doi:10.1016/j.jmb.2006.01.048 J. Mol. Biol. (2006) 357, 964–974

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Crystal Structure of the N Domain of Human Somatic Angiotensin I-converting Enzyme Provides a Structural Basis for Domain-specific Inhibitor Design

Hazel R. Corradi[1], Sylva L. U. Schwager[2], Aloysius T. Nchinda[2] Edward D. Sturrock[2]* and K. Ravi Acharya[1]*

[1]*Department of Biology and Biochemistry, University of Bath, Claverton Down, Bath BA2 7AY, UK*

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

**Corresponding author*

Human somatic angiotensin I-converting enzyme (sACE) is a key regulator of blood pressure and an important drug target for combating cardiovascular and renal disease. sACE comprises two homologous metallopeptidase domains, N and C, joined by an inter-domain linker. Both domains are capable of cleaving the two hemoregulatory peptides angiotensin I and bradykinin, but differ in their affinities for a range of other substrates and inhibitors. Previously we determined the structure of testis ACE (C domain); here we present the crystal structure of the N domain of sACE (both in the presence and absence of the antihypertensive drug lisinopril) in order to aid the understanding of how these two domains differ in specificity and function. In addition, the structure of most of the inter-domain linker allows us to propose relative domain positions for sACE that may contribute to the domain cooperativity. The structure now provides a platform for the design of "domain-specific" second-generation ACE inhibitors.

Keywords: angiotensin I-converting enzyme; cardiovascular disease; crystal structure; hypertension; inhibitor design

Introduction

Angiotensin I-converting enzyme (ACE, EC 3.4.15.1) is a zinc-dependent dipeptidyl carboxypeptidase with diverse physiological functions, including principally that of blood pressure regulation *via* angiotensin II production and bradykinin inactivation. Somatic ACE (sACE), a type I transmembrane protein, is composed of two homologous catalytic domains (N and C domains), arising from a gene duplication event[1] (Figure 1(a)). Recent studies have highlighted the unique physiological roles of the N and C domains as well as evidence for negative cooperativity between them.[2–4] The germinal form of ACE (testis ACE)[5] originates from the same gene as somatic ACE, but has a tissue-specific promoter located within intron 12. Testis ACE (tACE) plays a crucial role in reproduction,[6] although the nature of its role is still under investigation.[7–9]

Enzymatically active forms of the N domain have been found in human ileal fluid[10] and in the urine of mild hypertensive patients where it could have an important role in the development of hypertension.[11] Despite sharing ~60% sequence identity with the C domain, the N domain has its own distinctive physicochemical and functional properties. It is thermally more stable than the C domain,[12] more resistant to proteolysis under denaturing conditions[13] and is less dependent on chloride activation as compared with the C domain.[14,15] Both domains are heavily glycosylated, a feature that has hampered the 3D structure determination of the protein.

The N domain has ten N-linked glycosylation sites of which seven are unique to the N domain. The different glycan profile of the N domain is likely to be responsible for the carbohydrate-mediated dimerisation of the somatic form, which has been described under certain conditions.[16,17] Moreover, the difference in glycosylation could impact on the structural basis for epitope recognition. Indeed, epitope mapping of the N domain has revealed

Abbreviations used: ACE, angiotensin I-converting enzyme; sACE, somatic angiotensin I-converting enzyme; tACE, testis angiotensin I-converting enzyme.

E-mail addresses of the corresponding authors: e.d.s-sturrock@curie.uct.ac.za; k.r.a-k.r.acharya@bath.ac.uk

a region that might play a role in the relatively inefficient ectodomain shedding of sACE compared to its germinal isoform.[18] Substrates such as the hemoregulatory peptide AcSDKP,[19] angiotensin 1–7,[20] and the enkephalin precursor Met^5-Enk-Arg^6-Phe^7[21] are specific for the N domain, whereas the physiological substrates bradykinin and angiotensin I are hydrolysed with similar catalytic efficiency as the C domain. Interestingly, the N domain preferentially hydrolyses the Aβ peptide of the amyloid precursor protein resulting in inhibition of Aβ aggregation and cytotoxicity in cell-based assays,[22] although this is not necessarily the case *in vivo*.[23] The widely used ACE inhibitor captopril shows modest selectivity for the N domain.[14] Also, the phosphinic peptide inhibitor RXP-407 has a dissociation constant three orders of magnitude lower for the N domain of the enzyme.[24]

Despite the physiological importance of ACE and the fact that it is a well-validated therapeutic target, the first 3D structures were only determined almost half a century after its discovery.[25–27] The crystal structures of tACE and the *Drosophila* homologue AnCE reveal a predominantly α-helical structure with two internal chambers linked by a constriction where the active-site zinc is coordinated by the two histidine residues of the HEXXH motif and a single downstream glutamate. Comparison of the ACE structure with that of the ACE 2 homologue,[28,29] which more closely resembles the N domain, provides a molecular basis for the different mechanisms involved in the dipeptidase and carboxypeptidase activities of these two metalloenzymes.

We have now determined the crystal structures of both native N domain of sACE and N domain complexed with the inhibitor lisinopril at a resolution of 3.0 Å (Table 1). The structure, including that of the inter-domain linker region reveals differences in the active site and chloride binding. The well ordered linker region enables us to speculate on the relative domain orientations in somatic ACE, which provides some insight into the basis for domain cooperativity.

Table 1. Crystallographic data

	Native	Lisinopril complex
Resolution range (Å)	48.5–3.0	50–3.0
No. of reflections measured	178,961	151,057
No. of unique reflections	36,858	35,264
$I/\sigma(I)$ (outer shell[a])	9.0 (2.2)	10.2 (2.2)
Completeness (outer shell[a]) (%)	99.8 (99.7)	95.5 (85.3)
R_{symm} (outer shell[a]) (%)	14.1 (59.5)	12.2 (34.9)
R_{cryst} (%)	22.4	29.4
R_{free}[b] (%)	27.3	31.8
Average temperature factor (Å^2)		
Protein (mol A/mol B)	35/39	39/40
Carbohydrate (mol A/mol B)	72/75	72/70
Ligand		49/43
Solvent (no. of water molecules)	25 [25]	13 [19]
Zn^{2+} ion/Cl^{2-} ion	32/40	39/35
RMSD[c] from ideal values		
Bond lengths (Å)	0.02	0.017
Bond angles (deg.)	1.7	2.4

[a] Outer shell, 3.11–3.00 Å and 3.16–3.00 Å for the native data and lisinopril complex, respectively.
[b] R_{free} calculation used 4% and 2.3% of the reflections for the native data and lisinopril complex, respectively.
[c] RMSD, root-mean-square deviation.

Results and Discussion

Overall structure of the N domain

The N domain was crystallised in the space group $C222_1$ with two molecules per asymmetric unit. The overall fold consists of a mainly helical secondary structure, with the same topology as tACE (Figure 1(b)). Briefly, the N domain has an ellipsoid shape with a central groove dividing it into two subdomains, one of which contains the N-terminal region that covers the central binding cavity (Figure 1(b) and (e)). There are 27 helices, of which 18 are α-helices, five are short 3_{10} helices and four are mixed; and only six short β-strands. The structures of both molecules in the asymmetric unit are very similar with a root-mean-square deviation for the C^α atoms of 0.50 Å. Analysis of the Ramachandran plot using the program PROCHECK[30] shows that 87% of the residues lie in the most favoured region, with none in the disallowed region. The N terminus was well ordered with the residues (1–612) being modelled for molecule A. The N-terminal residue missing in molecule B and residues 613–629 were missing in both molecules A and B.

The N domain has been implicated in dimerisation of human sACE in membranes[17] and the packing of the N domain includes two large dimer interfaces. Both have buried surface areas, encompassing helices 13, 14 and 35 on the one hand and 26 and 27 on the other with areas of 2310.2 Å^2 and 2811.5 Å^2, respectively. A contact region that might be involved in dimerisation has been proposed on the basis of epitope mapping,[18] but the arrangement of the monomers in the asymmetric unit is somewhat different. There is only partial overlap between the interface around helices 13,14 and 35 and the area proposed to be involved by Balyasnikova *et al.* However, the dimer interaction could be mediated by a carbohydrate recognition domain[17] and the N domain carbohydrates in the crystal structure were enzymatically truncated to facilitate crystallisation.

The catalytic zinc ion was observed at the active site and one chloride ion, equivalent to chloride II of tACE, adjacent to Arg500. The human N domain protein contains ten putative N-glycosylation sites and Fourier difference density was observed at five of these sites in the native structure (Figure 1(b)). Seven *N*-acetyl glucosamine residues were modelled on molecule A at asparagine residues 25, 45, 117, 289 and 480, and six on molecule B at residues 25, 45, 117 and 480. Although the *B*-factors for these

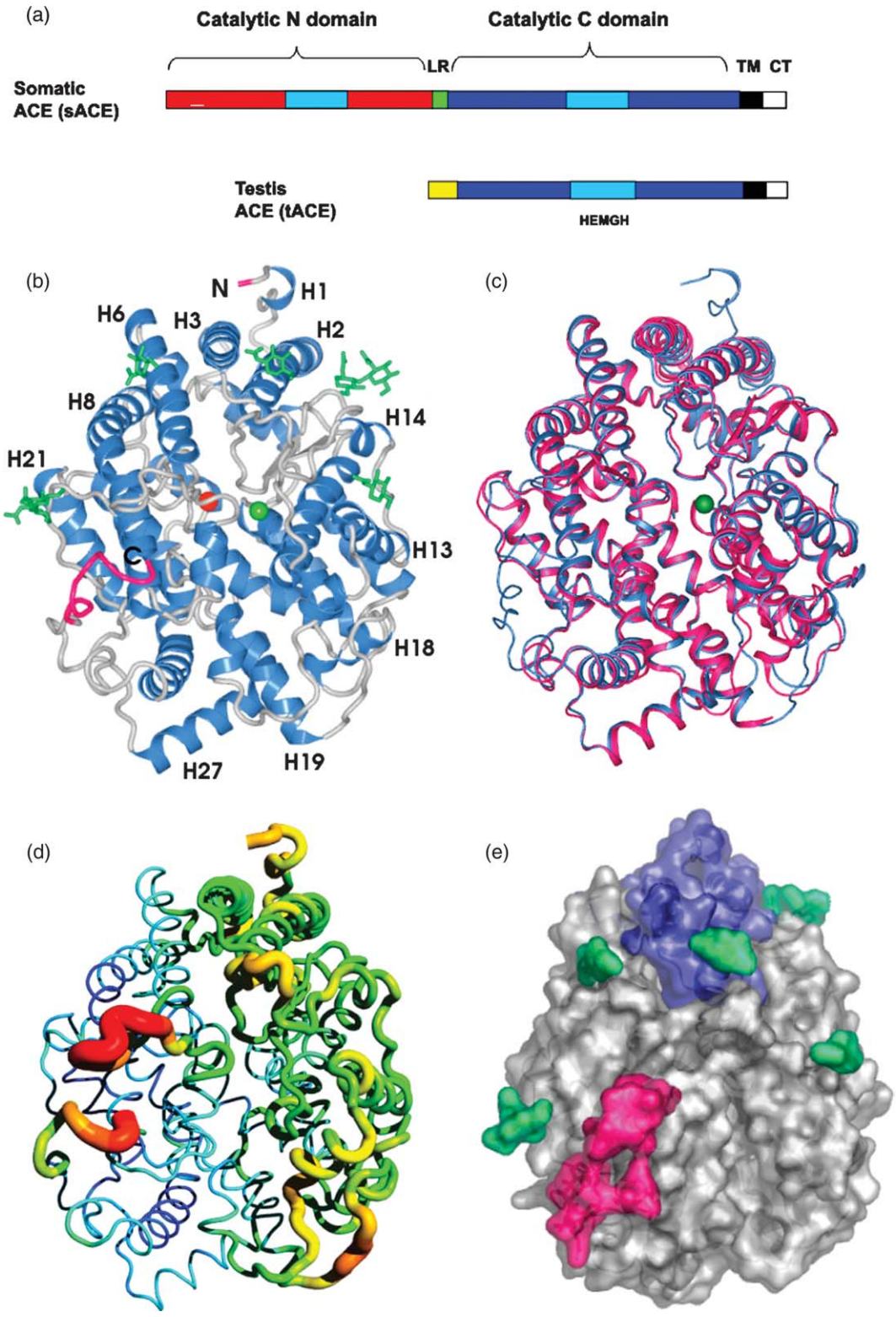

Figure 1. (a) The organisation of the different protein domains of the two isoforms of ACE, somatic ACE (sACE) and testis ACE (tACE). The two catalytic domains of sACE each contain the HEXXH zinc-binding motif and are connected by a short linker region (LR). Both forms comprise C-terminal juxtamembrane, transmembrane (TM), and cytosolic (CT) domains. (b) Overall structure of the N domain with the zinc ion in green, the chloride ion in red and the sugars in green. The N terminus and the inter-domain linker are shown in pink. Helices are numbered sequentially whether α or 3_{10} using the letter H. (c) Superposition of the N domain (blue) and tACE (pink) with the zinc ion in green. (d) Backbone coloured according to *B* factor with blue for the lowest *B* factors and red for the highest. The end of the inter-domain linker and the "flexible loop" are in red. (e) Surface diagram showing the lid covering the central binding cavity (dark blue), sugars (green) and the protruding surface patch (pink) comprising the linker and the flexible loop.

sugars are high, they are modelled on the basis of clear difference density, at least for the ring portions. A total of 25 water molecules were modelled by visual inspection and a glycerol molecule was modelled on the surface of each molecule adjacent to Glu219. Three acetate molecules were also modelled, one on the surface next to the symmetry axis and one in each of the active sites adjacent to Lys489, where a carboxyalanine was modelled in the tACE structure.[25]

The N domain protein was also crystallised in the presence of 5 mM lisinopril, for which clear difference density was observed upon refinement with the native N domain structure. The overall structure is the same as the native N domain, although the lack of completeness in the data results in poorly defined density for the N-terminal 30 residues. The lisinopril is bound at the active site in the same position and conformation as observed for tACE.[25] Four *N*-acetyl glucosamine residues on each molecule were modelled at asparagine 25, 117, and 480. Nineteen water molecules, two glycerols and one acetate ion were also modelled. The acetate ion was modelled near the symmetry axis and the two glycerol molecules at equivalent surface positions on the two molecules, similar to those observed in the native structure.

Comparison of the N and C domains

The N and C domains of somatic ACE have ~60% sequence identity and hence share the same overall topology as well as the highly conserved zinc-binding motif at the active site (Figure 2). As the sequence of tACE and the C domain of sACE are identical, except for the 36 N-terminal residues encoded by exon 13, the N and C domain structures can be compared using tACE. The most easily observable difference between the N and C domains (based on tACE) when superimposed (Figure 1(c)) is the extra length of the N domain at the N and C termini, the latter of which includes the inter-domain linker. The N terminus of the N domain protein, whilst having higher than average *B* factors, is well defined and packs against helix 3. Also, the loop between helices 19 and 20 (residues 409–417) that was not visible in the tACE structure, is well defined. Three other flexible loops, between helices 3 and 4, strands 1 and 2 and strand 6 and helix 23, show small differences between the two domains.

The N domain is activated at lower chloride concentrations and to a lesser extent than the C domain.[14] Consistent with this is the observation of only one chloride bound to the N domain, rather than the two observed for tACE. The chloride ion is observed at the identical site as chloride II in tACE, bound between Tyr202 and Arg500. At the equivalent position to the tACE chloride I site, a crucial arginine is substituted by a histidine in the N domain. Interestingly, ACE2, which was also only observed to bind one chloride ion in the crystal structure, binds chloride in the equivalent position to chloride I in tACE.[29]

The three (largest) N-terminal helices form a lid-like structure, referred to as the lid, that partly covers the substrate channel (Figure 1), as seen in the structure of tACE.[25] It was suggested that a change in conformation of the lid might be necessary for entry of the substrate and might also contribute to the substrate specificity of the domains. The unliganded form of ACE2, which has 41% sequence identity with the N domain of sACE, was observed in an open conformation, not seen for tACE or the N domain. This open conformation resulted from the hinge bending and subsequent movement of a sub-domain, referred to here as the lid sub-domain, comprising the lid plus a small β-sheet and five smaller helices (equivalent to helices 12–15, 18 and strands 4 and 5 in the N domain). It is possible that sACE also undergoes a similar flexing of this region during catalysis, but there is no evidence for this from the ACE crystal structures under the crystallisation conditions employed. The *B* factors for the equivalent lid sub-domain in the N domain are observed to be higher than those seen for the rest of the domain (Figure 1(d)) in both molecules of the asymmetric unit, suggesting that this region may be more flexible, although this could also be due to the smaller area of this region involved in crystal packing. Interestingly, a similar pattern is also seen in tACE, which has a different packing arrangement.

Comparison of the N and C domain active sites

The lisinopril bound to the active site was clearly defined, despite the poor electron density map in some other regions of the structure. The binding followed the same orientation as observed for tACE with the central C3 carboxylate coordinating the zinc ion and the phenyl group pointing up towards the lid. The active site residues and the lisinopril molecules for the two structures superimpose well (RMS deviation for 82 C^{α} atoms, 0.45 Å) with nearly all the lisinopril binding residues conserved in the N domain (Figure 3(a)).

Despite the structural homology between the two domains, there are some notable differences between the active sites, which might explain why lisinopril is more C-domain selective. The $S_{1'}$ subsite, where the lysyl group of lisinopril binds, shows some minor differences in backbone position compared to tACE (Figure 3(a)). For example, if Arg350 was aligned similarly to the equivalent residue in tACE, Val372, it might point inwards repelling the lysyl group of lisinopril. In our structure though, in molecule A where Arg350 is visible, the backbone is rotated towards the solvent, with the side-chain blocked from the active site by Gln355. The $S_{1'}$ subsite also shows some residue changes (Table 2) that appear to decrease the number of interactions with lisinopril. Firstly, Glu162 that was observed to interact electrostatically with the lysyl side-chain in the tACE complex, is replaced by Asp140 in the N domain. Asp140

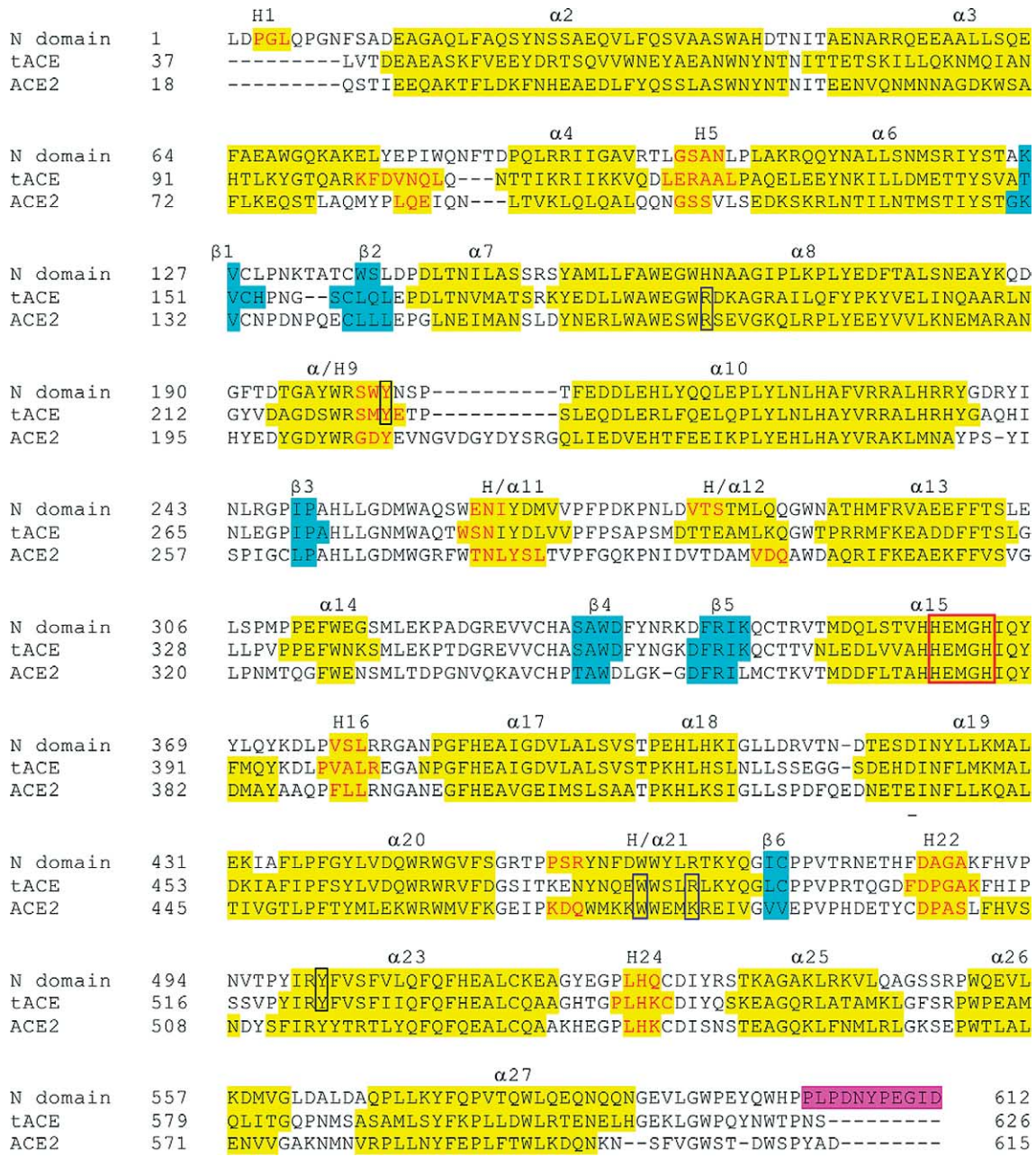

Figure 2. Sequence alignment of the sACE N domain with tACE and ACE2. Helices are highlighted in yellow and strands in blue. The linker region is shown in pink. Helices are numbered sequentially whether α or 3_{10}, and 3_{10} helices are shown with red letters. Chloride I coordinating residues are boxed in dark blue, chloride II in black, and the zinc binding motif in red.

does not make a similar contact with lisinopril as it is 6.5 Å away from the lysyl group, partly due to the shorter length of the residue, but also due to a slight shift in the helix 7 backbone. Secondly, Asp377 of tACE that was observed to form water-mediated interactions with the lysyl group, is also not conserved in the N domain and is replaced by Gln355. Gln355, 4 Å away from the lysyl residue, does not directly interact with it but might form water-mediated interactions, although no water molecules were visible in this region due to limited resolution of our structure. An important development in the field of ACE inhibitor design is the generation of domain-specific inhibitors to more precisely regulate the actions of the N and C domains of sACE. The two phosphinic peptide inhibitors RXP407 and RXPA380 are highly specific for the N and C domains, respectively.[24,31] The structural determinants of RXP407 important for N domain selectivity are known to be the N-terminal aspartate and *N*-acetyl groups in the P_2 position and the C-terminal amide, adjacent to the $P_{2'}$ position.[24]

The S_2 and $S_{2'}$ subsite residues of the N domain that differ in tACE (equivalent to the C domain) are also shown in Table 2. The S_2 residues most likely to confer N domain selectivity and contribute to

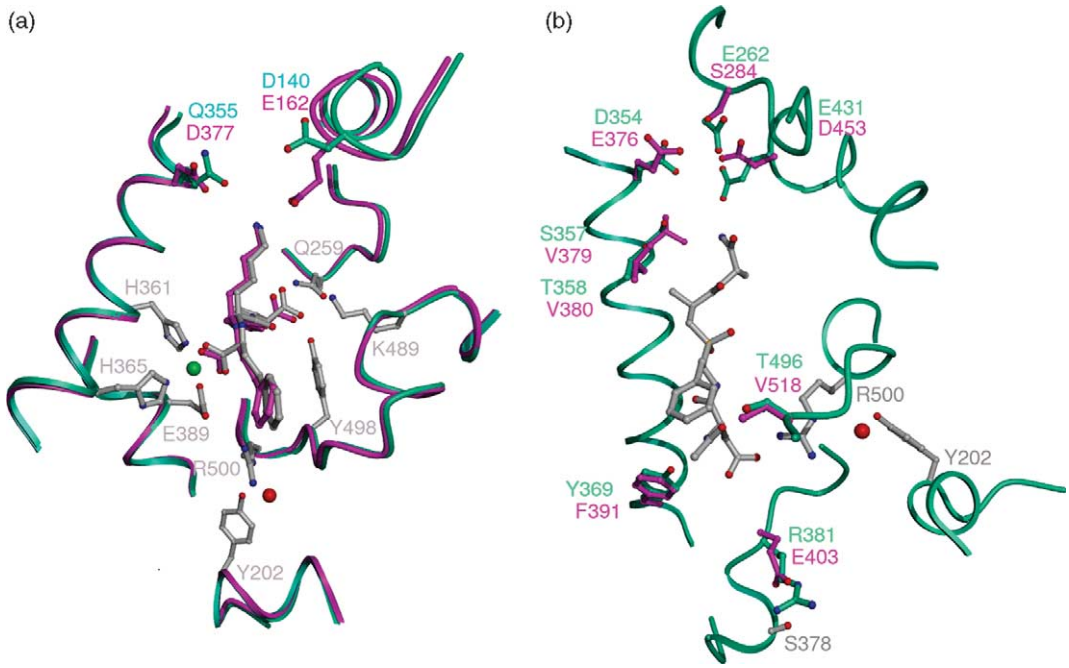

Figure 3. (a) Close up of the active site of the N domain (blue/green) and tACE (pink) with the zinc ion in green and the conserved chloride ion in red. Lisinopril is shown in grey/pink for the N domain/tACE, respectively. The lisinopril binding residues are shown in ball and stick. Residues common between the N domain and tACE are in grey (N domain numbering) with differing residues for these two proteins shown in blue/green and pink, respectively. (b) A different orientation of the native N domain active site (blue/green) with RXP407 (grey) approximately positioned using molecular docking. Residues differing between the N domain and tACE (equivalent to the C domain) in the $S_{2'}$ and S_2 subsites and the chloride coordinating residues are shown in ball and stick. Residues conserved between the N domain and tACE are shown in grey (N domain numbering), N domain residues are shown in blue/green and tACE residues in pink. The chloride ion is shown in red.

specific binding to the N-terminal groups of RXP407, are Tyr369, Arg381 and Thr496 (Phe391, Glu403 and Val518 in tACE) (Figure 3(b)). A common arginine, Arg500 in the N domain, which forms part of the chloride-binding site, is another candidate for RXP407 interaction. In our structure, although Arg500 is pointing towards subsite S_2, Arg381 is pointing away from the subsite in the lisinopril structure and molecule B of the native structure. However, in molecule A of the native structure Arg500 points towards the active site, showing that this conformation is also possible. Tyr369 also points towards the binding site and is observed to have a patch of difference density adjacent to it in both the native and lisinopril-complex structures. In both cases, this was modelled as water. As the inhibitor has both a charged and polar moiety, it is possible that both arginine residues (381 and 500), plus either the tyrosine or the threonine, interact with the inhibitor.

The $S_{2'}$ subsite also has several differences between the N and C domains, of which the charged residues are located at the far end of the cavity. This suggests an elongated conformation for the inhibitor if the C-terminal amide is a determinant for specificity, although the position of Ser357 and Thr358 on helix 15, provide a more polar environment than the two valine residues present in the C domain and may also contribute to the appropriate binding cavity for RXP407 (Figure 3(b)).

These subsite changes are also likely to play a role in the enzyme's specificity for N domain substrates. The case is more complex here, however, as N domain specific substrates such as AcSDKP, Ang1-7 and Met^5-Enk-Arg^6-Phe^7 also bind the C domain, but are hydrolysed much less efficiently. This suggests that specific determinants are responsible for positioning for optimal catalysis rather than just "tight binding". In the case of these substrates, the

Table 2. Active site residues that differ between the N and C domains (C domain numbering is as for tACE)

S_2	tACE/C domain	E403	F391	V518					
	N domain	R381	Y369	T496					
S_1	tACE/C domain	V518	S516	E143					
	N domain	T496	N494	S119					
$S_{1'}$	tACE/C domain	E162	N277	S284	E372	N374	E376	D377	V380
	N domain	D140	D255	E262	R350	T352	D354	N355	T358
$S_{2'}$	tACE/C domain	T282	S284	E376	V379	V380	D453		
	N domain	S260	D262	D354	S357	T358	E431		

composition of the S_1 sites in accommodating the aspartyl and methionyl groups in AcSDKP and Met[5]-Enk-Arg[6]-Phe[7], respectively, probably also plays a role in specificity. In particular, the N domain has Thr496 with a phenylalanine (490) in the proximal region of the S_1 cavity rather than valine and phenylalanine in the C domain, which decreases the hydrophobicity and provides a hydrogen bonding partner.

The domain arrangement of sACE

The complex kinetics of somatic ACE catalysis are partly due to the presence of an active site in both the N and C domains and also the potential for interaction between them. The N and C domains of sACE are joined by a linker that is susceptible to proteolysis[13] and is assumed to be partly flexible. The inclusion of the linker in the N domain construct allowed us to visualise this region for the first time. The linker (residues 602–612) was well defined in the electron density map after the first couple of rounds of refinement and was built into the model without ambiguity. The first part of the linker appears to be rigid in our structure and is held in place by a hydrogen bond between Tyr607 and Glu161. The last four residues are more flexible as they have high *B* factors and their side-chains are not visible (Figure 1(d)). They form a prominent surface patch, away from the core of the N domain, with a loop, anchored by the disulphide bond between Cys128 and Cys136, that appears to be very flexible (Figure 1(d) and (e)) in both molecules of the asymmetric unit. The equivalent loop in tACE also has high *B* factors, suggesting that there may be some other functional significance to the flexibility. However, its position suggests that it would form part of the inter-domain interface, and its flexibility suggests it could adapt its conformation to mould to the C domain.

There is some evidence that only one domain of sACE is capable of catalysis at a time[3,4] and that the N domain may be a negative regulator of the C domain,[2] suggesting some cooperativity between the two active sites. This can be explained in terms of the C domain, if indeed flexing of the lid or lid sub-domain are involved in catalysis, as the lid is at the N terminus of the C domain and must lie in close proximity to the linker and the N domain. The interaction of the N domain with the C domain lid might reduce the flexibility of this region and regulate access of substrate to the C domain active site or even form a surface that partially blocks the cavity opening. In contrast, the position of the linker ensures that the C domain is unlikely to interact with the lid of the N domain. However, the flexible loop adjacent to the linker is close to the basolateral opening of the active site and adjacent to helix 7, which forms part of the $S_{1'}$ subsite. It is also adjacent to the large loop between helix 14 and strand 4 that was part of the lid sub-domain in ACE2.

Precise modelling of sACE is not possible as the three N-terminal residues of the C domain were not visible in this structure of tACE. Furthermore, several side-chains on the linker and the flexible loop, which one might assume to play a part in the interaction, are disordered and not visible. In particular, there is an N-glycosylation sequon on the flexible loop of the N domain and the top part of the C domain lid, suggesting that inter-domain interactions and movements may be mediated or aided by sugars. However, based on our current structure, we have constructed two models of sACE in order to visualise possible interacting regions between the domains (Figure 4). These domain arrangements illustrate the two extremes of a possible rotation around a proposed hinge between the two domains. Although the precise boundaries of the hinge cannot be determined without the full sACE structure, based on the *B*-factors, only four to seven residues display any significant flexibility and this would limit any hinged-rotation for such large domains. A much larger structural reorganisation of the whole linker, or other parts of the structure, cannot be dismissed, but a small flexible region between the two compact domains seems likely from the available structural information. These two orientations are discussed briefly in order to highlight areas of the N and C domains that might be responsible for inter-domain communication.

The first, more extended orientation, allows the inter-domain linker to pack against the lid of the C domain, in a similar manner to how the extra N-terminal residues of the N domain pack against the N domain lid (Figure 4). This produces a snug fit between the lid of the C domain and the linker and flexible loop of the N domain but over a small contact area. Rotation of the C domain around this contact point to increase the contact area is possible. In this position, the positively charged lysine on the flexible loop and the three negatively charged residues on the linker may interact with some of the charged residues protruding from the C domain lid. The glycosylation on the C domain lid is likely to contribute to limiting the possible orientations of this kind.

The second orientation is more compact, following a 180 degree rotation from the first, allowing the C domain to continue in line with the linker rather than doubling back. This enables the C domain to contact parts of the N domain lid sub-domain as shown in Figure 4. However, a turn in the end of the linker would allow the C domain to rotate so that its lid could interact directly with the N domain lid sub-domain. In this kind of orientation, the N domain glycosylation at Asn45 and 289 on the lid sub-domain may also participate in the interaction. However, as the glycosylation of somatic ACE will have much more extensive glycosylation than can be viewed in this crystallographic analysis (due to truncation and disorder of the glycosyl chains), any possible role of glycosylation in domain interactions requires further investigation.

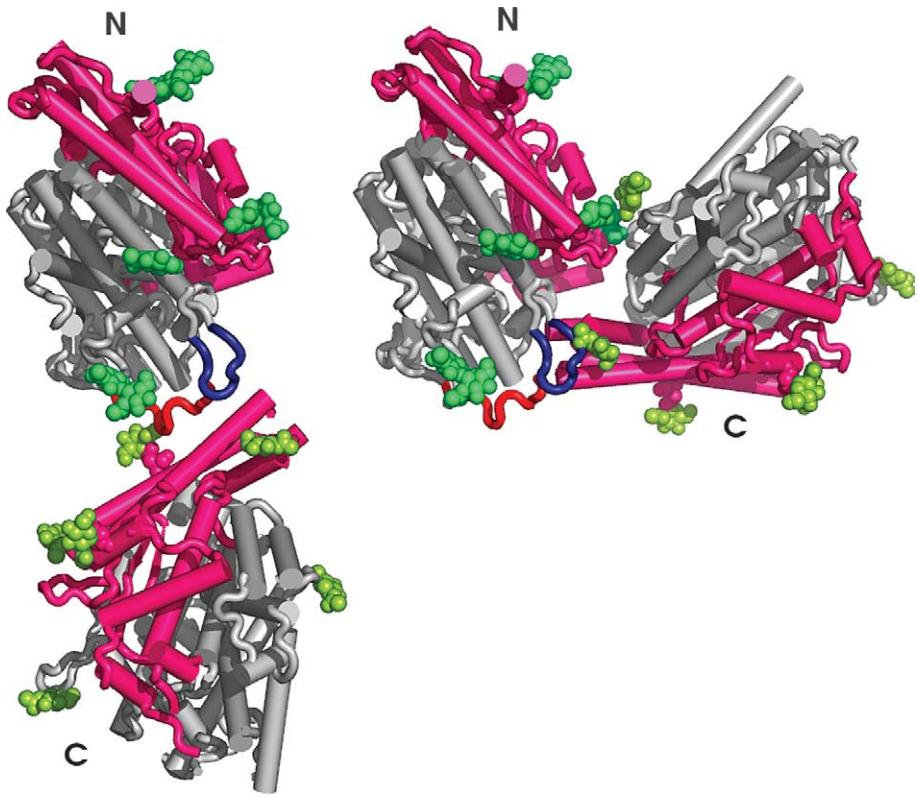

Figure 4. Two possible types of domain orientation for sACE. The sub-domain containing the lid, identified in ACE2, is shown in pink for both the N and C domains with the inter-domain linker shown in red and the flexible loop in dark blue. A bulky sphere representation is used for the sugar moieties from the N domain (green) and tACE/C domain (green/yellow).

Conclusion

The determination of the N domain structure is an important advance in the understanding of the overall structure of human sACE. It allows insight into the specificity and the physiological significance of the N and C domain active sites. There has been a renewed interest in the molecular basis for the enzyme activity of ACE, a cardiovascular drug target of enormous clinical importance. The structural details of previously determined tACE (C domain of sACE) along with the N domain of sACE presented here will clearly aid future rational design of domain-selective ACE inhibitors. Furthermore, the additional structure of the sACE linker region provides a springboard for studying the intriguing manner in which these two enzymatically active domains are likely to interact.

Materials and Methods

Expression and purification

An N domain construct D629 cloned into the vector pECE,[32] encoding the first 629 residues of somatic ACE, was subcloned into pBlueScript *via* XbaI and EcoRI and sequenced. D629 was subcloned into pcDNA3.1(+) using the same restriction sites. This N domain construct was then subsequently introduced into the CHO cell glutamine synthetase (GS) expression vector pEE14 using HindIII and XbaI. The identity of the construct and its correct orientation was confirmed *via* restriction enzyme digests.

CHO-K1 cells were co-transfected with pEE N domain and pSV2neo (10:1). Geneticin (G418) (Sigma)-resistant clones expressing soluble N domain were further selected for resistance to methionine sulfoximine (MSX).[33] Cells stably expressing N domain were grown in Glasgow Minimum Essential Medium (GMEM) supplemented with 10% (v/v) dialysed foetal calf serum (Gibco BRL) and 20 μM MSX. When confluent, growth medium was changed to 1% dialysed FCS, 0.05% albumax I (Gibco BRL), 20 μM MSX and 1.5 mM *N*-butyldeoxynojirimycin (NBDNJ) (Toronto Research Chemicals Inc.). Medium containing soluble N domain was harvested and purified in tandem over a pre-protein-G Agarose (Sigma) column followed by immunoaffinity chromatography using an N domain specific monoclonal antibody (5C5) (a kind gift from Dr Sergei Danilov) coupled to protein G Agarose. N domain was eluted with 50 mM ethanolamine (pH 11.5). It was dialysed against 5 mM Hepes (pH 7.5), 0.1 mM PMSF and concentrated in a 30 kDa Amicon concentrator at 1000–2000*g* and 4 °C to a concentration of 10 mg/ml, and stored at 4 °C. The purity of the protein was checked using 10% (w/v) SDS-PAGE.

Crystallisation, X-ray data collection and structure determination

Native N domain

One microliter N domain at 4 mg/ml was mixed with 1 μl reservoir solution (0.2 M lithium sulphate, 0.1 M sodium acetate (pH 4.9), 10 μM zinc sulphate, 15% (w/v)

polyethylene glycol 4000) and suspended above the reservoir as a hanging drop at 16 °C. Crystals grew within 1–2 weeks. A single crystal was cryocooled (100 K) using reservoir solution plus 25% (v/v) glycerol as a cryoprotectant. Diffraction data to 3.0 Å were collected on station PX14.1 of the synchrotron radiation source (Daresbury, UK) using a Quantum 4 charge-coupled-device detector (Area Detection Systems, Poway, CA). The data were processed and scaled using the HKL2000 software package (HKL Research, Charlottesville, VA).[34] The symmetry and systematic absences were consistent with the $C222_1$ space group (unit cell dimensions, $a=101.1$ Å; $b=211.3$ Å; and $c=171.3$ Å) with two protein molecules per asymmetric unit. The crystals contained ~54% solvent. Data reduction was carried out by using the CCP4 program TRUNCATE.[35]

The structure of the N domain was solved with the program MOLREP[36] using a homology model of the N domain based on the tACE structure (Protein Data Bank entry 1O8A) as a search model. Initial refinement was performed using REFMAC[37] through the CCP4i interface.[35] Four percent of reflections was kept aside for R_{free} calculation.[38] A large part of the C-terminal region not present in the tACE structure could be built after the first round of refinement using the program Coot.[39] Further rounds of refinement and model building allowed the building of the N terminus and the addition of water, sugar and glycerol molecules plus three acetate ions. Final refinement was performed using the CNS suite[40] (Table 1).

N domain–lisinopril complex

N domain, 4 mg/ml, was incubated with 5 mM lisinopril for 5 h at 4 °C. One microlitre N domain with lisinopril was mixed with 1 µl reservoir solution (0.2 M lithium sulphate, 0.1 M sodium acetate (pH 4.9), 10 µM zinc sulphate, 18% polyethylene glycol 4000) and suspended above the reservoir as a hanging drop at 16 °C. Crystals grew within one to two weeks in the same form as the native (unit cell dimensions, $a=101.3$ Å; $b=210.9$ Å; and $c=171.0$ Å). Diffraction data to 3.0 Å were collected as for the native data set. The data were processed with MOSFLM,[41] scaled with SCALA[35] and data reduction carried out by TRUNCATE.[35]

Refinement (with 2.3% of the reflection kept for the R_{free} calculation) of the native N domain structure with the CNS suite[40] showed clear difference density for the lisinopril, which was modelled with the program Coot,[39] along with a few water, sugar and glycerol molecules and an acetate ion (Table 1).

RXP407 modelling procedure

Docking exercises with RXP407 and the N domain were carried out using INSIGHT II (Accelrys Inc., Version 98.0). The Consistent Valence Force Field and the Extensible Systematic Force Field (metal adapted)† were used in all energy minimizations and dynamic runs. Conjugate gradient minimization was used after the initial 1000 steps, followed by 3000 cycles of molecular dynamics and 3000 cycles of energy minimization in an NTV (constant volume) ensemble, at a temperature of 300 K. All calculations were carried out in a dielectric constant of 1.00 and a cut-off distance of 9.5 Å.

† http://www.accelrys.com/support/life/discover/forcefield/esffSBL.html

Protein Data Bank accession codes

The atomic coordinates have been deposited with the RCSB Protein Data Bank‡, and the accession codes are 2C6F and 2C6N for the native and lisinopril complexes for sACE N domain, respectively.

Acknowledgements

We thank the staff at the Synchrotron Radiation Source, Daresbury (UK) for their support during X-ray data collection. This work was supported by the Wellcome Trust, UK (Project Grant 071047 to K.R.A., a Senior International Research Fellowship 070060 to E.D.S.) and the National Research Foundation, South Africa (grant to E.D.S.). The authors declare that they do not have conflicting financial interests.

References

1. Soubrier, F., Alhenc-Gelas, F., Hubert, C., Allegrini, J., John, M., Tregear, G. & Corvol, P. (1988). Two putative active centers in human angiotensin I-converting enzyme revealed by molecular cloning. *Proc. Natl Acad. Sci. USA*, **85**, 9386–9390.
2. Woodman, Z. L., Schwager, S. L. U., Redelinghuys, P., Carmona, A. K., Ehlers, M. R. W. & Sturrock, E. D. (2005). The N domain of somatic angiotensin-converting enzyme negatively regulates ectodomain shedding and catalytic activity. *Biochem. J.* **389**, 739–744.
3. Binevski, P. V., Sizova, E. A., Pozdnev, V. F. & Kost, O. A. (2003). Evidence for the negative cooperativity of the two active sites within bovine somatic angiotensin-converting enzyme. *FEBS Letters*, **550**, 84–88.
4. Andújar-Sánchez, M., Cámara-Artigas, A. & Jara-Pérez, V. (2004). A calorimetric study of the binding of lisinopril, enalaprilat and captopril to angiotensin-converting enzyme. *Biophys. Chem.* **111**, 183–189.
5. Ehlers, M. R. W., Fox, E. A., Strydom, D. J. & Riordan, J. F. (1989). Molecular cloning of human testicular angiotensin-converting enzyme: the testis isozyme is identical to the C-terminal half of endothelial angiotensin-converting enzyme. *Proc. Natl Acad. Sci. USA*, **86**, 7741–7745.
6. Hagaman, J. R., Moyer, J. S., Bachman, E. S., Sibony, M., Magyar, P. L., Welch, J. E. *et al*. (1998). Angiotensin-converting enzyme and male fertility. *Proc. Natl Acad. Sci. USA*, **95**, 2552–2557.
7. Kondoh, G., Tojo, H., Nakatani, Y., Komazawa, N., Murata, C., Yamagata, K. *et al*. (2005). Angiotensin-converting enzyme is a GPI-anchored protein releasing factor crucial for fertilisation. *Nature Med.* **11**, 160–166.
8. Leslie, L., Parkin, E. T., Turner, A. J. & Hooper, N. M. (2005). Angiotensin-converting enzyme as a GPIase: a critical reevaluation. *Nature Med.* **11**, 1139–1140.
9. Fuchs, S., Frenzel, K., Hubert, C., Lyng, R., Muller, L., Michaud, A. *et al*. (2005). Male fertility is dependent on dipeptidase activity of testis ACE. *Nature Med.* **11**, 1140–1142.

‡ www.rcsb.org

10. Deddish, P. A., Wang, J., Michel, B., Morris, P. W., Davidson, N. O., Skidgel, R. A. & Erdos, E. G. (1994). Naturally occurring active N-domain of human angiotensin I-converting enzyme. *Proc. Natl Acad. Sci. USA*, **91**, 7807–7811.
11. Casarini, D. E., Plavinik, F. L., Zanella, M. T., Marson, O., Krieger, J. E., Hirata, I. Y. & Stella, R. C. (2001). Angiotensin converting enzymes from human urine of mild hypertensive untreated patients resemble the N-terminal fragment of human angiotensin I-converting enzyme. *Int. J. Biochem. Cell Biol.* **33**, 75–85.
12. Voronov, S., Zueva, N., Orlov, V., Arutyunyan, A. & Kost, O. (2002). Temperature-induces selective death of the C-domain within angiotensin-converting enzyme molecule. *FEBS Letters*, **522**, 77–82.
13. Sturrock, E. D., Danilov, S. M. & Riordan, J. F. (1997). Limited proteolysis of human kidney angiotensin-converting enzyme and generation of catalytically active N- and C-terminal domains. *Biochem. Biophys. Res. Commun.* **236**, 16–19.
14. Wei, L., Clauser, E., Alhenc-Gelas, F. & Corvol, P. (1992). The two homologous domains of human angiotensin I-converting enzyme interact differently with competitive inhibitors. *J. Biol. Chem.* **267**, 13389–13405.
15. Jaspard, E., Wei, L. & Alhenc-Gelas, F. (1993). Differences in the properties and enzymatic specificities of the two active sites of the two active sites of angiotensin I-converting enzyme (kininase II). Studies with bradykinin and other natural peptides. *J. Biol. Chem.* **268**, 9496–9503.
16. Kost, O. A., Bovin, N. V., Chemodanova, E. E., Nasonov, V. V. & Orth, T. A. (2000). New feature of angiotensin-converting enzyme: carbohydrate-recognizing domain. *J. Mol. Recognit.* **13**, 360–369.
17. Kost, O. A., Balyasnikova, I. V., Chemodanova, E. E., Nikolskaya, I. I., Albrecht, R. F. I. & Danilov, S. M. (2003). Epitope-dependent blocking of the angiotensin-converting enzyme dimerization by monoclonal antibodies to the N-terminal domain of ACE: possible link of ACE dimerization and shedding from the cell surface. *Biochemistry*, **42**, 6965–6976.
18. Balyasnikova, I. V., Woodman, Z. L., Albrecht, R. F. I., Natesh, R., Acharya, K. R., Sturrock, E. D. & Danilov, S. M. (2005). Localization of an N-domain region of angiotensin-converting enzyme involved in the regulation of ectodomain shedding using monoclonal antibodies. *J. Proteome Res.* **4**, 258–267.
19. Rousseau, A., Michaud, A., Chauvet, M. T., Lenfant, M. & Corvol, P. (1995). The hemoregulatory peptide *N*-acetyl-Ser-Asp-Lys-Pro is a natural and specific substrate of the N-terminal active site of human angiotensin-converting enzyme. *J. Biol. Chem.* **270**, 3656–3661.
20. Deddish, P. A., Marcic, B., Jackman, H. L., Wang, H. Z., Skidgel, R. A. & Erdos, E. G. (1998). N-domain-specific substrate and C-domain inhibitors of angiotensin-converting enzyme: angiotensin-(1-7) and keto-ACE. *Hypertension*, **31**, 912–917.
21. Deddish, P. A., Jackman, H. L., Skidgel, R. A. & Erdos, E. G. (1997). Differences in the hydrolysis of enkephalin congeners by the two domains of angiotensin converting enzyme. *Biochem. Pharmacol.* **53**, 1459–1463.
22. Oba, R., Igarashi, A., Kamata, M., Nagata, K., Takano, S. & Nakagawa, H. (2005). The *N*-terminal active centre of human angiotensin-converting enzyme degrades Alzheimer amyloid beta-peptide. *Eur. J. Neurosci.* **21**, 733–740.
23. Hemming, M. L. & Selkoe, D. J. (2005). Amyloid beta-protein is degraded by cellular angiotensin-converting enzyme (ACE) and elevated by an ACE inhibitor. *J. Biol. Chem.* **280**, 37644–37650.
24. Dive, V., Cotton, J., Yiotakis, A., Michaud, A., Vassiliou, S., Jiracek, J. *et al*. (1999). RXP 407, a phosphinic peptide, is a potent inhibitor of angiotensin I converting enzyme able to differentiate between its two active sites. *Proc. Natl Acad. Sci. USA*, **96**, 4330–4335.
25. Natesh, R., Schwager, S. L. U., Sturrock, E. D. & Acharya, K. R. (2003). Crystal structure of the human angiotensin-converting enzyme–lisinopril complex. *Nature*, **421**, 551–554.
26. Natesh, R., Schwager, S. L. U., Evans, H. R., Sturrock, E. D. & Acharya, K. R. (2004). Structural details on the binding of antihypertensive drugs captopril and enalaprilat to human testicular angiotensin I-converting enzyme. *Biochemistry*, **43**, 8718–8724.
27. Kim, H. M., Shin, D. R., Yoo, O. J., Lee, H. & Lee, J. O. (2003). Crystal structure of *Drosophila* angiotensin I-converting enzyme bound to captopril and lisinopril. *FEBS Letters*, **538**, 65–70.
28. Guy, J. L., Jackson, R. M., Acharya, K. R., Sturrock, E. D., Hooper, N. M. & Turner, A. J. (2003). Angiotensin-converting enzyme-2 (ACE2): comparative modeling of the active site, specificity requirements, and chloride dependence. *Biochemistry*, **42**, 13185–13192.
29. Towler, P., Staker, B., Prasad, S. G., Menon, S., Tang, J., Parsons, T. *et al*. (2004). ACE2 structures reveal a large hinge-bending motion important for inhibitor binding and catalysis. *J. Biol. Chem.* **279**, 17996–18007.
30. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thronton, J. M. (1993). PROCHECK—a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
31. Georgiadis, D., Cuniasse, P., Cotton, J., Yiotakis, A. & Dive, V. (2004). Structural determinants of RXPA380, a potent and highly selective inhibitor of angiotensin-converting enzyme C-domain. *Biochemistry*, **43**, 8048–8054.
32. Balyasnikova, I. V., Metzger, R., Franke, F. E. & Danilov, S. M. (2003). Monoclonal antibodies to denatured human ACE (CD 143), broad species specificity, reactivity on paraffin sections, and detection of subtle conformational changes in the C-terminal domain of ACE. *Tissue Antigens*, **61**, 49–62.
33. Yu, X. C., Sturrock, E. D., Wu, Z., Biemann, K., Ehlers, M. R. W. & Riordan, J. F. (1997). Identificaton of *N*-linked glycosylation sites in human testis angiotensin-converting enzyme and expression of an active deglycosylated form. *J. Biol. Chem.* **272**, 3511–3519.
34. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
35. Collaborative Computational Project Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
36. Vagin, A. & Teplyakov, A. (2000). An approach to multi-copy search in molecular replacement. *Acta Crystallog. sect. D*, **56**, 1622–1624.

37. Murshudov, G. N. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D*, **53**, 240–255.
38. Brünger, A. T. (1992). Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472–475.
39. Emsley, P. & Cowtan, K. (2004). Coot: model building tools for molecular graphics. *Acta Crystallog.* **60**, 2126–2132.
40. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., GrosseKunstleve, R. W. *et al*. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
41. Leslie, A. G. W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4+ESF-EAMCB Newsletter on Protein Crystallography*, **26**.

Edited by I. Wilson

(Received 29 November 2005; received in revised form 6 January 2006; accepted 10 January 2006)
Available online 31 January 2006