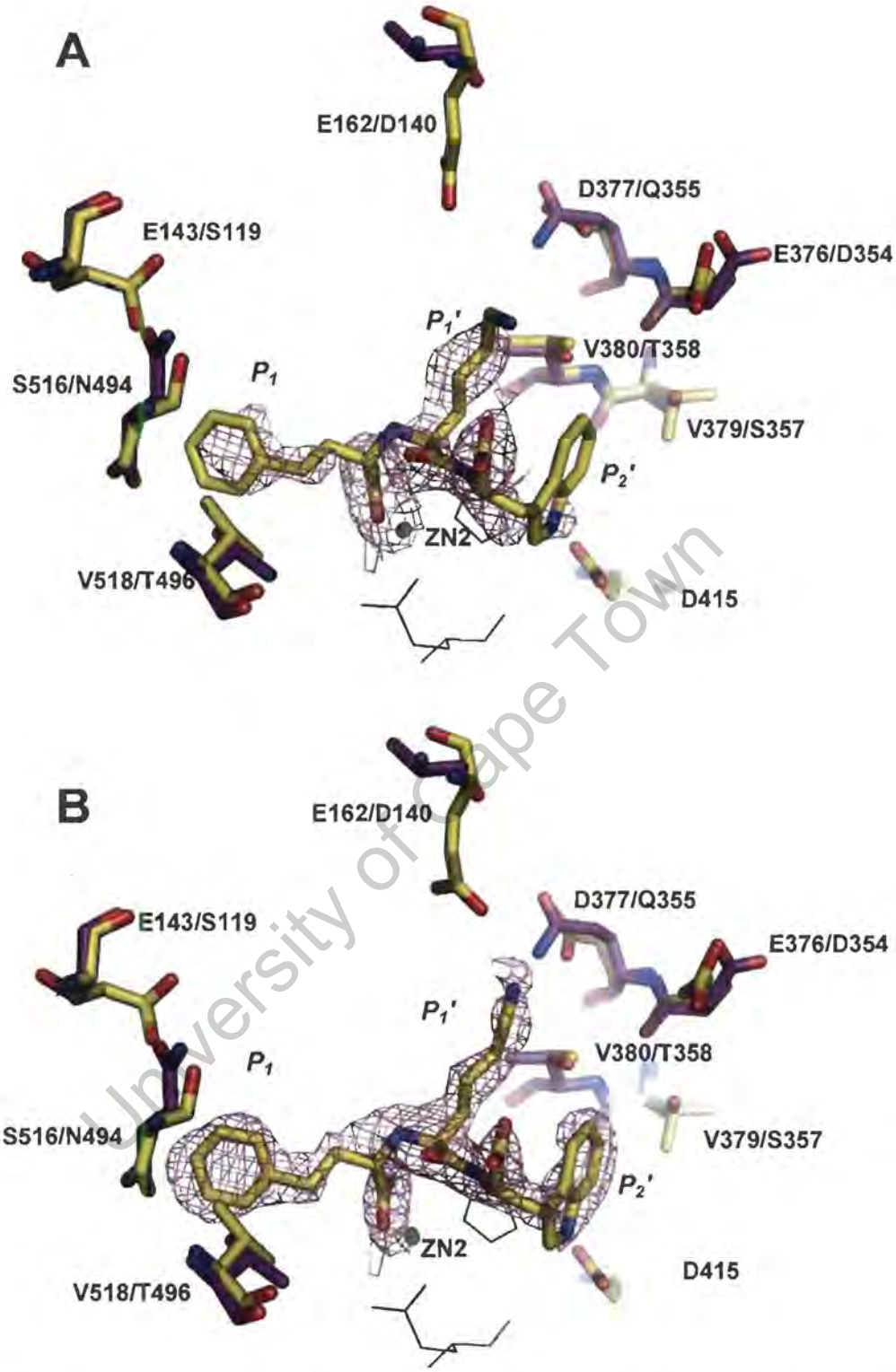


Figure 3.15 Stick representation of the active sites of models lisW-A (A) and lisW-B (B), showing lisW-S in the first σ_A-weighted F_{obs}-F_{calc} difference density map (purple mesh) into which the inhibitor model was built, contoured at 2σ. The active site zinc ion (ZN2) is shown as a green non-bonded sphere, and zinc-binding residues are shown in black lines in the background. Inhibitor-contacting residues that differ between domains are shown in yellow (tACE residues from the co-crystal structure) and purple (N domain residues from the aligned N domain structure, PDB ID 2C6F) sticks. D415, which forms a hydrogen bond with the P_2' Trp ring N, is also shown in yellow sticks. Residue labels are given as tACE/N domain, and the P_1, P_1' and P_2' moieties are labelled.

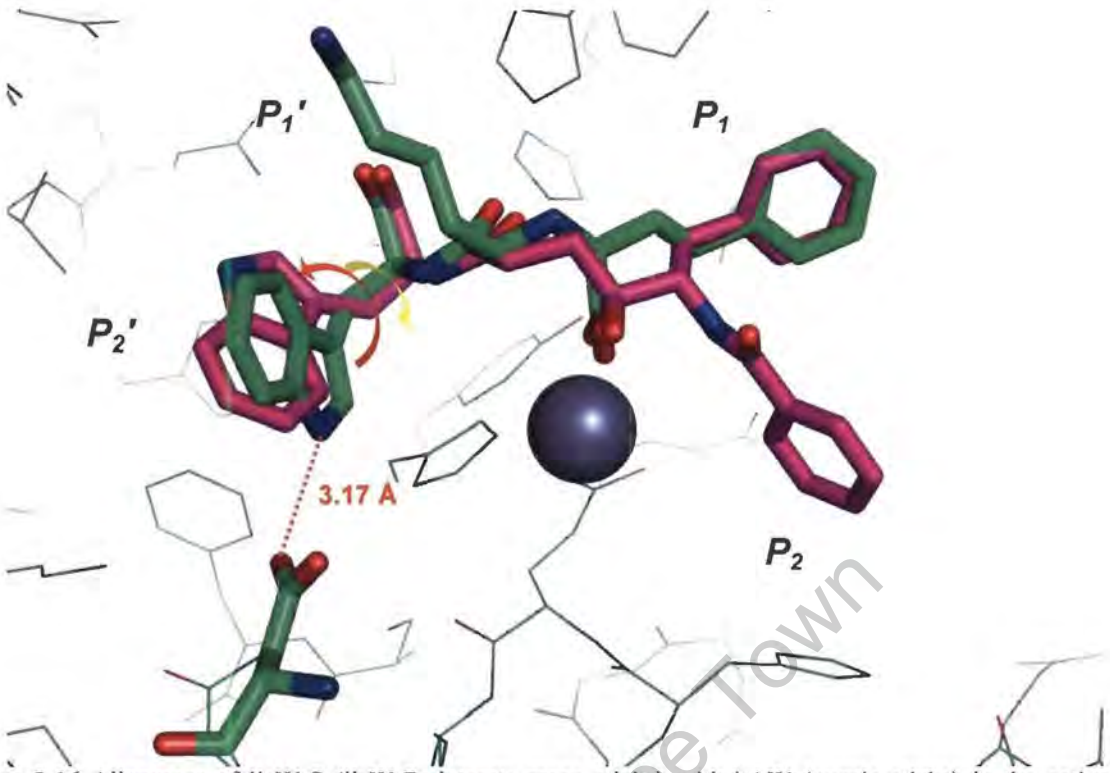


Figure 3.16 Alignment of lisW-S (lisW-B dataset; green sticks) with kAW (purple sticks) in the active site, showing differences in the orientation of the P_2' Trp moiety. The structures were aligned by pair-wise alignment of the alpha carbons of the zinc-binding residues. The P_2' Trp of lisW-S is rotated about χ_1 (yellow arrow) and χ_2 (orange arrow), relative to kAW. D415 is shown in green stick representation, and the hydrogen bond with the lisW-S Trp ring is marked as an orange dotted line. The active zinc ion is shown as a grey sphere, and the surrounding residues are shown in green line representation.

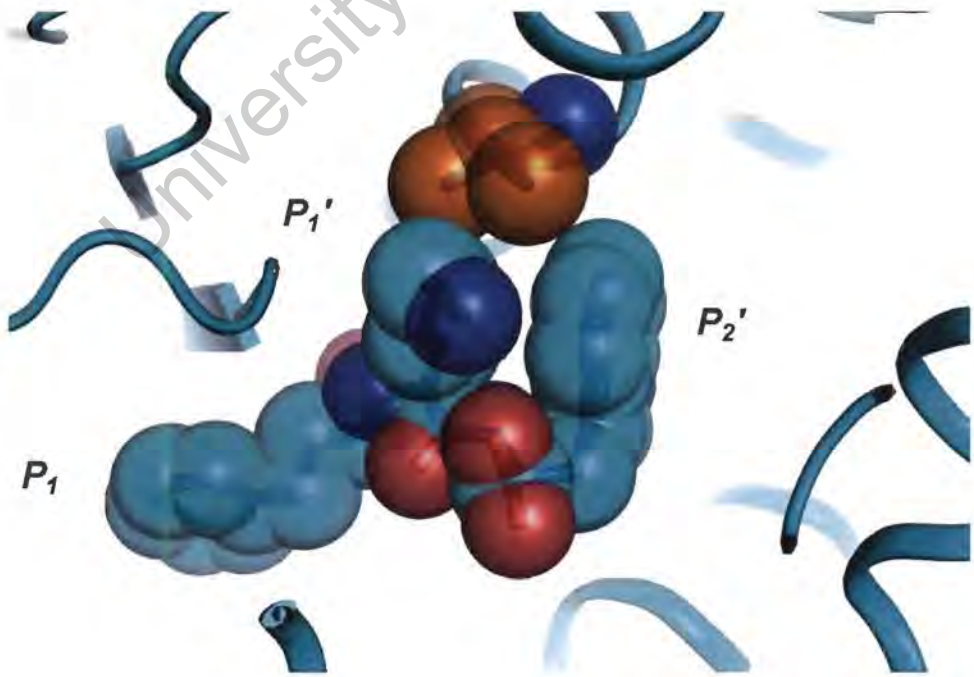


Figure 3.17 Hydrophobic stabilisation of a new conformation of the P_2' Trp of lisW-S (light blue sticks and transparent spheres) in the active site of tACE-G13 (lisW-B dataset, light blue cartoon representation). The side chain of V380 (orange sticks and transparent spheres) is intercalated between the P_1' Lys and P_2' Trp moieties.

3.6.4. Active site interactions of lisW-S

lisW-S makes 11 polar contacts (potential hydrogen bonds and ionic interactions) with 11 protein side chains (Table 3.10). In comparison with kAW and kAF, four additional contacts are contributed by the P_1' Lys side chain, while P_1 and P_2 polar contacts are absent, due to the absence of a polar moiety in the P_1 position, and to the lack of a P_2 sidechain. Hydrogen bonds are also made with four water molecules, two of which are conserved in the kAW and kAF structures, and two of which are specific to this inhibitor. One of the lisW-S-specific waters is hydrogen-bonded to the P_1' Lys amine, and the other occupies the position of the P_2 carbonyl oxygen of kAW and kAF. The remaining two conserved waters in the kAW and kAF structures are excluded from this structure by the hydrophobic contact between the P_1' Trp and P_2' Lys side chains. No glycerol or acetate molecules were observed in the active site.

Table 3.10 Polar contacts and close proximities between lisW-S (lisW-B dataset) and protein or solvent atoms.

Inhibitor moiety	Inhibitor atom	Interacting residue, atom	N domain equivalent	Distance (Å) [a]		Polar and *nonpolar* interactions via solvent atom
				lisW-S	LPR [b]	
P_2'	OXT	LYS 511, Nζ	LYS 489	2.82	2.93	-
	OXT	GLN 281, Nε2	GLN 259	3.02	3.22	-
	OXT	TYR 520, OH	TYR 498	2.70	2.56	-
	OXT	HOH 1177, O	-	3.19	(3.58)	***W279***, Q281, K511, water
	O	HOH 1177, O	-	2.43	2.68	***W279***, Q281, K511, water
	NE1	ASP 415, Oδ2	ASP 393	3.17	-	-
P_1'	OAA	HIS 513, Nε2	HIS 491	2.76	3.11	-
	OAA	HIS 353, Nε2	HIS 331	2.75	2.76	-
	NAA	GLU 162, Oε2	ASP 140	(3.80)	(3.45)	-
	NAA	ASP 377, Oδ1	GLN 355	(3.79)	(4.23)	-
	NAA	HOH 1159, O	-	3.07	(3.63)	E162, water
	NAX	ALA 354, O	ALA 332	2.90	2.92	-
Carboxylate	OAE	ZN2	ZN2	2.22	2.14	-
	OAE	TYR 523, OH	TYR 501	2.48	2.77	-
	OAE	HOH 1041, O	-	3.19	3.39	E411, ***V518***, R522
	OAF	ZN2	ZN2	2.38	2.61	-
	OAF	GLU 384, Oε2	GLU 362	2.81	2.70	-
	OAF	HOH 1109, O	-	3.11	2.82	A356, ***S355***, water

[a] Values in parentheses are greater than the maximum hydrogen-bonding distance (3.4 Å). [b] Distances between equivalent atoms in the structure of tACE with lisinopril (PDB ID 1O86).

Most of these polar interactions are conserved in the structure of tACE with lisinopril (PDB ID 1O86), which binds in a nearly identical conformation to lisW-S (Figure 3.18). There are some differences in the orientation of the amine group of the P_1' Lys moiety, which might be due to the influence of the P_2' Trp moiety, and in the zinc-co-ordination distances, however these could equally be attributed to the lower resolution of the lisW-B dataset in comparison with that used to solve the lisinopril structure.

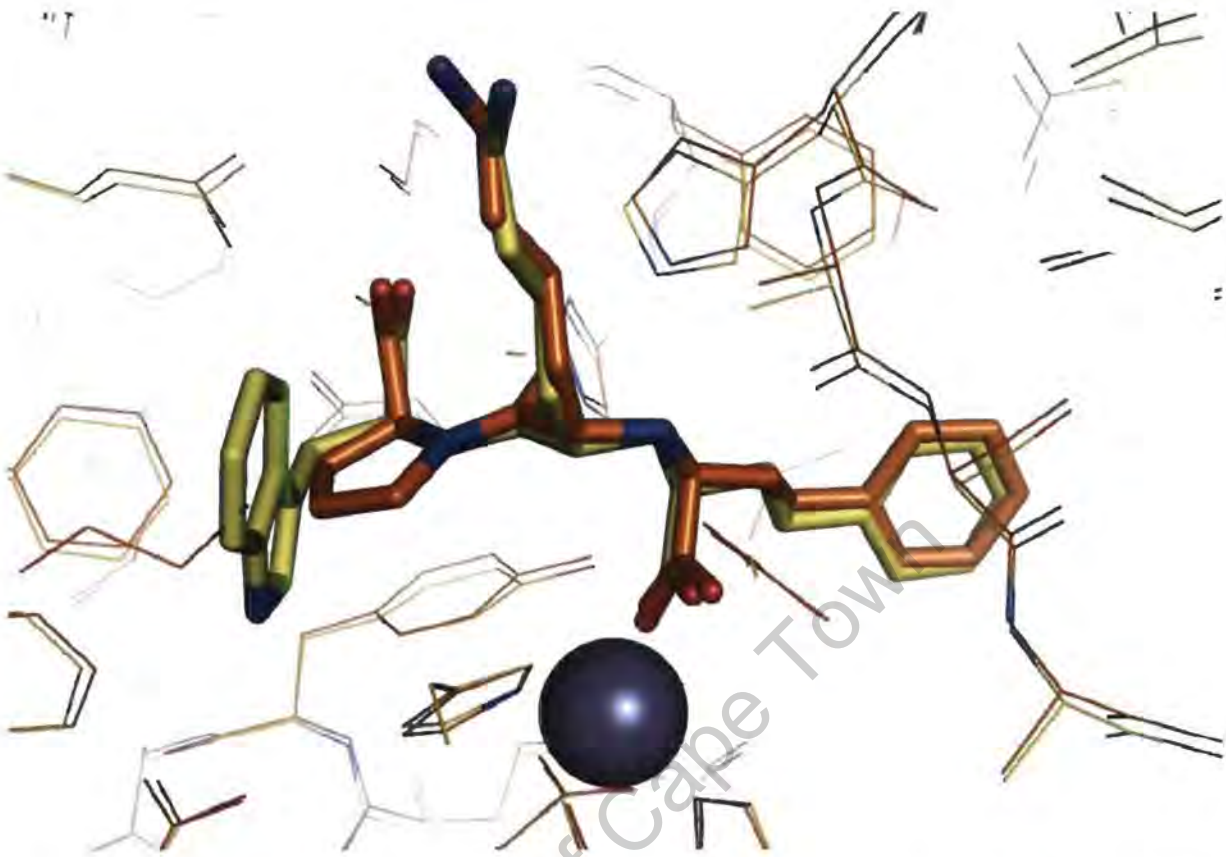


Figure 3.18 Alignment of the co-crystal structures of tACE with lisW-S (yellow sticks) and parent molecule lisinopril (PDB ID 1O86; orange sticks) in the active site. Structures were aligned by pair-wise alignment of the alpha carbons of the zinc-binding residues. The zinc ion is shown as a grey sphere, and adjacent protein components of both structures are shown in lines of the same colour as the inhibitor to indicate the degree of alignment.

Close proximities are observed between the nonpolar atoms of lisW-S and 14 protein side chains (Table 3.11). Most of these contacts (H353, A354, S355, V379, V380, H383, F457, F512, S516, V518, Y523, F527) are probably entropically favourable, although there are some close contacts with charged moieties (H383, D415 and E143). In addition to these, nine water molecules are in close proximity to nonpolar inhibitor moieties in the S_1 and S_2' pockets. Contacts in the S_1' and S_1 pockets are conserved in the structure of tACE with lisinopril (PDB ID 1O86).

Table 3.11 Distances between nonpolar atoms of lisW-S (lisW-B dataset) and other atoms in close proximity, including hydrophobic contacts, unfavourable contacts and waters. Equivalent distances in the co-crystal structure of tACE with lisinopril (PDB ID 1O86) are given for comparison.

Inhibitor Moiety	Inhibitor Atom	Adjacent residue, atom	N domain equivalent	Distance (Å)		Polar and *nonpolar* interactions via solvent atom
				lisW-S	LPR [a]	
P_2'	CA*	TYR 523, Cε2	TYR 501	3.76	3.74	-
	CB *	PHE 457, Cζ	PHE 435	3.72	3.72	-
	CD1*	PHE 527, Cε1	PHE 505	3.64	-	-
	CD1	TYR 523, Cδ2	TYR 501	3.57	-	-
	NE1*	HIS 383, Nδ1	HIS 361	3.45	-	-
	CZ2	ASP 415, Oδ1	ASP 393	3.86	-	-
	CH2*	VAL 379, Cγ1	SER 357	4.11	-	-
	CH2	GLU 376, O	ASP 354	5.21	-	-
	CZ3*	VAL 380, Cγ2	THR 358	3.36	-	-
	CH2*	HOH 1079, O	-	3.82	-	***E376***, ***V379***, water
	CZ3	HOH 1043, O	-	3.93	-	E376, water
	CE3*	lisW, CAU	-	4.33	-	-
	CE3	HOH 1121, O	-	3.93	-	Q281, T282, water
	N	TYR 523, Cε2	TYR 501	3.74	3.81	-
P_1'	CAV	HIS 383, Cδ2	HIS 361	4.30	4.21	-
	CAU*	lisW, Cε3	-	4.33	-	-
	CAS*	ALA 354, Cβ	ALA 332	4.04	4.23	-
	CAS*	HIS 353, Nε2	HIS 331	3.41	3.55	-
	CAR*	VAL 380, Cγ1	THR 358	3.41	3.99	-
	CAQ	HOH 1034, O	-	4.44	3.91	***L161***, E162, ***T166***, ***W279***, water
	CAQ	HOH 1177, O	-	4.48	3.97	lisW O, K511, ***W279***, Q281, water
P_1	CBL	ALA 354, O	ALA 332	3.20	3.43	-
	CBL	HIS 353, Cδ2	HIS 331	3.65	3.94	-
	CBL	HOH 1109, O	-	3.75	3.69	A356, E384, lisW OAF, water
	CAX*	VAL 518, Cγ1	THR 496	4.10	4.09	-
	CAX	HOH 1041, O	-	3.50	3.66	E411, ***V518***, R522, Y523
	CAP	HOH 1174, O	-	4.09	4.53	A356, ***H387***, ***F391***, water
	CAJ *	VAL 518, Cγ2	THR 496	3.89	3.97	-
	CAI	SER 516, Cβ	ASN 494	5.64	5.71	-
	CAI	GLU 143, Oε1	SER 199	5.60	5.81	
	CAI	HOH 1074, O	-	3.53	3.79	N70, ***V351***, H353, S355, K368, ***F512***
	CAO*	PHE 512, Cε2	PHE 490	3.69	3.73	-
	CAO	SER 355, Cβ	SER 333	4.00	3.98	-

[a] Distances between equivalent atoms in the structure of tACE with lisinopril (PDB ID 1O86). * More than one close proximity between the side chain moieties concerned (e.g. aromatic ring stacking); shortest interatomic distance quoted.

3.7. Determinants of lisW-S domain-selectivity

3.7.1. C-domain-selectivity of lisW-S and selection against lisW-R

The incorporation of the P_2' Trp into lisW-S reduces the affinity for the N domain dramatically relative to the parent compound, lisinopril, with neither enantiomer inhibiting the N domain at concentrations of up to 500 μM (Table 3.12). This reduction in affinity is similar to that observed with kAW and kAF, relative to kAP (Table 3.7). Both enantiomers of lisW were shown to be C-domain-selective (Nchinda *et al.*, 2006a), although the affinity of lisW-S

for the C domain is approximately 140-fold lower than that of lisinopril. This emphasises again the favourable contribution of a P_2' prolyl moiety to high affinity binding.

Table 3.12 Domain-selectivity of two enantiomers of novel inhibitor lisW, in comparison with the parent molecule lisinopril (Nchinda *et al.*, 2006a).

Inhibitor	$K_{iC\ domain}$ (μM) [a]	$K_{iN\ domain}$ (μM) [a]	C-selectivity ($K_{iN\ domain}$ /$K_{iC\ domain}$)
lisinopril	0.051	0.13	2.6
lisW-S	7.0	> 500 [b]	>100
lisW-R	26.3	> 500 [b]	>25

[a] K_i data were determined using the substrate z-Phe-His-Leu. [b] No inhibition detected up to a concentration of 500 μM.

The present crystallographic data suggest that the active site of tACE selects strongly for the S enantiomer, which showed the greater affinity for the C domain, and consequently the higher selectivity (Table 3.12). Selection against the R enantiomer is probably due to steric interference between the beta carbon of the P_1 side chain and the hydroxyl moiety of Y523, which were only 3.2 Å apart when the R enantiomer was built into the active site. Y523 makes a hydrogen bond with one of the zinc-binding inhibitor oxygens, and is conserved in both domains. This observation is in keeping with the activity of ACE on biological peptides, since the P_1 pseudo-Phe of lisW-S has the chirality of an L-amino acid, while that of lisW-R has D-amino acid chirality. The parent compound, lisinopril, shares this chirality.

3.7.2. Interactions accounting for the domain-selectivity of lisW-S

lisW-S makes contact with nine residues that differ between the N and C domains (Figure 3.15; Tables 3.10 and 3.11). Two of these, E162 and D377 of the C domain, are close enough to the P_1' Lys sidechain amine nitrogen to make ionic interactions or weak hydrogen bonds (Table 3.10). E162 is replaced by the shorter D140 in the N domain, and D377 by the uncharged Q355, both substitutions which have previously been suggested to play a role in the domain-selectivity of lisinopril (Natesh *et al.*, 2004).

In the S_2' and S_1 pockets, lisW-S contacts the same inter-domain substitutions that contribute to the C-domain-selectivity of kAW, namely E376D, V379S and V380T in the S_2' pocket, and S516N and V518T in the S_1 pocket (Figure 3.15). There is no steric clash with Y396 of the N domain since the P_2 position is not occupied. In addition to these contacts seen in kAW,

E413, which is replaced by S119 in the N domain, comes within distant hydrophobic interaction distance of the P_1 Phe ring (Figure 3.15; Table 3.11). However since this residue is charged it is unlikely to contribute favourably to C-domain-selectivity.

Because of the conformational change seen in the P_2' moiety, the importance of the S_2' interactions is likely to be different in comparison with kAW. For instance, V380 can be expected to play an important role in the affinity of this inhibitor since it stacks together with the P_2' Trp and P_1' Lys side chains. Equally, E376 is not likely to contribute favourably to binding to the C domain, since the Trp hydrogen bond is made directly to D415 and not to the water network as seen in kAW.

The C-domain-selectivity of lisW-S is thus probably due to favourable interactions with E162 (P_1'), D377 (P_1'), E376 (P_2'), V379 (P_2'), V380 (P_2'), S516 (P_1) and V518 (P_1), which are substituted in the N domain.

3.8. Water molecules in the active site of tACE

It has been suggested that replacement of conserved active site waters could be used as a strategy for ACE inhibitor design (Corradi *et al.*, 2007). Many of the water molecules contacting the inhibitors in these structures are conserved in other tACE co-crystal structures. Notably, the positions of waters 1006, 1016, 1058 and 1163 of the kAW structure (Tables 3.5 and 3.6) are conserved in all of the published tACE co-crystal structures (PDB IDs 1O86, 1UZE, 1UZF and 2OC2) as well as the kAF and lisW-S structures. In addition, waters 1020, 1161, 1165 and 1169 of the kAW structure are conserved in all except the RXPA380 co-crystal structure.

Water 1165 of the kAW structure is particularly interesting since it is within hydrogen bonding distance of one of the zinc-binding inhibitor oxygens, the P_1 backbone amide (of kAW and kAF), the zinc ligand E411, Y523 which participates in stabilisation of the gem-diolate, and R522 which is bound to the chloride ion CL2 (Figure 3.19). This water molecule thus seems to be important for substrate binding, and provides a possible explanation for the chloride dependence of the C domain (Wei *et al.*, 1991), which has been shown to be linked to R522 (R1098 of the full-length enzyme; Liu *et al.*, 2001). The interaction of R522 with

CL2 may serve to position this residue for optimal interaction with the substrate via water 1165.

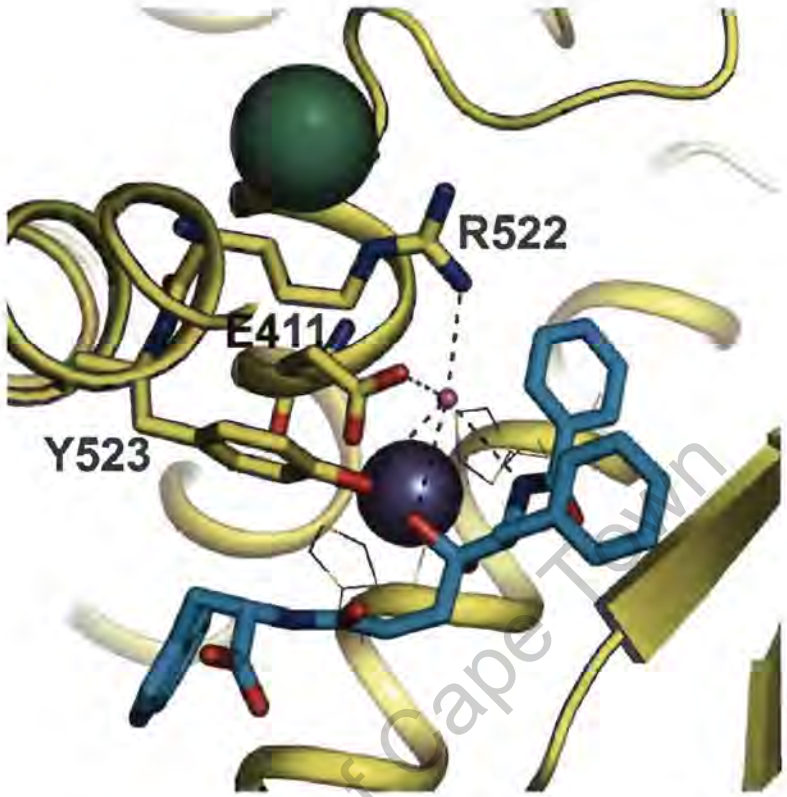


Figure 3.19 Cartoon representation of conserved water molecule 1165 in the active site of the co-crystal structure of tACE-G13 with kAW. The position of the oxygen atom of water 1165 is shown as a light purple non-bonded sphere, and residues within hydrogen-bonding distance as yellow sticks. kAW is shown as cyan sticks, and potential hydrogen bonds to water 1165 are shown as dark blue dashed lines. The protein backbone is shown in yellow cartoon representation, the active site zinc ion as a grey sphere, chloride ion CL2 as a green sphere, and the zinc-binding residues H383 and H387 as yellow lines.

The design of inhibitor compounds containing hydrogen bonding moieties that would occupy C-domain-specific water binding sites would enhance domain-selective binding. However comparison cannot be made with water sites in the N domain until an N domain structure is solved at high enough resolution to determine accurate water positions.

3.9. Comments on affinity and suggestions for further inhibitor design

3.9.1. The effect of the zinc-binding moiety

It should be noted that the binding affinities of the inhibitors studied here are not of the same order as those of currently available drugs such as lisinopril which have *Ki* values in the low

nanomolar range (Acharya *et al.*, 2003). This is likely due to a number of factors, including the weaker ketone zinc-binding moiety of kAP and kAW, and the loss of the P_2' Pro moiety.

Because the ketone zinc binding moiety only takes on its zinc-binding hydrated state in the active site, it may prove that ketone inhibitors with the right side chain moieties are highly specific *in vivo*. The ketone group does have the disadvantage of eliminating the scissile amide nitrogen however, so that the carbonyl oxygen of A354 is left without hydrogen bonding partners in an arrangement which cannot favour binding. Replacement of the ketone with a different zinc binding group would allow the amide nitrogen to be replaced and introduce an additional hydrogen bond, however the resulting increase in affinity would not be domain-selective.

3.9.2. Selectivity and affinity comparisons with other inhibitors

Pro in the P_2' position has been known to be favoured by both domains since the discovery of the BPPs and the development of the Ala-Pro analogue inhibitor captopril, and is thought to be an effect of the conformational constraint on the prolyl backbone (Michaud *et al.*, 1999). This effect is particularly evident in the affinity of lisW-S for the C domain, which is approximately 140-fold lower than that of lisinopril, despite the fact that both compounds have a strong carboxyl zinc-binding group (Table 3.12).

Although lisW-S and kAW do have lower affinities than their parent compounds with a P_2' Pro, a P_2' Trp moiety is not necessarily unfavourable for C domain binding, as shown by the high affinity of the C-domain-selective inhibitor RXPA380 (Georgiadis *et al.*, 2004). This compound shares P_2' Trp, P_1 Phe and P_2 benzyl moieties with kAW, and is only two-fold more C-domain-selective, however its affinity for the C domain is approximately 230-fold higher than that of kAW (Georgiadis *et al.*, 2004). Drug compounds trandolaprilat, perindoprilat and ramiprilat, which bind with high affinity and slight C-domain-selectivity, also have bicyclic P_2' groups combined with a proline-like backbone rigidity in the P_2' position (Perich *et al.*, 1994; Acharya *et al.*, 2003).

RXPA380 differs from kAW in its P_1' pseudo-Pro moiety, phosphinic zinc-binding group and P_2 pseudo-Phe. The P_1' pseudo-proline moiety of RXPA380 may alter binding kinetics by imposing a greater degree of rigidity on the substrate, thus reducing the unfavourable entropic

effects of binding, and perhaps additionally enforcing a P_2' conformation that is unfavourable for binding to the N domain. The pseudo-proline probably also has higher affinity for the S_1' subsite than the pseudo-glycine of kAW, as does the phosphinic zinc-binding moiety in comparison with a ketone group. The P_2 pseudo-Phe, which is bulkier than the benzoyl moiety of kAW, probably also has an effect on binding, although this group should be more flexible than the CBz group and thus less likely to clash with Y369 of the N domain.

Interestingly, the affinity of lisW-S for the C domain is approximately 10-fold lower than that of kAW and kAF, despite the additional polar interactions contributed by the P_1' moiety. This may be the result of the absence of the favourable contribution of a P_2 group and perhaps also the unfavourable conformational entropy contribution required to constrain the flexible Lys sidechain.

3.9.3. Suggestions for future inhibitor design

To take full advantage of the weak zinc-binding ketone group, C-domain-selectivity could be enhanced by increasing the affinity and selectivity of the side chain moieties. Attention should be paid to increasing the size of the P_2' moiety, since it is evident from the structures that a P_2' Trp or Phe does not fill the pocket sufficiently to take full advantage of all the potential domain-specific contacts. Trp, Phe and Arg at this position have been shown to favour the C domain (Dive *et al.*, 1999; Acharya *et al.*, 2003). Backbone rigidity might also be introduced as in trandolaprilat, perindoprilat and ramiprilat (de Lima, 1999).

The close interaction between the P_1' and P_2' moieties of lisW-S, in contrast with kAW in which the P_2' Trp takes a different conformation, reveals cooperativity between side chains binding to the S_1' and S_2' pockets. The participation of the C-domain-specific residue, V380, in this interaction means that this characteristic could be exploited for enhanced selectivity and affinity. This could be achieved by introducing a fused moiety that would bind to both sites, or by experimenting with different combinations of P_1' and P_2' moieties.

The proximity of the P_1 and P_1' moieties to the zinc-binding group limits the degrees of freedom available to these side chains such that modifications to these groups which favour binding to one domain over the other, are likely to have a marked effect on domain-selectivity. Possible candidate moieties for the P_1' position are Lys, Phe and Trp, which have

been shown to be cleaved more efficiently by the C domain than the N domain, and Pro, which resists hydrolysis (Bersanetti *et al.*, 2004). At the P_1 position, Phe, Leu and Arg have been shown to favour the C domain (Araujo *et al.*, 2000; Bersanetti *et al.*, 2004).

Introduction of a nitrogen into the P_2 benzoyl ring would allow the formation of an additional hydrogen bond with the carbonyl oxygen of A356. Since the benzoyl ring is in close proximity to F391 (Y369 in the N domain), this could enhance both affinity and domain-selectivity. Ile, Ser and Lys at this position are also favoured by the C domain (Bersanetti *et al.*, 2004).

The combination of some or all of these alterations, together with the use of a weak zinc-binding group such as a ketone or phosphinate, should lead to the development of a new generation of C-domain-selective ACE inhibitors, presenting a different pharmacological profile to the inhibitors currently available for clinical use.

CHAPTER 4: Conclusions

4.1. tACE-G13 co-crystal structures

This study describes the first use of the minimally glycosylated tACE mutant tACE-G13 for co-crystal structure determination in the drug design process towards the development of domain-selective ACE inhibitors for clinical use. The structures of novel C-domain-selective inhibitors kAW, kAF and lisW-S have been solved bound to the active site of tACE-G13 using X-ray crystallography, to 2.18, 2.17 and 2.30 Å, respectively. No change was observed in the conformation of the protein component in comparison with previously determined structures, although an additional glycan residue could be built into the models for kAW and kAF.

4.2. Visualisation of the gem-diol transition state in ketone inhibitors

The ketone zinc-binding moiety of kAW and kAF was found to be in the hydrated *gem*-diol state in the active site, equivalent to the transition state. The stabilisation of this state in the absence of a leaving amide group demonstrates the ability of the active site to stabilise the transition state transiently during peptide bond cleavage. These structures double the number of reported visualisations of the formation of a *gem*-diol hydrated state from a ketone precursor in the active site of an enzyme, the other two cases being co-crystal structures of carboxypeptidase A published 20 years ago and not deposited in the PDB (Christianson *et al.*, 1987; Shoham *et al.*, 1988).

4.3. Selection for L-amino acid enantiomer of lisW

Although two enantiomers of the lisinopril derivative, lisW-S and lisW-R, were used during co-crystallisation, only one isoform was detected in the active site. This suggests that the active site of ACE selects strongly for the lisW-S enantiomer, which has L-amino acid chirality in the P_1 position, probably by steric interference between the beta carbon of the P_1 side chain and the hydroxyl moiety of Y523. The D-amino acid enantiomer lisW-R is thus excluded from binding to the active site.

4.4. Identification of inhibitor interactions important for domain-selectivity

All three inhibitors take a similar backbone conformation to that observed in the co-crystal structures of tACE with other peptide inhibitors. The positions of their side chains allowed a detailed description of the S_2, S_1, S_1' and S_2' pockets and the identification of residues which could contribute to domain-selective binding. The moderate domain-selectivity of kAP can be attributed to four mutations in the P_1 and P_2 pockets of the N domain relative to the C domain, while additional selectivity of the derivatives kAW and kAF in the P_2' pocket can be attributed to five inter-domain mutations. The lisinopril derivative inhibitor lisW-S derives its domain-selectivity from the same five P_2' inter-domain mutations, as well as from mutations in the P_1 and P_1' pockets. Several of these residues, E376, V380, Y391 and V518, have subsequently been confirmed by mutagenesis to be critical for the selectivity of kAW and kAF (Watermeyer *et al.*, Appendix D).

4.5. Multiple conformations in the P_2' position

The P_2' Trp moiety takes different conformations in kAW and lisW-S, due to hydrophobic stacking interactions of the Trp in the lisW-S structure with the P_1' Lys moiety and V380. In kAW the Trp indole nitrogen makes a hydrogen bond to the water network, while in lisW-S the indole nitrogen is hydrogen-bonded directly to D415. Moreover, the Phe in this position in kAF was found to be partially disordered, and a glycerol molecule shares this pocket with the Trp of kAW. This indicates that the P_2' pocket could accommodate still larger side chains, and highlights the importance of including hydrogen-bonding groups in this position for selective binding.

4.6. Conserved water molecules

Several conserved active site water molecules have been identified which might be displaced by hydrogen bonding groups in the next generation of domain-selective inhibitors. One of these, water 1165 of the kAW structure, is conserved in all the structures solved to date apart from the co-crystal structure with RXPA380. This water molecule connects the inhibitor backbone and crucial active site residues, E411 and Y523, via hydrogen bonds to R522, which is bound to chloride CL2. The positioning of this conserved water molecule may be critical for substrate binding, suggesting a reason for the R522-dependent chloride sensitivity of the C domain.

4.7. Suggestions for further inhibitor design

From these results, proposals can be made for future inhibitor design. The potential of the ketone zinc-binding group should examined, since its weak zinc affinity means that the potency of the inhibitor depends to a greater extent on side-chain interactions. These side chain interactions might be made more C-domain-selective by, for example, exploitation of water positions or by introduction of previously-determined C-domain-selective residues. P_2' residues should be bulkier and more rigid than Trp in order to take full advantage of selective interactions in this pocket, and consideration of cooperativity between the P_1' and P_2' positions might reveal novel means of achieving selective binding.

4.8. Broader perspective

By allowing the N domain to continue to function unimpeded, a drug compound targeted to the C domain of ACE is predicted to result in the alleviation of side effects associated with ACE inhibitor therapy. The results reported here represent a step along the pathway towards achieving this goal.

CHAPTER 5: References

Acharya KR, Sturrock ED, Riordan JF, and Ehlers MR (2003). ACE revisited: a new target for structure-based drug design. *Nature Reviews Drug Discovery 2:* 891-902.

Almquist RG, Chao WR, Ellis ME, and Johnson HL (1980). Synthesis and biological activity of a ketomethylene analogue of a tripeptide inhibitor of angiotensin converting enzyme. *Journal of Medicinal Chemistry 23:* 1392-1398.

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

Araujo MC, Melo RL, Cesari MH, Juliano MA, Juliano L, and Carmona AK (2000). Peptidase Specificity Characterization of C- and N-Terminal Catalytic Sites of Angiotensin I-Converting Enzyme. *Biochemistry 39:* 8519-8525.

Arndt JW, Hao B, Ramakrishnan V, Cheng T, Chan SI, and Chan MK (2002). Crystal structure of a novel carboxypeptidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Structure 10:* 215-224.

Bersanetti PA, Andrade MC, Casarini DE, Juliano MA, Nchinda AT, Sturrock ED, Juliano L, and Carmona AK (2004). Positional-scanning combinatorial libraries of fluorescence resonance energy transfer peptides for defining substrate specificity of the angiotensin I-converting enzyme and development of selective C-domain substrates. *Biochemistry 43:* 15729-15736.

Binevski PV, Sizova EA, Pozdnev VF, and Kost OA (2003). Evidence for the negative cooperativity of the two active sites within bovine somatic angiotensin-converting enzyme. *FEBS Letters 550:* 84-88.

Bingham RJ, Dive V, Phillips SE, Shirras AD, and Isaac RE (2006). Structural diversity of angiotensin-converting enzyme. *FEBS Journal 273:* 362-373.

Brown CK, Madauss K, Lian W, Beck MR, Tolbert WD, and Rodgers DW (2001). Structure of neurolysin reveals a deep channel that limits substrate access. *Proceedings of the National academy of Sciences of the United States of America 98:* 3127-3132.

Brünger AT (1992). Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature 335:* 472-475.

Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, and Warren GL (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallographica Section D: Biological Crystallography 54:* 905-921.

Bunning P, Holmquist B, and Riordan JF (1983). Substrate specificity and kinetic characteristics of angiotensin converting enzyme. *Biochemistry 22:* 103-110.

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

Christianson DW, David PR, and Lipscomb WN (1987). Mechanism of Carboxypeptidase A: Hydration of a Ketonic Substrate Analogue. *Proceedings of the National Academy of Sciences of the United States of America 84:* 1512-1515.

Coates L, Erskine PT, Mall S, Williams PA, Gill RS, Wood SP, and Cooper JB (2003). The structure of endothiapepsin complexed with the gem-diol inhibitor PD-135,040 at 1.37 angstrom. *Acta Crystallographica Section D-Biological Crystallography 59:* 978-981.

Cohen GE (1997). ALIGN: A program to superimpose protein coordinates, accounting for insertions and deletions. *Journal of Applied Crystallography 30:* 1160-1161.

Collaborative Computational Project No. 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallographica Section D-Biological Crystallography 50:* 760-763.

Comellas-Bigler M, Lang R, Bode W, and Maskos K (2005). Crystal structure of the *E. coli* dipeptidyl carboxypeptidase Dcp: further indication of a ligand-dependent hinge movement mechanism. *Journal of Molecular Biology 349:* 99-112.

Corradi HR, Chitapi I, Sewell BT, Georgiadis D, Dive V, Sturrock ED, and Acharya KR (2007). The structure of testis angiotensin-converting enzyme in complex with the C domain-specific inhibitor RXPA380. *Biochemistry 46:* 5473-5478.

Corradi HR, Schwager SL, Nchinda AT, Sturrock ED, and Acharya KR (2006). Crystal structure of the N domain of human somatic angiotensin I-converting enzyme provides a

structural basis for domain-specific inhibitor design. *Journal of Molecular Biology 357:* 964-974.

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

Davis IW, Murray LW, Richardson JS, and Richardson DC (2004). MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Research 32:* W615-W619.

de Lima DP (1999). Synthesis of angiotensin-converting enzyme (ACE) inhibitors: An important class of antihypertensive drugs. *Quimica Nova 22:* 375-381.

Deddish PA, Marcic B, Jackman HL, Wang HZ, Skidgel RA, and Erdös EG (1998). N-domain-specific substrate and C-domain inhibitors of angiotensin-converting enzyme: angiotensin-(1-7) and keto-ACE. *Hypertension 31:* 912-917.

Dive V, Cotton J, Yiotakis A, Michaud A, Vassiliou S, Jiracek J, Vazeux G, Chauvet MT, Cuniasse P, and Corvol P (1999). RXP 407, a phosphinic peptide, is a potent inhibitor of angiotensin I converting enzyme able to differentiate between its two active sites. *Proceedings of the National Academy of Sciences of the United States of America 96:* 4330-4335.

Donoghue M, Hsieh F, Baronas E, Godbout K, Gosselin M, Stagliano N, Donovan M, Woolf B, Robison K, Jeyaseelan R, Breitbart RE, and Acton S (2000). A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9. *Circulation Research 87:* E1-E9.

Ehlers MR, Chen YN, and Riordan JF (1991). Purification and characterization of recombinant human testis angiotensin-converting enzyme expressed in Chinese hamster ovary cells. *Protein Expression and Purification 2:* 1-9.

Ehlers MR, Fox EA, Strydom DJ, and Riordan JF (1989). Molecular cloning of human testicular angiotensin-converting enzyme: the testis isozyme is identical to the C-terminal half of endothelial angiotensin-converting enzyme. *Proceedings of the National Academy of Sciences of the United States of America 86:* 7741-7745.

Ehlers MR and Riordan JF (1989). Angiotensin-converting enzyme: new concepts concerning its biological role. *Biochemistry 28:* 5311-5318.

Ehlers MR and Riordan JF (1991). Angiotensin-converting enzyme: zinc- and inhibitor-binding stoichiometries of the somatic and testis isozymes. *Biochemistry 30:* 7118-7126.

Ehlers MR, Schwager SL, Scholle RR, Manji GA, Brandt WF, and Riordan JF (1996). Proteolytic release of membrane-bound angiotensin-converting enzyme: role of the juxtamembrane stalk sequence. *Biochemistry 35:* 9549-9559.

Erdös EG (1990). Angiotensin-I Converting Enzyme and the Changes in Our Concepts Through the Years - Dahl,Lewis,K. Memorial Lecture. *Hypertension 16:* 363-370.

Eriksson U, Danilczyk U, and Penninger JM (2002). Just the beginning: Novel functions for angiotensin-converting enzymes. *Current Biology 12:* R745-R752.

Esther CR, Jr., Semeniuk D, Marino EM, Zhou Y, Overbeek PA, and Bernstein KE (1997). Expression of testis angiotensin-converting enzyme is mediated by a cyclic AMP responsive element. *Laboratory Investigation 77:* 483-488.

Esunge PM (1991). From blood pressure to hypertension: the history of research. *Journal of the Royal Society of Medicine 84:* 621

Fleming I, Kohlstedt K, and Busse R (2006). The tissue renin-angiotensin system and intracellular signalling. *Current Opinion in Nephrology and Hypertension 15:* 8-13.

Friedland J and Silverstein E (1976). A sensitive fluorimetric assay for serum angiotensin-converting enzyme. *American Journal of Clinical Pathology 66:* 416-424.

Fuchs S, Xiao HD, Cole JM, Adams JW, Frenzel K, Michaud A, Zhao H, Keshelava G, Capecchi MR, Corvol P, and Bernstein KE (2004). Role of the N-terminal catalytic domain of angiotensin-converting enzyme investigated by targeted inactivation in mice. *Journal of Biological Chemistry 279:* 15946-15953.

Fuchs S, Xiao HD, Hubert C, Michaud A, Campbell DJ, Adams JW, Capecchi MR, Corvol P, and Bernstein KE (2007). Angiotensin-Converting Enzyme C-Terminal Catalytic Domain Is the Main Site of Angiotensin I Cleavage In Vivo. *Hypertension* (article in press).

Georgiadis D, Beau F, Czarny B, Cotton J, Yiotakis A, and Dive V (2003). Roles of the two active sites of somatic angiotensin-converting enzyme in the cleavage of angiotensin I and bradykinin - Insights from selective inhibitors. *Circulation Research 93:* 148-154.

Georgiadis D, Cuniasse P, Cotton J, Yiotakis A, and Dive V (2004). Structural determinants of RXPA380, a potent and highly selective inhibitor of the angiotensin-converting enzyme C-domain. *Biochemistry 43:* 8048-8054.

Gerdts CJ, Tereshko V, Yadav MK, Dementieva I, Collart F, Joachimiak A, Stevens RC, Kuhn P, Kossiakoff A, and Ismagilov RF (2006). Time-controlled microfluidic seeding in nL-volume droplets to separate nucleation and growth stages of protein crystallization. *Angewandte Chemie International Edition England 45:* 8156-8160.

Gordon K, Redelinghuys P, Schwager SL, Ehlers MR, Papageorgiou AC, Natesh R, Acharya KR, and Sturrock ED (2003). Deglycosylation, processing and crystallization of human testis angiotensin-converting enzyme. *Biochemical Journal 371:* 437-442.

Guy JL, Jackson RM, Acharya KR, Sturrock ED, Hooper NM, and Turner AJ (2003). Angiotensin-converting enzyme-2 (ACE2): comparative modeling of the active site, specificity requirements, and chloride dependence. *Biochemistry 42:* 13185-13192.

Hayakari M, Satoh K, Ookawa K, Kano H, Murakami S, Ikeda N, and Tsuchida S (1997). Kinetic evaluation of beta-neoendorphin hydrolysis by the somatic and testicular isozymes of human angiotensin-converting enzyme. *Biochimica et Biophysica Acta-Protein Structure and Molecular Enzymology 1339:* 31-38.

Hayashi MAF and Camargo ACM (2005). The Bradykinin-potentiating peptides from venom gland and brain of *Bothrops jararaca* contain highly site specific inhibitors of the somatic angiotensin-converting enzyme. *Toxicon 45:* 1163-1170.

Hemming ML and Selkoe DJ (2005). Amyloid beta-protein is degraded by cellular angiotensin-converting enzyme (ACE) and elevated by an ACE inhibitor. *Journal of Biological Chemistry 280:* 37644-37650.

Hogg T, Kuta S, I, Bezouska K, Ulbrich N, and Hilgenfeld R (2002). Sugar-mediated lattice contacts in crystals of a plant glycoprotein. *Acta Crystallographica Section D: Biological Crystallography 58:* 1734-1739.

Hooper NM (1994). Families of Zinc Metalloproteases. *FEBS Letters 354:* 1-6.

Hubert C, Houot AM, Corvol P, and Soubrier F (1991). Structure of the angiotensin I-converting enzyme gene. Two alternate promoters correspond to evolutionary steps of a duplicated gene. *Journal of Biological Chemistry 266:* 15377-15383.

Ibarrarubio ME, Gomezvelez LA, Cruz C, and Pedrazachaverri J (1993). Effect of Chloride and Diamide on Serum Angiotensin I-Converting Enzyme-Activity from 8 Mammalian-Species. *Comparative Biochemistry and Physiology C-Pharmacology Toxicology & Endocrinology 106:* 495-501.

Inokuchi J and Nagamatsu A (1981). Tripeptidyl carboxypeptidase activity of kininase II (angiotensin-converting enzyme). *Biochimica et Biophysica Acta 662:* 300-307.

Isaac RE, Lamango NS, Ekbote U, Taylor CA, Hurst D, Weaver RJ, Carhan A, Burnham S, and Shirras AD (2007). Angiotensin-converting enzyme as a target for the development of novel insect growth regulators. *Peptides 28:* 153-162.

Isaac RE, Michaud A, Keen JN, Williams TA, Coates D, Wetsel WC, and Corvol P (1999). Hydrolysis by somatic angiotensin-I converting enzyme of basic dipeptides from a cholecystokinin/gastrin and a LH-RH peptide extended at the C-terminus with gly-Arg/Lys-arg, but not from diarginyl insulin. *European Journal of Biochemistry 262:* 569-574.

Jaspard E, Wei L, and Alhencgelas F (1993). Differences in the Properties and Enzymatic Specificities of the 2 Active-Sites of Angiotensin I-Converting Enzyme (Kininase-Ii) - Studies with Bradykinin and Other Natural Peptides. *Journal of Biological Chemistry 268:* 9496-9503.

Jones TA (1978). A graphics model building and refinement system for macromolecules. *Journal of Applied Crystallography 11:* 268-272.

Jullien ND, Cuniasse P, Georgiadis D, Yiotakis A, and Dive V (2006). Combined use of selective inhibitors and fluorogenic substrates to study the specificity of somatic wild-type angiotensin-converting enzyme. *FEBS Journal 273:* 1772-1781.

Junot C, Gonzales MF, Ezan E, Cotton J, Vazeux G, Michaud A, Azizi M, Vassiliou S, Yiotakis A, Corvol P, and Dive V (2001). RXP 407, 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. *Journal of Pharmacology and Experimental Therapeutics 297:* 606-611.

Kim HM, Shin DR, Yoo OJ, Lee H, and Lee JO (2003). Crystal structure of Drosophila angiotensin I-converting enzyme bound to captopril and lisinopril. *FEBS Letters 538:* 65-70.

Kissinger CR, Gehlhaar DK, and Fogel DB (1999). Rapid automated molecular replacement by evolutionary search. *Acta Crystallographica Section D: Biological Crystallography 55:* 484-491.

Kramers C, Danilov SM, Deinum J, Balyasnikova IV, Scharenborg N, Looman M, Boomsma F, de Keijzer MH, van Duijn C, Martin S, Soubrier F, and Adema GJ (2001). Point mutation in the stalk of angiotensin-converting enzyme causes a dramatic increase in serum angiotensin-converting enzyme but no cardiovascular disease. *Circulation 104:* 1236-1240.

Langford KG, Zhou Y, Russell LD, Wilcox JN, and Bernstein KE (1993). Regulated expression of testis angiotensin-converting enzyme during spermatogenesis in mice. *Biology of Reproduction 48:* 1210-1218.

Lim EJ, Sampath S, Coll-Rodriguez J, Schmidt J, Ray K, and Rodgers DW (2007). Swapping the substrate specificities of the neuropeptidases neurolysin and thimet oligopeptidase. *Journal of Biological Chemistry 282:* 9722-9732.

Lipscomb WN and Sträter N (1996). Recent advances in zinc enzymology. *Chemical Reviews 96:* 2375-2433.

Liu X, Fernandez M, Wouters MA, Heyberger S, and Husain A (2001). Arg(1098) is critical for the chloride dependence of human angiotensin I-converting enzyme C-domain catalytic activity. *Journal of Biological Chemistry 276:* 33518-33525.

Macours N, Poels J, Hens K, Francis C, and Huybrechts R (2004). Structure, evolutionary conservation, and functions of angiotensin- and endothelin-converting enzymes. *International Reviews in Cytology 239:* 47-97.

Mangani S, Carloni P, and Orioli P (1992). Crystal structure of the complex between carboxypeptidase A and the biproduct analog inhibitor L-benzylsuccinate at 2.0 A resolution. *Journal of Molecular Biology 223:* 573-578.

Marcic B, Deddish PA, Jackman HL, Erdös EG, and Tan FL (2000). Effects of the N-terminal sequence of ACE on the properties of its C-domain. *Hypertension 36:* 116-121.

McDonald IK and Thornton JM (1994). Satisfying hydrogen bonding potential in proteins. *Journal of Molecular Biology 238:* 777-793.

Mehndiratta P, Walton WJ, Hare JT, Pulido S, Parthasarathy G, Emmett MR, Marshall AG, and Logan TM (2004). Expression, purification, and characterization of avian Thy-1 from Lec1 mammalian and Tn5 insect cells. *Protein Expression and Purification 33:* 274-287.

Michaud A, Chauvet MT, and Corvol P (1999). N-domain selectivity of angiotensin I-converting enzyme as assessed by structure-function studies of its highly selective substrate, N-acetyl-seryl-aspartyl-lysyl-proline. *Biochemical Pharmacology 57:* 611-618.

Moreau ME, Garbacki N, Molinaro G, Brown NJ, Marceau F, and Adam A (2005). The kallikrein-kinin system: current and future pharmacological targets. *Journal of Pharmacological Sciences 99:* 6-38.

Moretti S, Armougom F, Wallace IM, Higgins DG, Jongeneel CV, and Notredame C (2007). The M-Coffee web server: a meta-method for computing multiple sequence alignments by combining alternative alignment methods. *Nucleic Acids Research 35:* W645-W648.

Nagata K, Tsutsui S, Lee WC, Ito K, Kamo M, Inoue Y, and Tanokura M (2004). Crystallization and preliminary X-ray analysis of carboxypeptidase 1 from *Thermus thermophilus*. *Acta Crystallographica Section D: Biological Crystallography 60:* 1445-1446.

Naqvi N, Liu K, Graham RM, and Husain A (2005). Molecular basis of exopeptidase activity in the C-terminal domain of human angiotensin I-converting enzyme: insights into the origins of its exopeptidase activity. *Journal of Biological Chemistry 280:* 6669-6675.

Natesh R, Schwager SL, Evans HR, Sturrock ED, and Acharya KR (2004). Structural details on the binding of antihypertensive drugs captopril and enalaprilat to human testicular angiotensin I-converting enzyme. *Biochemistry 43:* 8718-8724.

Natesh R, Schwager SL, Sturrock ED, and Acharya KR (2003). Crystal structure of the human angiotensin-converting enzyme-lisinopril complex. *Nature 421:* 551-554.

Nchinda AT, Chibale K, Redelinghuys P, and Sturrock ED (2006a). Synthesis and molecular modeling of a lisinopril-tryptophan analogue inhibitor of angiotensin I-converting enzyme. *Bioorganic and Medicinal Chemistry Letters 16:* 4616-4619.

Nchinda AT, Chibale K, Redelinghuys P, and Sturrock ED (2006b). Synthesis of novel keto-ACE analogues as domain-selective angiotensin I-converting enzyme inhibitors. *Bioorganic and Medicinal Chemistry Letters 16:* 4612-4615.

Oba R, Igarashi A, Kamata M, Nagata K, Takano S, and Nakagawa H (2005). The N-terminal active centre of human angiotensin-converting enzyme degrades Alzheimer amyloid beta-peptide. *European Journal of Neuroscience 21:* 733-740.

Ocaranza MP, Godoy I, Jalil JE, Varas M, Collantes P, Pinto M, Roman M, Ramirez C, Copaja M, Diaz-Araya G, Castro P, and Lavandero S (2006). Enalapril attenuates downregulation of Angiotensin-converting enzyme 2 in the late phase of ventricular dysfunction in myocardial infarcted rat. *Hypertension 48:* 572-578.

Ondetti MA, Rubin B, and Cushman DW (1977). Design of specific inhibitors of angiotensin-converting enzyme: new class of orally active antihypertensive agents. *Science 196:* 441-444.

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

Pelmenschikov V, Blomberg MRA, and Siegbahn PE (2002). A theoretical study of the mechanism for peptide hydrolysis by thermolysin. *Journal of Biological Inorganic Chemistry 7:* 284-298.

Perich RB, Jackson B, and Johnston CI (1994). Structural constraints of inhibitors for binding at two active sites on somatic angiotensin converting enzyme. *European Journal of Pharmacology 266:* 201-211.

Rawlings ND and Barrett AJ (1995). Evolutionary families of metallopeptidases. *Methods in Enzymology 248:* 183-228.

Ray K, Hines CS, Coll-Rodriguez J, and Rodgers DW (2004). Crystal structure of human thimet oligopeptidase provides insight into substrate recognition, regulation, and localization. *Journal of Biological Chemistry 279:* 20480-20489.

Redelinghuys P, Nchinda AT, Chibale K, and Sturrock ED (2006). Novel ketomethylene inhibitors of angiotensin I-converting enzyme (ACE): inhibition and molecular modelling. *Biological Chemistry 387:* 461-466.

Redelinghuys P, Nchinda AT, and Sturrock ED (2005). Development of domain-selective Angiotensin I-converting enzyme inhibitors. *Annals of the New York Academy of Science 1056:* 160-175.

Rigat B, Hubert C, Alhenc-Gelas F, Cambien F, Corvol P, and Soubrier F (1990). An insertion/deletion polymorphism in the angiotensin I-converting enzyme gene accounting for half the variance of serum enzyme levels. *Journal of Clinical Investigation 86:* 1343-1346.

Riviere G, Michaud A, Corradi HR, Sturrock ED, Ravi AK, Cogez V, Bohin JP, Vieau D, and Corvol P (2007). Characterization of the first angiotensin-converting like enzyme in bacteria: Ancestor ACE is already active. *Gene 399:* 81-90.

Sayed-Tabatabaei FA, Oostra BA, Isaacs A, Van Duijn CM, and Witteman JCM (2006). ACE polymorphisms. *Circulation Research 98:* 1123-1133.

Schüttelkopf AW and van Aalten DM (2004). PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallographica Section D: Biological Crystallography 60:* 1355-1363.

Schwager SL, Carmona AK, and Sturrock ED (2006). A high-throughput fluorimetric assay for angiotensin I-converting enzyme. *Nature Protocols 1:* 1961-1964.

Shoham G, Christianson DW, and Oren DA (1988). Complex between carboxypeptidase A and a hydrated ketomethylene substrate analogue. *Proceedings of the National academy of Sciences of the United States of America 85:* 684-688.

Sica DA (2004). Angiotensin-converting enzyme inhibitors side effects--physiologic and non-physiologic considerations. *Journal of Clinical Hypertension 6:* 410-416.

Skeggs LT, Jr., Kahn JR, and Shumway NP (1956). The purification of hypertensin II. *Journal of Experimental Medicine 103:* 301-307.

Skeggs LT, Jr., Marsh WH, Kahn JR, and Shumway NP (1954). The existence of two forms of hypertensin. *Journal of Experimental Medicine 99:* 275-282.

Skidgel RA, Engelbrecht S, Johnson AR, and Erdös EG (1984). Hydrolysis of Substance-P and Neurotensin by Converting Enzyme and Neutral Endopeptidase. *Peptides 5:* 769-776.

Skidgel RA and Erdös EG (1985). Novel Activity of Human Angiotensin-I Converting Enzyme - Release of the Nh2-Terminal and Cooh-Terminal Tripeptides from the Luteinizing-Hormone-Releasing Hormone. *Proceedings of the National Academy of Sciences of the United States of America 82:* 1025-1029.

Skidgel RA and Erdös EG (2004). Angiotensin converting enzyme (ACE) and neprilysin hydrolyze neuropeptides: a brief history, the beginning and follow-ups to early studies. *Peptides 25:* 521-525.

Skirgello OE, Balyasnikova IV, Binevski PV, Sun ZL, Baskin II, Palyulin VA, Nesterovitch AB, Albrecht RF, Kost OA, and Danilov SM (2006). Inhibitory antibodies to human angiotensin-converting enzyme: fine epitope mapping and mechanism of action. *Biochemistry 45:* 4831-4847.

Soubrier F, Alhenc-Gelas F, Hubert C, Allegrini J, John M, Tregear G, and Corvol P (1988). Two putative active centers in human angiotensin I-converting enzyme revealed by molecular cloning. *Proceedings of the National Academy of Sciences of the United States of America 85:* 9386-9390.

Spyroulias GA, Galanis AS, Pairas G, Manessi-Zoupa E, and Cordopatis P (2004). Structural features of angiotensin-I converting enzyme catalytic sites: conformational studies in solution, homology models and comparison with other zinc metallopeptidases. *Current Topics in Medicinal Chemistry 4:* 403-429.

Sträter N and Lipscomb WN (1995). Transition-State Analog L-Leucinephosphonic Acid Bound to Bovine Lens Leucine Aminopeptidase - X-Ray Structure at 1.65 Angstrom Resolution in A New Crystal Form. *Biochemistry 34:* 9200-9210.

Sturrock ED, Natesh R, van Rooyen JM, and Acharya KR (2004). Structure of angiotensin I-converting enzyme. *Cellular and Molecular Life Sciences 61:* 2677-2686.

Tipnis SR, Hooper NM, Hyde R, Karran E, Christie G, and Turner AJ (2000). A human homolog of angiotensin-converting enzyme - Cloning and functional expression as a captopril-insensitive carboxypeptidase. *Journal of Biological Chemistry 275:* 33238-33243.

Towler P, Staker B, Prasad SG, Menon S, Tang J, Parsons T, Ryan D, Fisher M, Williams D, Dales NA, Patane MA, and Pantoliano MW (2004). ACE2 X-ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis. *Journal of Biological Chemistry 279:* 17996-18007.

Turner AJ and Hooper NM (2002). The angiotensin-converting enzyme gene family: genomics and pharmacology. *Trends in Pharmacological Sciences 23:* 177-183.

Turner AJ, Tipnis SR, Guy JL, Rice GI, and Hooper NM (2002). ACEH/ACE2 is a novel mammalian metallocarboxypeptidase and a homologue of angiotensin-converting enzyme insensitive to ACE inhibitors. *Canadian Journal of Physiology and Pharmacology 80:* 346-353.

Tyndall JD, Nall T, and Fairlie DP (2005). Proteases universally recognize beta strands in their active sites. *Chemical Reviews 105:* 973-999.

Tzakos AG, Galanis AS, Spyroulias GA, Cordopatis P, Manessi-Zoupa E, and Gerothanassis IP (2003). Structure-function discrimination of the N- and C- catalytic domains of human angiotensin-converting enzyme: implications for Cl- activation and peptide hydrolysis mechanisms. *Protein Engineering 16:* 993-1003.

Tzakos AG and Gerothanassis IP (2005). Domain-selective ligand-binding modes and atomic level pharmacophore refinement in angiotensin I converting enzyme (ACE) inhibitors. *Chembiochem 6:* 1089-1103.

van Esch JHM, Tom B, Dive V, Batenburg WW, Georgiadis D, Yiotakis A, van Gool JMG, de Bruijn RJA, de Vries R, and Danser AHJ (2005). Selective angiotensin-converting enzyme c-domain inhibition is sufficient to prevent angiotensin I-induced vasoconstriction. *Hypertension 45:* 120-125.

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

Warner FJ, Smith AI, Hooper NM, and Turner AJ (2004). Angiotensin-converting enzyme-2: a molecular and cellular perspective. *Cellular and Molecular Life Sciences 61:* 2704-2713.

Watermeyer JM, Sewell BT, Schwager SL, Natesh R, Corradi HR, Acharya KR, and Sturrock ED (2006). Structure of testis ACE glycosylation mutants and evidence for conserved domain movement. *Biochemistry 45:* 12654-12663.

Wei L, Alhencgelas F, Corvol P, and Clauser E (1991). The 2 Homologous Domains of Human Angiotensin-I-Converting Enzyme Are Both Catalytically Active. *Journal of Biological Chemistry 266:* 9002-9008.

Wei L, Clauser E, Alhencgelas F, and Corvol P (1992). The 2 Homologous Domains of Human Angiotensin-I-Converting Enzyme Interact Differently with Competitive Inhibitors. *Journal of Biological Chemistry 267:* 13398-13405.

Williams TA, Michaud A, Houard X, Chauvet MT, Soubrier F, and 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. *Biochemical Journal 318:* 125-131.

Woodman ZL, Schwager SL, Redelinghuys P, Carmona AK, Ehlers MR, and Sturrock ED (2005). The N domain of somatic angiotensin-converting enzyme negatively regulates ectodomain shedding and catalytic activity. *Biochemical Journal 389:* 739-744.

Yu XC, Sturrock ED, Wu Z, Biemann K, Ehlers MR, and Riordan JF (1997). Identification of N-linked glycosylation sites in human testis angiotensin-converting enzyme and expression of an active deglycosylated form. *Journal of Biological Chemistry 272:* 3511-3519.

Zaman MA, Oparil S, and Calhoun DA (2002). Drugs targeting the renin-angiotensin-aldosterone system. *Nature Reviews Drug Discovery 1:* 621-636.

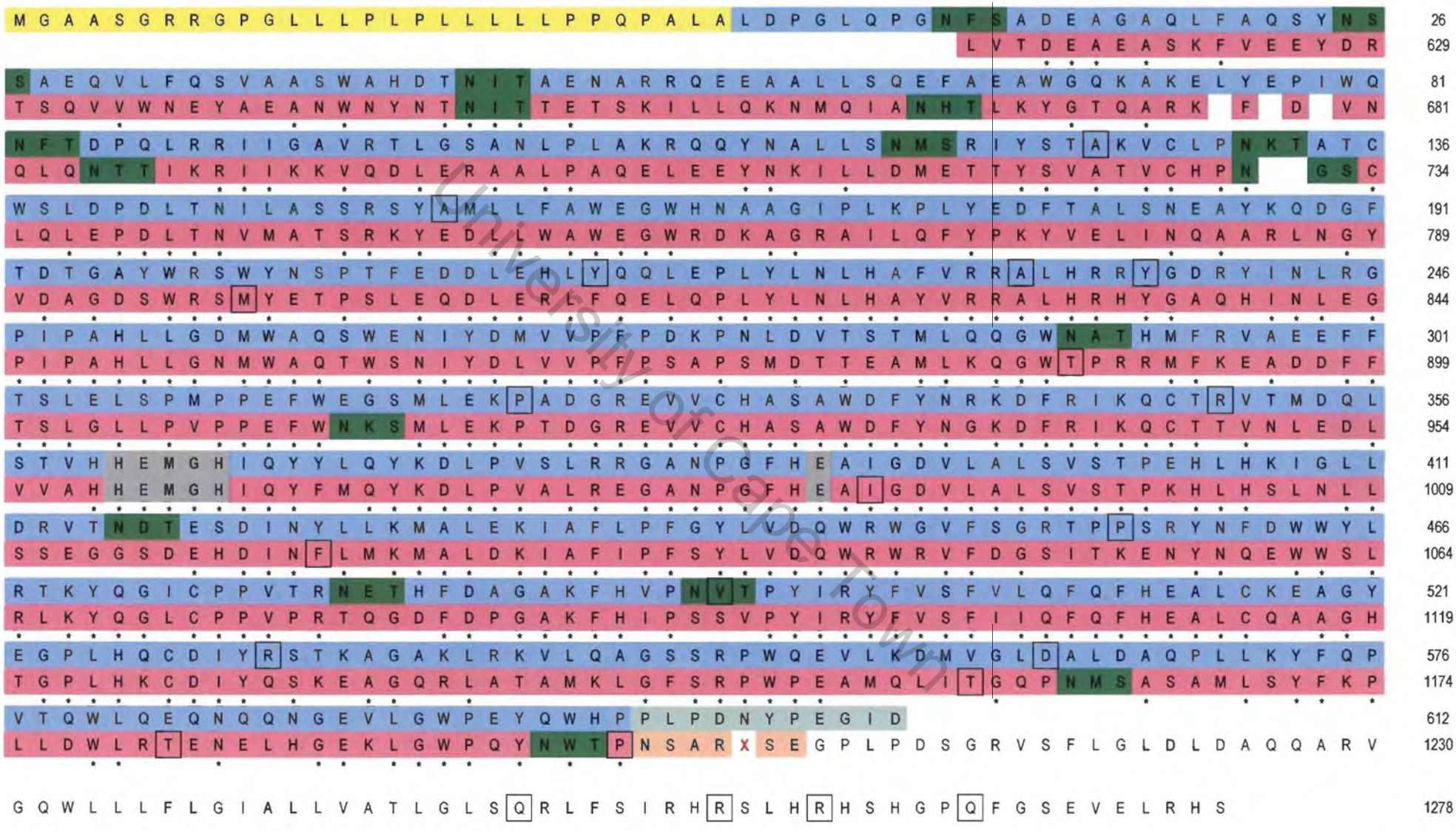
Zhang Y, Desharnais J, Marsilje TH, Li C, Hedrick MP, Gooljarsingh LT, Tavassoli A, Benkovic SJ, Olson AJ, Boger DL, and Wilson IA (2003). Rational design, synthesis,

evaluation, and crystal structure of a potent inhibitor of human GAR Tfase: 10-(trifluoroacetyl)-5,10-dideazaacyclic-5,6,7,8-tetrahydrofolic acid. *Biochemistry 42:* 6043-6056.

CHAPTER 6: Appendices

Appendix A: Structure-based sequence alignment of the N and C domains of human ACE

The following table shows a sequence alignment of the N and C domains of human somatic ACE (GI: 23238217). The table was generated using the structural alignment of the N and C domain structures, as calculated by *ALIGN* (Cohen, 1997), as a reference for aligning residues in regions of low homology. The signal sequence is highlighted yellow, the N domain blue, the linker peptide light green, the C domain pink, and the stalk region with cleavage site (red X) orange. Stars under the alignment mark residues that are conserved between the N and C domains. The transmembrane and cytoplasmic domains are not highlighted. The conserved zinc-binding motif is printed in bold text and highlighted grey, and putative glycosylation sites are highlighted green. The last N-linked glycosylation site in the C domain has been shown to be unglycosylated (Yu *et al.*, 1997). Numbering does not include the signal peptide, which is cleaved off during processing. The sites of human SNPs are boxed. Testis ACE is identical to the C domain but has an additional 56-residue N-terminal signal sequence and O-glycosylated region (not shown).



Appendix B: Polar interactions of tACE and the N domain with inhibitors lisinopril, captopril, enalapril and RXPA380

The following table describes the polar interactions of inhibitors lisinopril, captopril, enalapril and RXPA380, with the active sites of ACE, as determined from their co-crystal structures (Natesh *et al.*, 2004; Corradi *et al.*, 2006; Corradi *et al.*, 2007).

tACE interactions with captopril (CAP) enalapril (ENA), RXPA380 (RXP) and lisinopril (LPR)									N domain interactions with lisinopril (LPR)		
tACE atom [a]	**CAP atom, distance (Å)**		**ENA atom, distance (Å)**		**RXP atom, distance (Å)**		**LPR atom, distance (Å)**		**N domain atom** [b]	**LPR atom, distance (Å)**	
His 353 NE2	O1	2.5	O1	2.7	O5	2.9	O1	2.8	His 331 NE2	O1	3.0
Ala 354 O	-	-	N1	3.0	-	-	N1	2.9	Ala 332 O	N1	2.6
Glu 384 OE2	S1	3.1	O2	2.7	O4	2.6	O2	2.7	Glu 362 OE2	O2	2.7
Lys 511 NZ	O2	2.7	O4	2.8	O6	2.6	O4	2.9	Lys 489 NZ	O4	2.7
His 513 NE2	O1	2.7	O1	3.1	O5	2.9	O1	3.1	His 491 NE2	O1	2.8
Tyr 520 OH	O2	2.7	O4	2.6	O6	2.6	O4	2.6	Tyr 498 OH	O4	3.2
Tyr 523 OH	-	-	O3	2.8	O3	2.7	O3	2.8	Tyr 501 OH	O3	2.7
Glu 162 OE2	-	-	-	-	O2	2.7	N3	3.4	-	-	-
Wat 747 O	-	-	O2	2.9	-	-	O2	2.8	-	-	-
Wat 759 O	O3	2.6	O5	2.7	OXT	3.0	O5	2.7	-	-	-
Wat 787 O	-	-	-	-	-	-	N3	3.0	-	-	-
Wat 835 O	-	-	-	-	-	-	N3	3.3	-	-	-
Wat 866 O	O3	3.0	O5	2.7	-	-	O5	2.8	-	-	-
Wat 1038 O	-	-	-	-	-	-	N3	3.1	-	-	-

[a] Waters (Wat) are numbered according to the co-crystal structures of tACE with lisinopril, captopril and enalaprilat. [b] Residues are on the same line as their C domain equivalents.

Appendix C: Structural alignment of ACE homologues

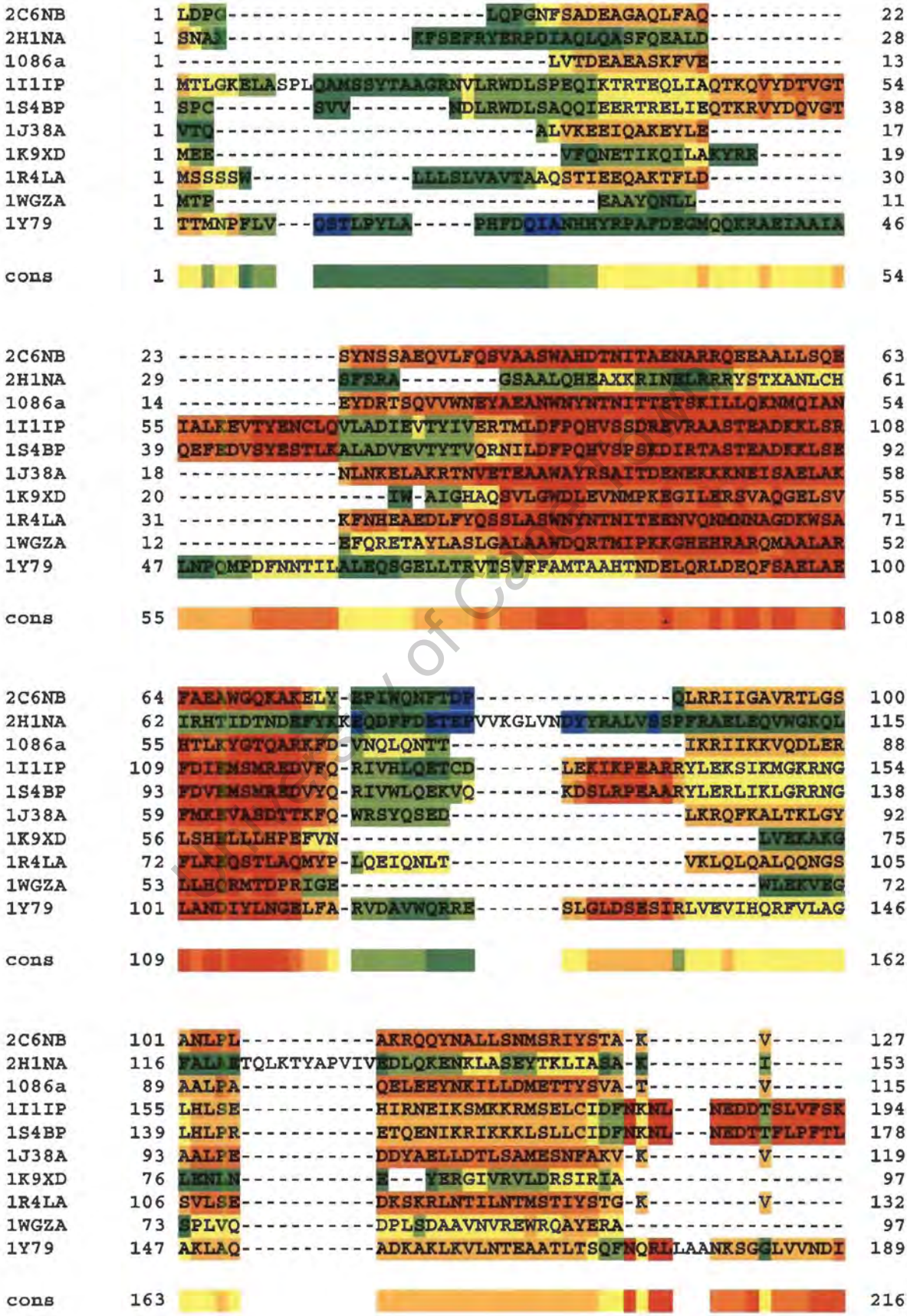
The following tables present a structure-based multiple sequence alignment of human ACE and ACE homologues that was generated using *M-Coffee* (Moretti *et al.*, 2007). This server uses a combination of eight different multiple sequence alignment methods and outputs an alignment score based on the agreement between the methods.

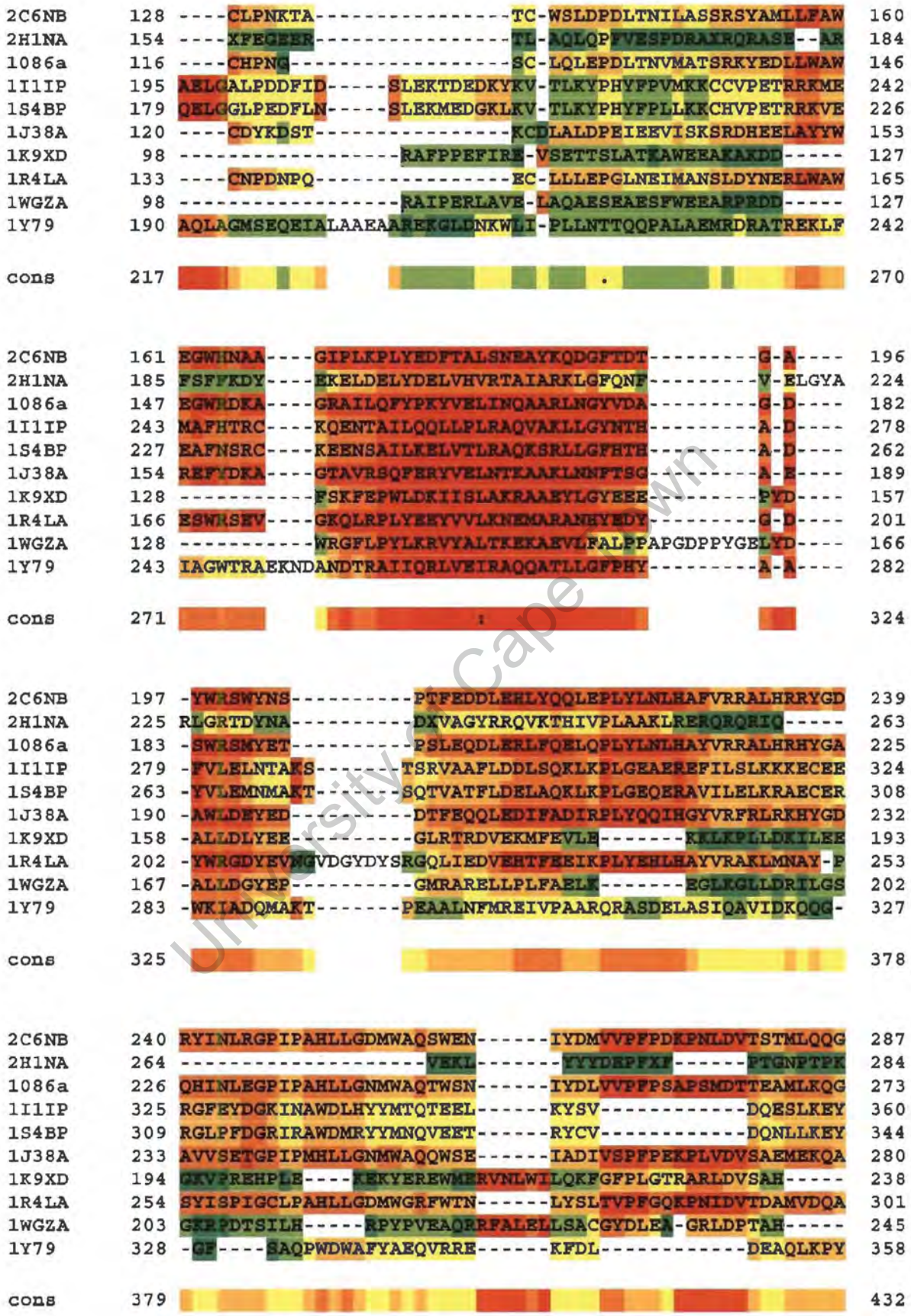
Table C.1. Overall alignment scores and key to Table C.2.

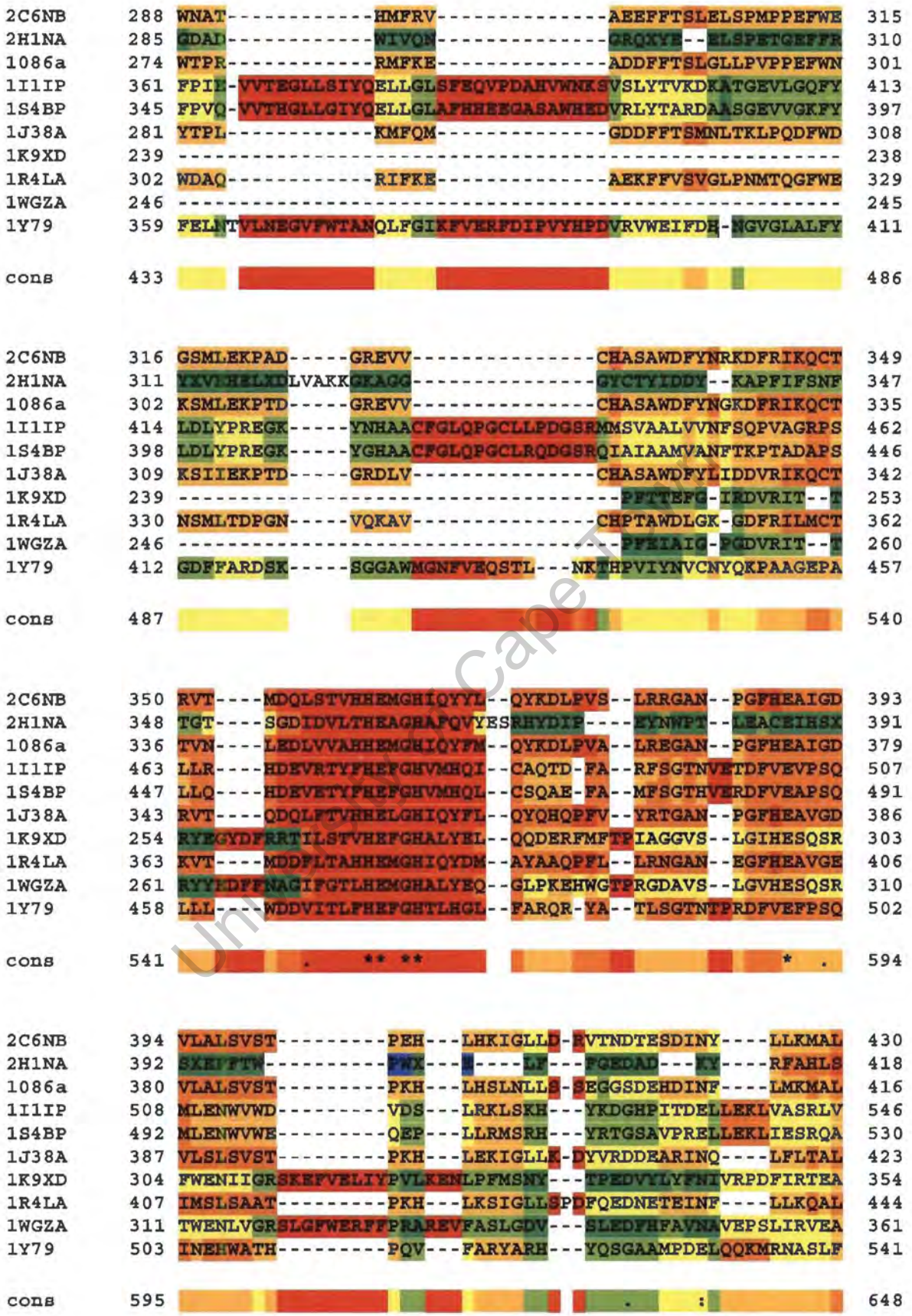
Structure	Code [a]	Alignment score [b]
Human tACE (C domain)	1086A	66
Drosophila melanogaster AnCE	1J38A	66
Human ACE N domain	2C6NB	65
Human ACE2	1R4LA	65
Rat neurolysin	1I1IP	58
Human thimet oligopeptidase	1S4BP	58
Escherischia coli dipeptidyl carboxypeptidase	1Y79	53
Pyrococcus furiosus carboxypeptidase 1	1K9XD	52
Thermus thermophilus carboxypeptidase 1	1WGZA	52
Bacillus stearothermophilus oligopeptidase F	2H1NA	40

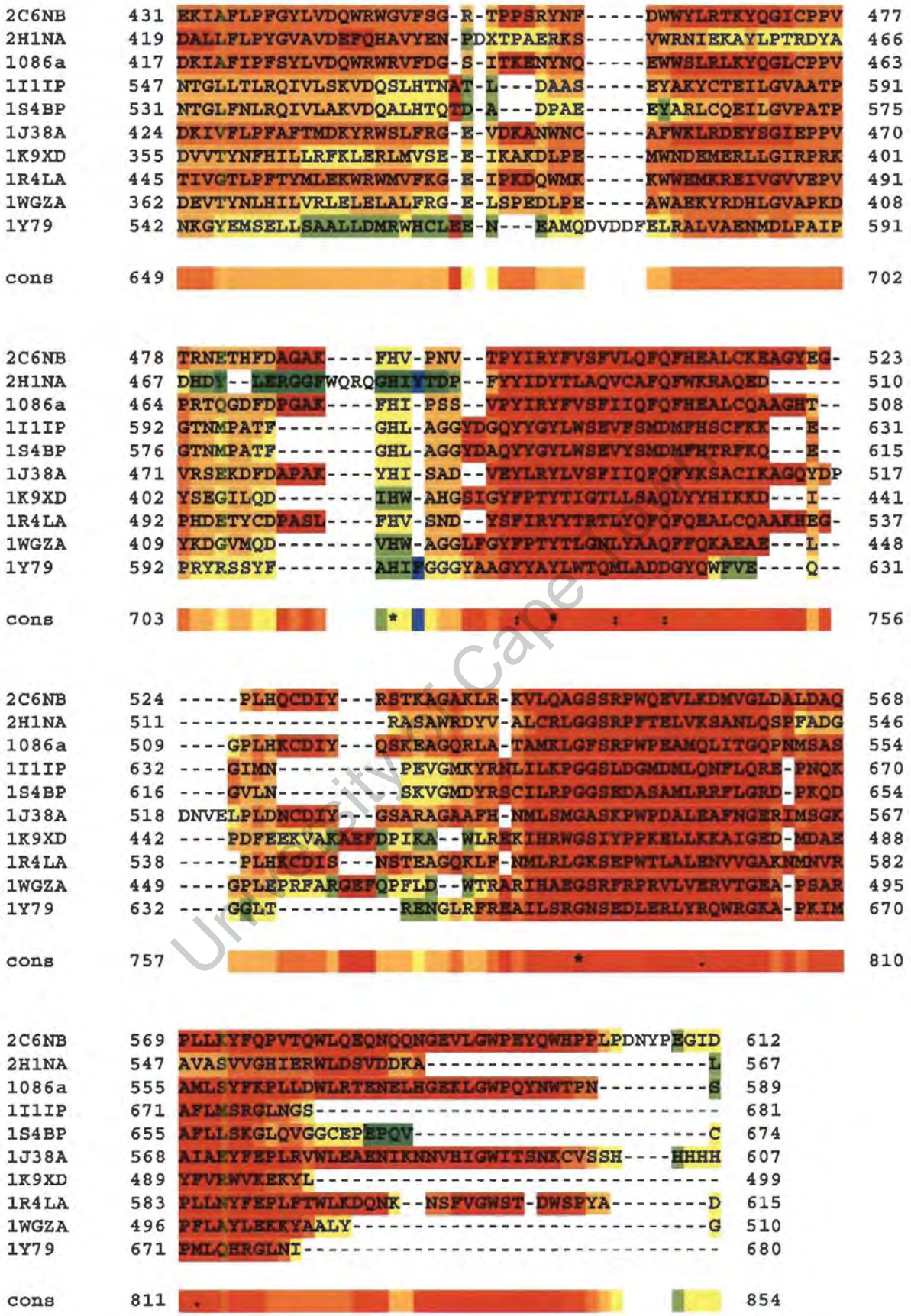
[a] Codes refer to the PDB ID and chain ID used for the alignment, where more than one chain was present in the structure. [b] Score based on the consistency of alignment between the different methods used. Colour key for alignment quality: BAD AVG GOOD

Table C.2. Structure-based multiple sequence alignment of human ACE and ACE homologues.









Appendix D: Watermeyer et al

The following manuscript has been submitted for publication in *Biochemistry*. It is appended here because it includes enzyme kinetic data determined by Wendy Kröger, another student in the group, which has been cited in this thesis in order to explain the results obtained for inhibitors kAW and kAF. The structural data described in the manuscript are the same data that are presented in this thesis and are the author's own work.

PROBING THE BASIS OF DOMAIN-DEPENDENT INHIBITION USING NOVEL KETONE INHIBITORS OF ANGIOTENSIN-CONVERTING ENZYME

Jean M. Watermeyer[1], Wendy L. Kröger[1], Hester G. O'Neill[1], B. Trevor Sewell[1], Edward D. Sturrock[1]

[1]Division of Medical Biochemistry, Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Observatory, 7925, South Africa

Running head: Domain-selectivity of ketone ACE inhibitors

Address correspondence to: Edward David Sturrock, Division of Medical Biochemistry, Wernher Beit Building North, University of Cape Town, Observatory, 7925, South Africa. Fax: +27 21 406 6470; Email: Edward.Sturrock@uct.ac.za

The atomic coordinates and structure factors (codes 3BKK and 3BKL) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

Human angiotensin-converting enzyme (ACE) has two homologous domains, the N and C domains, with differing substrate preferences. X-ray crystal structures of the C and N domains complexed with various inhibitors have indicated active site residues that might be important for the molecular basis of this selectivity. However it is unclear to what extent the different residues contribute towards substrate domain-selectivity. Here, co-crystal structures of human testis ACE, equivalent to the C domain, have been determined with two novel C-domain-selective ketomethylene inhibitors, (5S)-5-[(N-benzoyl)amino]-4-oxo-6-phenyl-hexanoyl-L-tryptophan (kAW) and (5S)-5-[(N-benzoyl)amino]-4-oxo-6-phenyl-hexanoyl-L-phenylalanine (kAF). The ketone groups of both inhibitors bind to the zinc ion as a hydrated geminal diolate, demonstrating the ability of the active site to catalyse the formation of the transition state. Moreover, active site residues involved in inhibitor binding have been mutated to their N domain counterparts and the effect of the mutations on inhibitor binding has been determined. The C-domain-selectivity of these inhibitors was found to result from interactions between bulky hydrophobic side chain moieties and the C-specific residues F391, V518, E376 and V380 (numbering for testis ACE). Mutation of these residues decreased the affinity for the inhibitors 4- to 20-fold. T282, V379, E403, D453, and S516 did not contribute individually to C-domain-selective inhibitor binding. These data emphasise the importance of key S_1, S_2 and S_2' residues for inhibitor domain-selectivity.

The occurrence of two homologous active sites in one polypeptide chain is a rare phenomenon, thought to come about by gene duplication leading to the development of new activities (1,2). Human somatic angiotensin-converting enzyme (ACE)[1] is one such case, having two 55% identical domains, N and C, each containing an active site with a zinc binding motif, HEMGH (3,4). ACE is a key target for the treatment of hypertension because of its role in increasing blood pressure via the renin-angiotensin system (5,6).

Further examples of such two-domain enzymes are sucrase isomaltase (SI) and lactase-phlorizin hydrolase (LPH), both involved in sugar metabolism (7,8). In these enzymes the homologous domains, which are ~40% identical, bind similar but

distinct substrates. The domains of SI cleave related disaccharides sucrose and isomaltose, while LPH cleaves lactose and the glucoside phlorizin. In the case of LPH, mechanism-based domain-selective inhibitors were used to elucidate the activities of the two domains (9). Other examples of enzymes containing homologous active domains include adenylyl transferase of *Escherischia coli* (10), which catalyses the adenylylation and deadenylylation of glutamine synthetase, and protein disulphide isomerase (11), which catalyses disulphide bond oxidation, reduction and isomerisation.

In all of these examples, a few active site substitutions correspond to significantly different, but functionally complementary, substrate specificities for the homologous domains. ACE differs from these enzymes in that its activity is not highly specific, with each domain being able to cleave a range of peptide sequences, however the efficiency of cleavage varies between domains and some substrates are domain-specific (12-14). The N and C domains of ACE also differ in chloride dependence (4,15) and thermostability (16).

These differences between the domains make it appealing to develop an antihypertensive ACE inhibitor therapy that selectively targets one domain. The C domain has been shown to produce the potent vasopressor angiotensin II from angiotensin I more efficiently than the N domain, *in vivo* (17-19). Moreover, build-up of bradykinin, a substrate of both domains, has been implicated in a number of ACE inhibitor-related side effects (20-22). It has been suggested that a C-domain-selective ACE inhibitor would reduce blood pressure while allowing the N domain to maintain physiological bradykinin levels, resulting in more effective treatment of hypertension with reduced side effects (14,23).

Keto-ACE (hereafter referred to as kAP), an analogue of the tripeptide ACE inhibitor Bz-Phe-Gly-Pro, has been shown to be 26 to 34 times more selective for the C domain with a variety of substrates (24,25). Using kAP as a platform and making C-terminal substitutions, our group has synthesised a number of novel compounds. Two of these, (5S)-5-[(N-benzoyl)amino]-4-oxo-6-phenyl-hexanoyl-L-tryptophan (kAW) and (5S)-5-[(N-benzoyl)amino]-4-oxo-6-phenyl-hexanoyl-L-phenylalanine (kAF) (Figure 1), are highly C-selective (26).

In this study, we determine the crystal structures of human testis ACE (tACE), which is equivalent to the C domain of ACE, in complex with the novel domain-selective inhibitors kAW and kAF. In addition, we assess the effect of a number of active site mutations on inhibitor binding. This combined structural and kinetic approach allows the determination of the extent to which particular residues contribute to the C-domain-selectivity of these inhibitors.

EXPERIMENTAL PROCEDURES

tACE constructs- A fully N-glycosylated tACE mutant, tACEΔ36NJ, truncated after S625 and lacking the 36 *O*-glycosylated, N-terminal residues (27) was cloned into the *Bam*HI and *Eco*RI restriction sites of pcDNA3.1(+) (Invitrogen) for expression purposes. A minimally glycosylated form of this mutant, tACE-G13 (28), was used for crystallisation.

Inhibitors- Inhibitors kAP, kAW and kAF were synthesised previously (26,29). Stocks were made by dissolving inhibitors in a small volume of DMSO and then diluting in water or buffer, for enzyme assays.

Cloning and mutagenesis- Site-directed mutagenesis was performed on an *Sph*I-*Eco*RI fragment of the C domain coding region in pGEM11zf(+) (Promega) as described previously (30), converting selected active site residues to their N domain counterparts. Oligonucleotides containing the mutation of interest were

synthesised by Inqaba Biotech, South Africa, and reagents used for site-directed mutagenesis were supplied by Promega. Colonies were screened with appropriate restriction enzymes, and positive clones were confirmed by nucleotide sequencing. C domain mutants were constructed in pcDNA3.1+ for protein expression.

Heterologous expression and purification- tACE constructs were expressed in Chinese hamster ovary cells and purified by Sepharose-lisinopril affinity chromatography, as described previously (28,31). ACE activity was detected using a fluorogenic assay with substrates Hippuryl-Histidyl-Leucine (HHL; Sigma) or z-Phenyl-Histidyl-Leucine (z-FHL; Sigma) (32). Pooled fractions were dialysed against 2 L 0.5 mM Hepes (pH 7.5) overnight. tACE-G13 was concentrated to 1-5 mg/mL for crystallisation using an Amicon Ultra-4 spin column (Millipore). The purified protein's molecular weight and purity were assessed using SDS-PAGE.

Crystallisation- Crystals were grown by vapour diffusion at 16 °C in hanging drops comprising 2 µl protein (1.45 mg/mL) or protein/inhibitor solution mixed with 2 µl reservoir buffer over 1 ml reservoirs covered with 500 µl oil (3 parts paraffin, 2 parts silicon oil, Hampton) (33). Inhibitor-free tACE-G13 microcrystals grew within two weeks over reservoirs containing 10 mM sodium acetate (pH 4.7; Merck), 15% PEG 4000 (Fluka) and 10 µM $ZnSO_4$. Co-crystals of tACE-G13 with inhibitors were grown by streak-seeding these microcrystals into fresh drops containing inhibitor mixed with tACE-G13, over reservoirs containing a higher concentration of sodium acetate (30 mM) but otherwise identical to the reservoirs used for growing microcrystals. Diffracting crystals were grown within two weeks in drops containing 1 mM kAW with 1 mg/ml tACE-G13 and 1.25 mM kAF with 1.4 mg/ml tACE-G13.

Data collection and processing- Data were collected at 100 K from a single crystal per inhibitor at beamline BM14(UK) of the European Synchrotron Radiation Facility, Grenoble, France, using a wavelength of 1.033 Å.

Data were processed using *HKL2000* (34) and phasing was carried out with *EPMR 2.5* (35) using protein atoms from the crystal structure of tACE-G13 (Protein Data Bank (PDB) code 2IUL) as a model.

Model building was carried out using *O v9.0.7* (36), against $2F_{obs}$-F_{calc}, F_{obs}-F_{calc} and composite omit maps (simulated annealing with 5% omission), calculated using *CNS* (37). Waters were added using the water_pick protocol of *CNS* (37), followed by visual inspection. Ligand molecules were built using *PRODRG* (38) and *ARP/wARP* (39). Maximum likelihood minimization and simulated annealing of selected residues was carried out using *CNS* (37). *PROCHECK* and *SFCHECK* from the *CCP4* programme suite (39), and *Molprobity* (40) were used for model validation. Hydrogen bonds and close contacts were identified using *HBplus* (41). Figures were generated using *Pymol v0.99* (DeLano Scientific, Palo Alto, CA, U.S.A).

Substrate hydrolysis- The hydrolysis of the fluorogenic peptide Abz-FRK(Dnp)P-OH (a kind gift from Dr Adriana Carmona) was characterised in a continuous assay at 25 °C under initial rate conditions, by measuring fluorescence at $\lambda_{ex} = 320$ nm and $\lambda_{em} = 420$ nm. A substrate concentration range of 0 – 12 µM was added to the reaction mixture of 50 mM Hepes buffer, pH 6.8, containing 200 mM NaCl, 10 µM $ZnCl_2$ and an enzyme concentration of 0.2 nM. Kinetic constants were calculated using the direct linear plot method (42,43).

Inhibition kinetics- ACE constructs (2 nM) were incubated with an appropriate concentration range of inhibitor in 50 mM Hepes buffer, pH 6.8, containing 200 mM NaCl and 10 µM $ZnCl_2$, at ambient temperature for 60 min. 20 µL of the enzyme-inhibitor mixture were added to a reaction mixture of 50 mM Hepes buffer,

pH 6.8, containing 200 mM NaCl, 10 μM $ZnCl_2$ and 4 or 8 μM Abz-FRK(Dnp)P-OH, to make up a total volume of 300 μL, in triplicate. Residual enzyme activity was monitored by determining fluorescence at $\lambda_{ex} = 320$ nm and $\lambda_{em} = 420$ nm, at 25°C. Inhibition constants were calculated using the Dixon method (44).

RESULTS

Co-crystal structures of tACE-G13 with novel inhibitors- The C-domain-selective compounds kAW and kAF were derived from the moderately domain-selective ketone ACE inhibitor, kAP, by substitution of the P_2' group (Figure 1) (26,29). In order to understand the mechanism of domain-selectivity, we determined co-crystal structures of these two novel inhibitors with tACE, which is identical in sequence to the C domain. To facilitate crystallisation we used a minimally glycosylated construct, tACE-G13, containing only two intact glycosylation sites (28,45). Structures were solved by molecular replacement and refined to final crystallographic *R*-factors of 20.0% in both cases. Data processing and refinement statistics are presented in Table 1.

In both the kAW and kAF co-crystal structures, the protein component is in an identical conformation to the inhibitor-free tACE-G13 structure (PDB code 2IUL), as evidenced by all-atom r.m.s. deviations of 0.61 Å and 0.67 Å, respectively. Overall, this structure is mostly alpha helical and ellipsoid in shape, containing two chloride ions, with the zinc ion buried deep in the active site cleft. As in the unliganded structure, glycan residues are present at the two intact glycosylation sites in tACE-G13, N72 and N109, where there is evidence of glycan-mediated stabilisation of the lid helices, as was observed previously (45).

One inhibitor molecule is bound to the active site of each tACE-G13 molecule, in an extended conformation (Figure 2) similar to that seen for the phosphinic peptide inhibitor, RXPA380 (46). Weak density around the P_2' side chain of kAF indicates a degree of flexibility, which is also evidenced by the higher B-factors of these atoms. These inhibitors have a ketone zinc-binding moiety analogous to the carbonyl oxygen of the scissile peptide bond, however the density maps surprisingly indicate the stabilisation of a hydrated geminal diolate (*gem*-diol) transition state at the active site, in which two oxygens are attached to the carbonyl carbon analogue (Figure 3). The zinc co-ordination distances for the *gem*-diol oxygens are 2.23 and 2.42 Å for kAW and 2.41 and 2.52 Å for kAF (see supplementary Table 1).

kAW makes nine polar contacts (potential hydrogen bonds and ionic interactions) with seven protein side chains, while kAF makes nine polar contacts with six protein side chains, the difference being small changes in the orientations of Q281 and E384 (Supplementary Table 1). Additional hydrogen bonds are made with four water molecules, whose positions are conserved between the structures, as well as with a glycerol molecule derived from the cryo-protectant in the kAW structure. The conformation of the Trp moiety of kAW does not appear to be effected by the presence of the glycerol, since it takes the same conformation as the P_2' Trp of RXPA380, which does not have a glycerol in this position. Moreover, no glycerol-like density was present at this position in composite omit maps generated using the RXPA380 structure factors from the PDB. These solvent molecules form part of a hydrogen bonding network connecting the inhibitor to other active site residues and to the bulk solvent (Figure 4).

Close contacts are also observed between the nonpolar moieties of the inhibitors and 17 active site residues as well as a number of water molecules (11 and 13 for kAW and kAF, respectively), and glycerol and acetate ions in the case of

kAW (Supplementary Table 2). Some of these active site residues (T282, S355, V379, V380, H387, F457, F512, V518 and F527) could contribute favourably to binding entropy by the hydrophobic effect, while others (Q281, H353, A354, E384, D415 and the water molecules) are polar or charged groups, and some (H383 and Y523) make both favourable and unfavourable contacts.

Nine residues that interact with kAW and kAF directly, or indirectly via solvent atoms, are replaced with a different amino acid in the N domain and may thus contribute to the domain-selectivity of these ketone inhibitors (Figure 2).

Kinetic analysis of active site ACE mutants- In order to investigate the role of the nine active site residues that differ between domains, a series of tACE mutants was generated in which these residues were converted to their N domain counterparts (Table 2).

Using the substrate Abz-FRK(Dnp)P-OH under the conditions described, kAW shows strong C-selectivity, having almost 1300-fold greater affinity for the C domain than for the N domain (Table 2). The greatest decreases in affinity for kAW observed among the mutants were with the S_1 pocket V518T and S_2 F391Y mutants (Table 2, Figure 5). The mutants E403R and S516N in these pockets displayed no change in affinity for kAW relative to wild-type tACE. In the S_2' pocket mutants E376D and V380T demonstrated similar increases in K_i, while the V379S mutation caused an increase in affinity for the inhibitor. The T282S and D453E mutants displayed similar K_i's to that of the wild-type C domain (Table 2, Figure 5).

The selectivity of kAW is only two-fold lower than that of RXPA380, which shares a P_2' Trp, P_1 Phe and P_2 benzyl moiety with kAW (14). However, the affinity of kAW for the C domain is approximately 230-fold lower than that of RXPA380, which has a C domain K_i of 3 nM (14). These data for RXPA380 were determined under the same reaction conditions as those used here, but using a different fluorogenic peptide substrate, Mca-ASDK-DpaOH (14). No change in selectivity was observed when using Abz-FRK(Dnp)P-OH (data not shown).

kAF was also shown to be highly C-domain-selective under these conditions, with a C domain K_i of 0.83 μM (Table 2). No inhibition of the N domain was observed using up to 500 μM kAF and at a range of chloride concentrations. Due to the limited solubility of this compound, inhibition could not be tested at higher concentrations. The V518T mutation again displayed the greatest decrease in affinity for kAF, while small decreases in affinity were also observed for the F391, E376D and V380T mutants (Table 2, Figure 5).

In order to understand the domain-selectivity of the parent compound, kAP, we tested the effects of the F391Y, S516N and V518T mutations in the S_1 and S_2 pockets on the binding affinity of this inhibitor. kAP had a 30-fold higher affinity for the C domain than for the N domain (Table 2). Notably, the F391Y mutation in the S_2 pocket caused an 8-fold decrease in affinity relative to the wild-type C domain, while the S_1 V518T mutant had a K_i similar to that observed for the N domain. The S516N substitution resulted in comparable binding affinity to that of the wild-type C domain (Table 2).

DISCUSSION

As an enzyme with two homologous catalytically active domains, ACE is an intriguing drug target. Differences in the activity profiles of the domains, particularly with reference to the cleavage of angiotensin I and bradykinin, render it desirable to design selective inhibitors to the C domain, however the high homology between the domains makes this a complex problem. That domain-selective inhibition is possible is evident from the selectivity of some known substrates and inhibitors, and from the recent identification of the

strongly domain-selective phosphinic peptide inhibitors RXPA380 and RXP407 (13,14,24,47,48). We have recently developed new C-selective inhibitors kAW and kAF, based on the moderately domain-selective inhibitor kAP (26,29).

In zinc metalloproteases, peptide bond hydrolysis is thought to occur via a tetrahedral transition state which forms upon nucleophilic attack of the carbonyl carbon by an activated zinc-bound water molecule at the active site (49,50). This is followed by protonation of the scissile amide nitrogen and subsequent peptide bond cleavage. We found that the replacement of the scissile peptide bond nitrogen with a carbon atom in these inhibitors (Figure 1) led to the trapping of the tetrahedral *gem*-diol transition state at the active site (Figure 3).

The hydrated transition state has been observed previously in the structures of carboxypeptidase A (CPA) co-crystallised with ketomethylene inhibitors 5-benzamido-2-benzyl-4-oxopentanoic acid (BOP) and 5-amino-(N-t-butoxycarbonyl)-2-benzyl-4-oxo-6-phenylhexanoic acid ((51,52); structures not available in the PDB). In these structures, the hydrated form of the ketone was observed at the enzyme active site despite the fact that this form does not occur (<1%) in aqueous solution. Active site stabilisation of *gem*-diols is commonly observed in co-crystal structures with aldehyde inhibitors and inhibitors containing activated electrophilic ketones, with examples from the PDB including co-crystal structures of leucine aminopeptidase (code 1LAN) (53), CPA (code 1CBX) (54), endothiapepsin (code 1OD1) (55) and human GAR transformylase (code 1NJS) (56). However, the hydration of a ketone in the absence of an activating group, as seen here, has to the best of our knowledge only been reported in two other crystal structures, those of CPA mentioned above.

The *gem*-diol oxygens in the structures of kAW and kAF at the active site of tACE-G13 have partial negative charges and interact with the active site zinc ion (Supplementary Table 1), in similar positions to the carboxyl oxygens of tACE-bound lisinopril and enalaprilat (PDB codes 1O86 and 1UZE). Y523 and E384, which has been proposed to shuttle a proton between the attacking nucleophilic water and the leaving amide, also make hydrogen bonds with the *gem*-diol (Figure 3). In the structure of CPA bound to BOP the *gem*-diol has longer zinc-coordinating distances (2.51 and 2.91 Å; (51)) than those observed here, however it is clear from the overlap of the *gem*-diol oxygens of kAW and kAF with the zinc-binding oxygens of other inhibitors that these positions are characteristic of the ACE active site.

Thus, we have shown that in the absence of a leaving amide nitrogen, the catalytic residues of ACE are able to stabilise the tetrahedral transition state. This stabilisation probably plays a transient role in overcoming the energy barrier between substrate and products during catalysis.

The ketone zinc-binding moiety of kAP and its derivative inhibitors has lower affinity for the active site zinc ion than the carboxyl, sulphydryl and phosphinyl groups of other previously-studied inhibitors (46,57,58). This results in a situation similar to the binding of a natural substrate, with the affinity of the inhibitor depending to a great extent on the contributions of other groups. These inhibitors are thus suitable tools for investigating the contributions of moieties other than the zinc ion to inhibitor binding.

kAP has a 30-fold higher affinity for the C domain than for the N domain (Table 2). This domain-selectivity does not reside in the P_1' position since there is no P_1' side chain in kAP, or in the P_2' moiety, since Pro at this position does not favour binding to one domain (59). Moreover, from the co-crystal structures of tACE and the N domain with lisinopril (PDB codes 1O86 and 2C6N, respectively), one can see that a P_2' Pro stacks against Y523 and its N

domain equivalent, Y501, making no interactions with residues that differ between domains. Thus, the modest selectivity of kAP must be attributed to its P_1 and P_2 moieties, which are retained in kAW and kAF.

The S_1 and S_2 pockets of the N and C domains differ in four residues that lie within 6 Å of kAW and kAF, of which S516, V518 (S_1 pocket) and F391 (S_2 pocket) make hydrophobic contacts with the inhibitor P_1 and P_2 side chains. These are replaced by the charged and polar residues N494, T496 and Y396, respectively, in the N domain (Figure 2).

Of these substitutions, F391Y and V518T were shown by our kinetic analysis to be important for the selectivity of kAP, reducing affinity 8-fold and 26-fold, respectively. The F391Y mutation introduces a hydroxyl group into the S_2' pocket, which would result in steric hindrance of the binding of the P_2 carboxybenzoyl moiety of kAP. This substitution has been identified as potentially contributing to the C-domain-selectivity of the phosphinic peptide inhibitor RXPA380, which also has a benzoyl group at this position (46).

The effect of the V518T mutation is not steric, however the association of the P_1 Phe side chain with T496 of the N domain is probably made unfavourable by the displacement of any water molecules hydrogen-bonded to the hydroxyl moiety, which would be necessary for binding of the bulky phenyl group. Displacement of waters around V518, in contrast, would result in a favourable increase in entropy. This S_1 subsite alteration may explain the moderate C-selectivity of lisinopril and enalaprilat, which also have P_1 phenyl groups, as well as the observation that a bulky P_1 moiety confers C-domain-selectivity (60).

The C-domain-selectivity of kAP is thus due to sterically and entropically favourable hydrophobic interactions between the bulky hydrophobic P_1 and P_2 benzoyl moieties and F391 and V518, which are replaced with polar groups in the N domain. These substitutions also had a dramatic effect on the binding of kAW and kAF (Figure 5).

The single residue replacement of Trp for the P_2' Pro of kAP reduces the affinity of kAW for the N domain dramatically, making it approximately 40-fold more C-selective than kAP (Table 2). Although the mutation of V518 to Thr did result in a large decrease in affinity for kAW (14-fold), the K_i did not equal that of the N domain, as occurred with kAP, suggesting that other factors play a role in the selectivity of this inhibitor (Table 2).

In the crystal structure, the P_2' Trp indole nitrogen of kAW makes a hydrogen bond to a highly ordered glycerol molecule, which is hydrogen-bonded to T282, Q281, N277, E376, and a network of water molecules in the S_2' subsite (Figure 4). Since this glycerol is derived from the cryo-protectant used for data collection, crystal growth must have occurred in its absence, suggesting that a water molecule was the original acceptor for the Trp hydrogen bond.

In addition to this hydrogen bond, the Trp interacts closely with a hydrophobic patch created by T282, F457, F527 and Y523 on one side, and with H383, V379, V380 and a water molecule on the other side (Figure 2, Supplementary Table 2). Four of these ten residues, T282, E376, V379 and V380, are substituted in the N domain (Figure 2).

From the kinetic data, the enhanced selectivity of kAW can be attributed to E376 and V380, with the E376D and V380T mutations demonstrating approximately 4-fold decreases in affinity (Table 2, Figure 5). Since E376 interacts with the Trp indole N indirectly via the glycerol (or water network), the effect of the E376D mutation is probably a result of the weakening of this interaction due to the lengthening of the distance between the negatively charged group and the indole nitrogen. In the N domain structure (PDB code 2C6F), this side chain is turned away

from the active site. The V380T mutation amounts to the introduction of a polar group in close proximity to the Trp ring, similar to the S_1 V518T mutation. The overall low affinity of the N domain for kAW is thus probably the combined effect of these two S_2' mutations together with those in the P_1 and P_2 pockets.

Interestingly, the V379S mutant actually displays considerably higher affinity for kAW than the wild-type. This effect is not readily explained by the crystal structure but may be due to a rearrangement of side chains in the S_2' pocket, possibly leading to the formation of a water-mediated hydrogen bond between the Ser and the Trp indole N. This is supported by the observation that this mutant did not have increased affinity for kAF, which lacks a P_2' hydrogen-bonding moiety (Table 2). The C-domain-like affinity of the double VV/ST mutant can be explained as the net result of the affinity-decreasing effect of the V380T mutation and the large affinity-increasing effect of the V379S mutation.

kAF is also strongly C-selective, exhibiting no inhibition of the N domain up to 500 μM inhibitor. Like the Trp of kAW, the P_2' sidechain contacts V379 and V380, as well as the conserved hydrophobic residues H383, Y523, F527 and F457, but many of these contacts are more distant than those of kAW because of the smaller size of the Phe side chain (Figures 2 and 4, Supplementary Table 2). Moreover, the aromatic ring is unable to make a hydrogen bond to the water network, and in place of the glycerol moiety seen in the kAW structure, a ring of ordered waters is present, interacting with the polar residues beyond (Figure 4; Supplementary Table 2). The effect of the smaller size of this side chain, together with the lack of a hydrogen-bonding moiety, is evident in the disorder of this group in the crystal structure, in comparison with the Trp of kAW, which displays lower B-factors and clearer density (Figures 2 and 4).

As with kAW, the mutation of V380 and E376 to T358 and D354, respectively, contributes to the low affinity of this inhibitor for the N domain (Table 2). The selectivity of kAF is thus the combined effect of these mutations, together with the clash of the bulky hydrophobic P_1 and P_2 moieties with F391 and V518.

From this investigation, it can be concluded that the C-domain-selectivity of the kAP P_2' derivatives kAW and kAF results from interactions between their bulky hydrophobic side chain moieties and the C-domain-specific residues F391 (P_2), V518 (P_1), E376 and V380 (P_2'). Candidate substituted residues T282, V379, E403, D453, and S516 do not individually contribute favourably to the binding of these inhibitors.

It should be noted that the binding affinities of these inhibitors are not of the same order of magnitude as those of currently available drugs such as lisinopril which have K_i values in the low nanomolar range (60). Replacing the zinc-binding ketone with another group having higher affinity for the zinc ion would probably improve the overall affinity, however this would be a non-selective effect. An approach more suited to domain-selective inhibitor design would be to increase the affinity and selectivity of the inhibitor side chain moieties while maintaining the ketone group. Attention should be paid to increasing the size of the P_2' moiety, since it is evident from the structures that a P_2' Trp or Phe does not fill the pocket sufficiently to take full advantage of all the potential domain-specific contacts. Novel domain-selective effects might also be achieved by introducing a P_1' side chain into these inhibitors.

These data will likely aid in the process of structure-based drug design of domain-selective ACE inhibitors suitable for clinical use, ideally eliminating some of the side effects associated with current ACE inhibitor therapy.

ACKNOWLEDGEMENTS

We thank Muhammed Sayed and Itai Chitapi for helping with the data collection, Ravi Acharya and Mario Ehlers for helpful comments on the manuscript, and David Christianson for the CPA co-ordinates. This work was supported by the Wellcome Trust, U.K. (Senior International Research Fellowship 070060 to E.D.S.), the National Research Foundation, South Africa, the Stella and Paul Loewenstein Charitable and Educational Trust, the Ernst and Ethel Erikson Trust and the University of Cape Town.

REFERENCES

1. Bergdoll, M., Eltis, L. D., Cameron, A. D., Dumas, P., and Bolin, J. T. (1998) *Protein Sci.* **7,** 1661-1670
2. Diaz-Mejia, J. J., Perez-Rueda, E., and Segovia, L. (2007) *Genome Biol.* **8,** R26
3. Soubrier, F., Alhenc-Gelas, F., Hubert, C., Allegrini, J., John, M., Tregear, G., and Corvol, P. (1988) *Proc.Natl.Acad.Sci.U.S.A* **85,** 9386-9390
4. Wei, L., Alhenc-Gelas, F., Corvol, P., and Clauser, E. (1991) *J.Biol.Chem.* **266,** 9002-9008
5. Erdos, E. G. (1990) *Hypertension* **16,** 363-370
6. Zaman, M. A., Oparil, S., and Calhoun, D. A. (2002) *Nat.Rev.Drug Discov.* **1,** 621-636
7. Hunziker, W., Spiess, M., Semenza, G., and Lodish, H. F. (1986) *Cell* **46,** 227-234
8. Mantei, N., Villa, M., Enzler, T., Wacker, H., Boll, W., James, P., Hunziker, W., and Semenza, G. (1988) *EMBO J.* **7,** 2705-2713
9. Arribas, J. C., Herrero, A. G., Martin-Lomas, M., Canada, F. J., He, S., and Withers, S. G. (2000) *Eur.J.Biochem.* **267,** 6996-7005
10. Anderson, W. B. and Stadtman, E. R. (1970) *Biochem.Biophys.Res.Commun.* **41,** 704-709
11. Ellgaard, L. and Ruddock, L. W. (2005) *EMBO Rep.* **6,** 28-32
12. Rousseau, A., Michaud, A., Chauvet, M. T., Lenfant, M., and Corvol, P. (1995) *J.Biol.Chem.* **270,** 3656-3661
13. Dive, V., Cotton, J., Yiotakis, A., Michaud, A., Vassiliou, S., Jiracek, J., Vazeux, G., Chauvet, M. T., Cuniasse, P., and Corvol, P. (1999) *Proc.Natl.Acad.Sci.U.S.A* **96,** 4330-4335
14. Georgiadis, D., Beau, F., Czarny, B., Cotton, J., Yiotakis, A., and Dive, V. (2003) *Circ.Res.* **93,** 148-154
15. Wei, L., Clauser, E., Alhenc-Gelas, F., and Corvol, P. (1992) *J.Biol.Chem.* **267,** 13398-13405
16. Voronov, S., Zueva, N., Orlov, V., Arutyunyan, A., and Kost, O. (2002) *FEBS Lett.* **522,** 77-82
17. Junot, C., Gonzales, M. F., Ezan, E., Cotton, J., Vazeux, G., Michaud, A., Azizi, M., Vassiliou, S., Yiotakis, A., Corvol, P., and Dive, V. (2001) *J.Pharmacol.Exp.Ther.* **297,** 606-611
18. Fuchs, S., Xiao, H. D., Cole, J. M., Adams, J. W., Frenzel, K., Michaud, A., Zhao, H., Keshelava, G., Capecchi, M. R., Corvol, P., and Bernstein, K. E. (2004) *J.Biol.Chem.* **279,** 15946-15953
19. Cotton, J., Hayashi, M. A. F., Cuniasse, P., Vazeux, G., Ianzer, D., De Camargo, A. C. M., and Dive, V. (2002) *Biochemistry* **41,** 6065-6071
20. Nussberger, J., Cugno, M., Amstutz, C., Cicardi, M., Pellacani, A., and Agostoni, A. (1998) *Lancet* **351,** 1693-1697

21. Beltrami, L., Zingale, L. C., Carugo, S., and Cicardi, M. (2006) *Expert.Opin.Drug Saf* **5,** 643-649
22. Ehlers, M. R. (2006) *Expert.Opin.Drug Saf* **5,** 739-740
23. Dickstein, K. and Kjekshus, J. (2002) *Lancet* **360,** 752-760
24. Deddish, P. A., Marcic, B., Jackman, H. L., Wang, H. Z., Skidgel, R. A., and Erdos, E. G. (1998) *Hypertension* **31,** 912-917
25. Almquist, R. G., Chao, W. R., Ellis, M. E., and Johnson, H. L. (1980) *J.Med.Chem.* **23,** 1392-1398
26. Nchinda, A. T., Chibale, K., Redelinghuys, P., and Sturrock, E. D. (2006) *Bioorg.Med.Chem.Lett.* **16,** 4612-4615
27. Yu, X. C., Sturrock, E. D., Wu, Z., Biemann, K., Ehlers, M. R., and Riordan, J. F. (1997) *J.Biol.Chem.* **272,** 3511-3519
28. Gordon, K., Redelinghuys, P., Schwager, S. L., Ehlers, M. R., Papageorgiou, A. C., Natesh, R., Acharya, K. R., and Sturrock, E. D. (2003) *Biochem.J.* **371,** 437-442
29. Redelinghuys, P., Nchinda, A. T., Chibale, K., and Sturrock, E. D. (2006) *Biol.Chem.* **387,** 461-466
30. Papworth, C., Bauer, J. C., Braman, J., and Wright, D. A. (1996) *Strategies* **9,** 3-4
31. Ehlers, M. R., Chen, Y. N., and Riordan, J. F. (1991) *Protein Expr.Purif.* **2,** 1-9
32. Friedland, J. and Silverstein, E. (1976) *Am.J.Clin.Pathol.* **66,** 416-424
33. Chayen, N. E. (1997) *Structure.* **5,** 1269-1274
34. Otwinowski, Z. and Minor, W. (1997) [20] Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology*, Academic Press,
35. Kissinger, C. R., Gehlhaar, D. K., and Fogel, D. B. (1999) *Acta.Cryst.* **D55 (Pt 2),** 484-491
36. Jones, T. A. (1978) *J. Appl. Crystallogr.* **11,** 268-272
37. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr.D Biol.Crystallogr.* **54 (Pt 5),** 905-921
38. Schuttelkopf, A. W. and van Aalten, D. M. (2004) *Acta Crystallogr.D Biol.Crystallogr.* **60,** 1355-1363
39. Collaborative Computational Project, N. 4. (1994) *Acta Crystallogr.D Biol.Crystallogr.* **50,** 760-763
40. Davis, I. W., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2004) *Nucleic Acids Res.* **32,** W615-W619
41. McDonald, I. K. and Thornton, J. M. (1994) *J.Mol.Biol.* **238,** 777-793
42. Eisenthal, R. and Cornish-Bowden, A. (1974) *Biochem.J.* **139,** 715-720
43. Carmona, A. K., Schwager, S. L., Juliano, M. A., Juliano, L., and Sturrock, E. D. (2006) *Nat.Protoc.* **1,** 1971-1976
44. Dixon, M. (1953) *Biochem.J.* **55,** 170-171
45. Watermeyer, J. M., Sewell, B. T., Schwager, S. L., Natesh, R., Corradi, H. R., Acharya, K. R., and Sturrock, E. D. (2006) *Biochemistry* **45,** 12654-12663
46. Corradi, H. R., Chitapi, I., Sewell, B. T., Georgiadis, D., Dive, V., Sturrock, E. D., and Acharya, K. R. (2007) *Biochemistry* **46,** 5473-5478
47. Hayashi, M. A. F. and Camargo, A. C. M. (2005) *Toxicon* **45,** 1163-1170
48. Georgiadis, D., Cuniasse, P., Cotton, J., Yiotakis, A., and Dive, V. (2004) *Biochemistry* **43,** 8048-8054
49. Lipscomb, W. N. and Strater, N. (1996) *Chemical Reviews* **96,** 2375-2433
50. Pelmenschikov, V., Blomberg, M. R., and Siegbahn, P. E. (2002) *J.Biol.Inorg.Chem.* **7,** 284-298

51. Christianson, D. W., David, P. R., and Lipscomb, W. N. (1987) *PNAS* **84,** 1512-1515
52. Shoham, G., Christianson, D. W., and Oren, D. A. (1988) *Proc.Natl.Acad.Sci.U.S.A* **85,** 684-688
53. Strater, N. and Lipscomb, W. N. (1995) *Biochemistry* **34,** 9200-9210
54. Mangani, S., Carloni, P., and Orioli, P. (1992) *J.Mol.Biol* **223,** 573-578
55. Coates, L., Erskine, P. T., Mall, S., Williams, P. A., Gill, R. S., Wood, S. P., and Cooper, J. B. (2003) *Acta Crystallographica Section D-Biological Crystallography* **59,** 978-981
56. Zhang, Y., Desharnais, J., Marsilje, T. H., Li, C., Hedrick, M. P., Gooljarsingh, L. T., Tavassoli, A., Benkovic, S. J., Olson, A. J., Boger, D. L., and Wilson, I. A. (2003) *Biochemistry* **42,** 6043-6056
57. Natesh, R., Schwager, S. L., Sturrock, E. D., and Acharya, K. R. (2003) *Nature* **421,** 551-554
58. Natesh, R., Schwager, S. L., Evans, H. R., Sturrock, E. D., and Acharya, K. R. (2004) *Biochemistry* **43,** 8718-8724
59. Michaud, A., Chauvet, M. T., and Corvol, P. (1999) *Biochemical Pharmacology* **57,** 611-618
60. Acharya, K. R., Sturrock, E. D., Riordan, J. F., and Ehlers, M. R. (2003) *Nat.Rev.Drug Discov.* **2,** 891-902

FOOTNOTES

[1] ABBREVIATIONS: ACE, angiotensin-converting enzyme; SI, sucrase-isomaltase; LPH, lactase-phlorizin hydrolase; kAP, keto-ACE; kAW, (5S)-5-[(N-benzoyl)amino]-4-oxo-6-phenyl-hexanoyl-L-tryptophan; kAF, (5S)-5-[(N-benzoyl)amino]-4-oxo-6-phenyl-hexanoyl-L-phenylalanine; tACE, testis angiotensin-converting enzyme; tACE-G13, minimally glycosylated tACE mutant; DMSO, dimethyl sulphoxide; HHL, Hippuryl-Histidyl-Leucine; z-FHL, z-Phenyl-Histidyl-Leucine; PDB, Protein Data Bank; CPA, carboxypeptidase A; BOP, 5-amino-(N-t-butoxycarbonyl)-2-benzyl-4-oxo-6-phenylhexanoic acid.

FIGURE LEGENDS

Figure 1. Chemical structures of kAP and its derivatives kAW and kAF, showing residue positions relative to the zinc-binding group (P_2, P_1, P_1', P_2').

Figure 2. Stick representation of kAW (A) and kAF (B) in the active site, showing the first F_{obs}-F_{calc} difference density map (purple mesh) into which the inhibitor model was built, contoured at a level of 3 σ. The position of the active site zinc ion (ZN2) is shown as a green non-bonded sphere, indicated with an arrow, and zinc-binding residues are shown as black lines in the foreground. Clear density for a tetrahedral *gem*-diol hydrated form of the zinc-binding ketone moiety is evident. Inhibitor-contacting residues that differ between domains are shown in yellow (tACE residues from the co-crystal structure) and purple (N domain residues from the aligned N domain structure, PDB code 2C6F). Residue labels are given as tACE/N domain, and the P_1, P_2 and P_2' moieties are labeled. In A, the active site-associated glycerol (GOL) and acetate (ACT) molecules are also shown as yellow sticks in F_{obs}-F_{calc} difference density.

Figure 3. Stabilisation of the hydrated transition state in the active site. The ketone inhibitor kAW (blue sticks) in F_{obs}-F_{calc} density (purple mesh, contoured at 3 σ) has been orientated to show the *gem*-diol moiety coordinating the active zinc ion (grey sphere). The zinc-coordinating residues and those involved in stabilising the intermediate are labeled and shown as yellow sticks.

Figure 4. Close-up views of the S_2' pocket with kAW (A) and kAF (B), showing the solvent network. The inhibitors and glycerol are shown as yellow sticks in F_{obs}-F_{calc} difference density (purple mesh), contoured at 3 σ. The active site zinc ion is shown as a green nonbonded sphere, water molecules as light blue nonbonded spheres, and solvent hydrogen bonds as blue dashed lines.

Figure 5. Comparison of the relative binding affinity of tACE active site mutants for ketone inhibitors kAW (black bars) and kAF (grey bars), with that of wild type tACE (C domain). Values above the line represent an N-domain-like decrease in affinity relative to that of tACE. The N domain K_i was 1258- and >602-fold higher than that of tACE for kAW and kAF, respectively.

Table 1. X-ray data collection and processing statistics.

Data set	**kAW**	**kAF**
Resolution (Å)[a]	43-2.18 (2.26-2.18)	43-2.17 (2.25-2.17)
No. reflections	130768	158385
Redundancy [a]	3.8 (3.2)	4.8 (4.1)
Completeness (%)[a]	95.0 (95.2)	88.3 (80.2)
$\langle I/\sigma(I)\rangle$ [a]	10.87 (2.37)	14.36 (2.61)
No. molecules per asymmetric unit	1	1
Space group	$P2_12_12_1$	$P2_12_12_1$
a (Å)	59.85	60.38
b (Å)	85.04	85.07
c (Å)	134.92	135.51
R_{merge} (%) [a, b]	10.9 (42.1)	8.7 (39.2)
No. reflections in working set (test set)	31173 (1753)	30340 (1459)
R_{cryst} (%) [a,c]	20.0 (26.6)	20.0 (25.0)
R_{free} (%) [a,d]	23.9 (29.7)	24.3 (26.4)
Mean B factor (A^2) [min. - max.]	21.6 [2.8-69.0]	22. 7 [5.9-72.5]
No. solvent atoms	254	230
No. protein atoms [e]	4750	4777
No. metal ions	3	3
No. inhibitor atoms	39	36
No. glycan atoms	67	67
r.m.s. deviations from ideal stereochemistry		
bonds (Å)	0.01	0.01
angles	1.4	1.3
dihedrals	22.2	21.2
improper angles	1.2	1.0
Ramachandran plot, % residues in favoured regions	97.40	96.23

[a] Values in parentheses refer to the highest resolution shell; [b] $R_{merge} = [\Sigma_h\Sigma_j |I_j(h) - \langle I(h)\rangle|] / \Sigma_h\Sigma_j |I_j(h)|$, where $I_i(h)$ and $\langle I(h)\rangle$ are the ith and the mean measurements of the intensity of reflection h, respectively. [c] $R_{cryst} = \sum_h |F_o - F_c| / \sum_h F_o$, where F_o and F_c are the observed and calculated structure factor amplitudes of reflection h of the working set, respectively; [d] R_{free} is equal to R_{cryst} for h belonging to the test set of reflections; [e] excluding duplicated atoms in alternate conformations.

Table 2. Kinetics of inhibition by kAW and kAP, of tACE mutants having active site residues mutated to their N domain counterparts. Inhibition constants were determined using the Dixon method (44) with fluorogenic peptide substrate Abz-FRK(Dnp)P-OH.

Pocket	tACE mutant construct [a]	kAW K_i (μM)	kAF K_i (μM)	kAP K_i (μM)
	tACE	0.679	0.83	0.05
S_2'	T282S(260)	0.920	-	-
S_2'	E376D(354)	2.33	2.4	-
S_2'	V379S(357)	0.064	0.81	-
S_2'	V380T(358)	2.65	1.9	-
S_2'	V379S/V380T	0.870	-	-
S_2'	D453E(431)	0.618	-	-
S_1	S516N(494)	0.497	-	0.012
S_1	V518T(496)	9.74	18.1	1.3
S_2	F391Y(369)	4.75	3.9	0.43
S_2	E403R(381)	0.24	-	-
	N-dom	854.2	>500	1.5

[a] tACE residue and number, followed by the corresponding N domain residue to which it was mutated, with N domain numbering in brackets.

Figure 1.

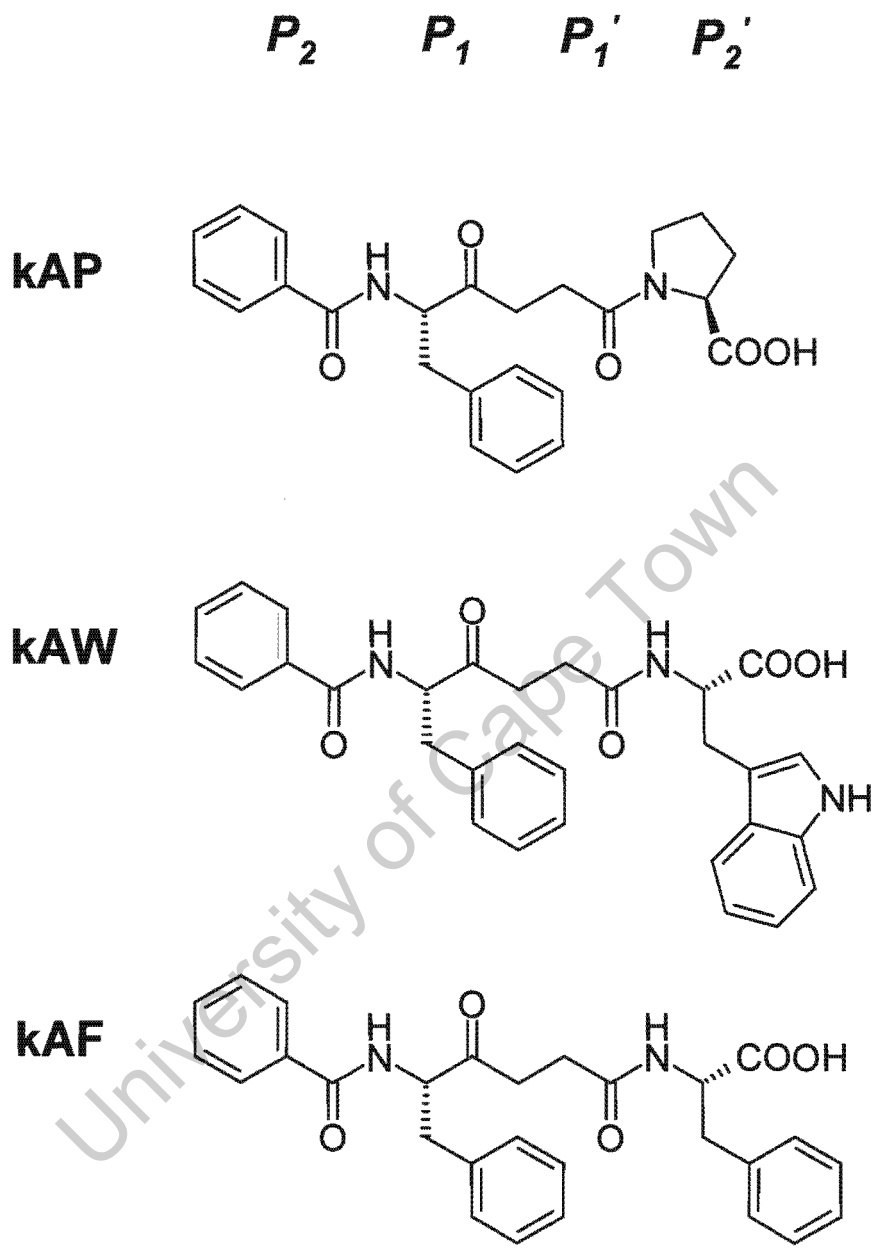


Figure 2.

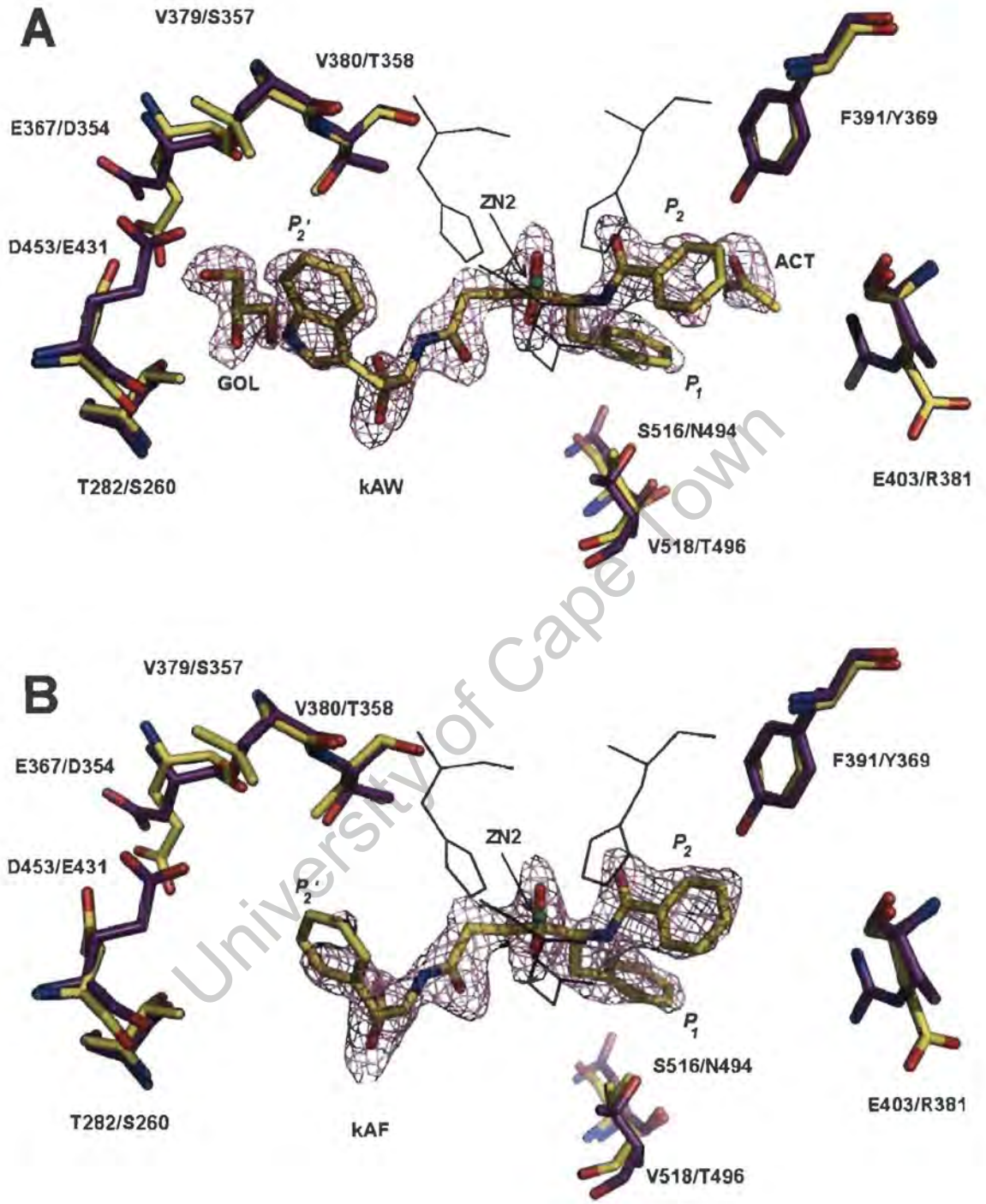


Figure 3.

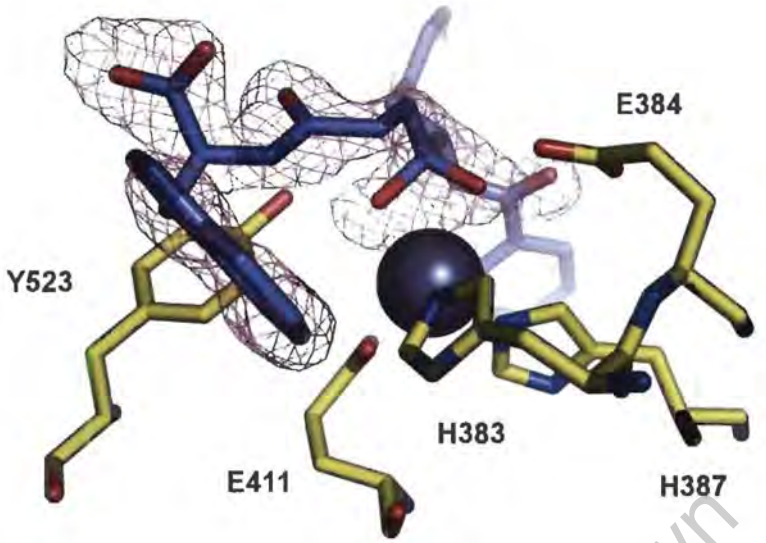


Figure 4.

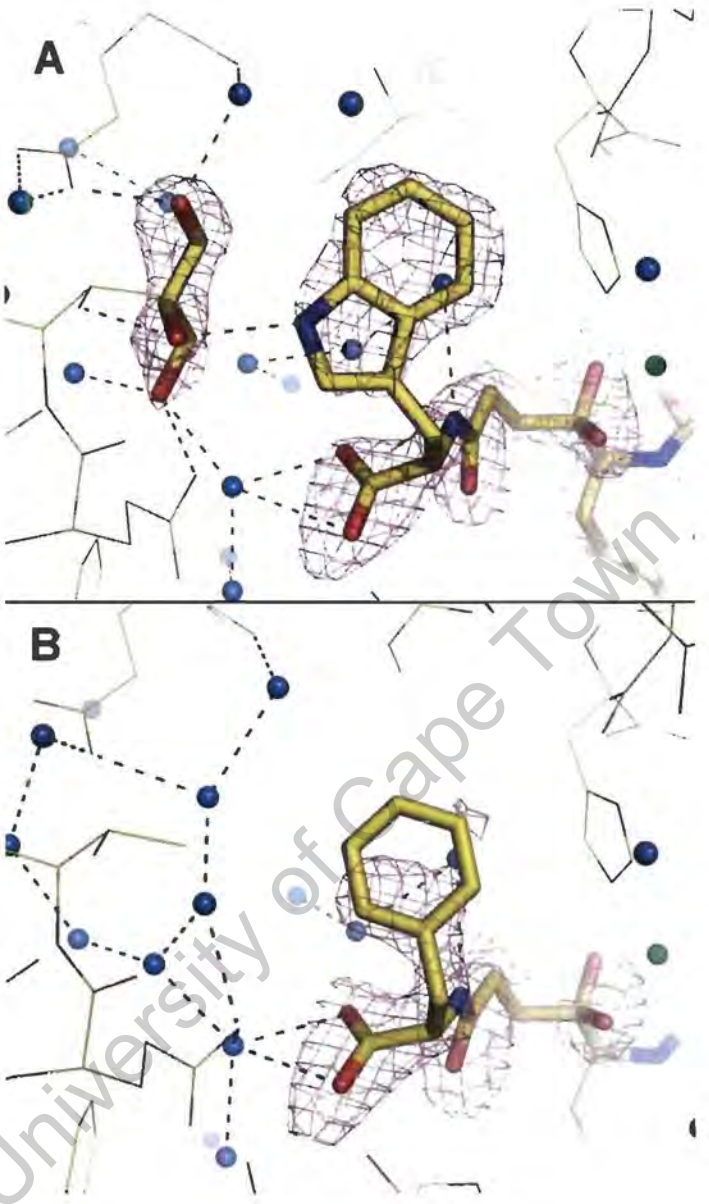


Figure 5.