Structure of Testis ACE Glycosylation Mutants and Evidence for Conserved Domain Movement[†,‡]

Jean M. Watermeyer,[§] B. Trevor Sewell,[§] Sylva L. Schwager,[§] Ramanathan Natesh,[||,⊥] Hazel R. Corradi,[||] K. Ravi Acharya,*[,||] and Edward D. Sturrock*[,§]

Division of Medical Biochemistry, Institute of Infectious Disease and Molecular Medicine, University of Cape Town, South Africa, and Department of Biology and Biochemistry, University of Bath, Claverton Down, Bath BA2 7AY, U.K.

Received June 9, 2006; Revised Manuscript Received August 14, 2006

ABSTRACT: Human angiotensin-converting enzyme is an important drug target for which little structural information has been available until recent years. The slow progress in obtaining a crystal structure was due to the problem of surface glycosylation, a difficulty that has thus far been overcome by the use of a glucosidase-1 inhibitor in the tissue culture medium. However, the prohibitive cost of these inhibitors and incomplete glucosidase inhibition makes alternative routes to minimizing the *N*-glycan heterogeneity desirable. Here, glycosylation in the testis isoform (tACE) has been reduced by Asn-Gln point mutations at N-glycosylation sites, and the crystal structures of mutants having two and four intact sites have been solved to 2.0 Å and 2.8 Å, respectively. Both mutants show close structural identity with the wild-type. A hinge mechanism is proposed for substrate entry into the active cleft, based on homology to human ACE2 at the levels of sequence and flexibility. This is supported by normal-mode analysis that reveals intrinsic flexibility about the active site of tACE. Subdomain II, containing bound chloride and zinc ions, is found to have greater stability than subdomain I in the structures of three ACE homologues. Crystallizable glycosylation mutants open up new possibilities for cocrystallization studies to aid the design of novel ACE inhibitors.

Since its isolation in 1956 as "hypertensin-converting enzyme", human angiotensin-converting enzyme (ACE[1]) has been known to play a key role in the regulation of blood pressure (*1*). ACE is a membrane-bound zinc metalloprotease of the M2 family, acting as a dicarboxypeptidase on a range of oligopeptide substrates. Its hypertension-inducing activity is mediated through cleavage of two of these substrates: inactive decapaptide angiotensin I is converted into the active vasopressor angiotensin II, while cleavage of bradykinin inactivates this vasodilatory nonapeptide (reviewed in ref *2*). Inhibitors of ACE have been shown to be effective in the treatment of hypertension and cardiac disease, and a number of these drugs are commonly prescribed today (reviewed in ref *3*).

Two isoforms of ACE occur in mammals: a somatic form (sACE) that is expressed in most mammalian tissues but primarily in the lungs, vascular endothelia, and kidneys, and a germinal form (tACE) that arises only in the adult male testis (*4*, *5*). sACE has two homologous domains, each containing an active site, while tACE has a single active domain arising by tissue-specific transcription from a promoter in intron 12 of the sACE gene (*6*, *7*). tACE is thus identical to the C-terminal domain of sACE, except for an additional 67-residue N-terminal signal peptide and O-glycosylated region (*8*). Both forms are shed from the cell membrane by an ACE sheddase that cleaves at a site in the juxtamembrane stalk region (*9*).

Despite ACE's importance as a drug target, the first crystal structure of an ACE isoform, human tACE, was only solved recently (*10*). The major obstacle to the solution of this structure had been the presence of a number of N-linked glycan chains on the surface of the enzyme, which prevented crystallization (*11*). Crystallizable material was eventually obtained by expressing truncated tACE (lacking the N-terminal O-glycosylated and transmembrane regions) in the presence of an α-glucosidase-1 inhibitor, *N*-butyl-deox-

[†] This work was supported by the Carnegie Corporation of New York, the University of Cape Town, the South Afican National Research Foundation, and the Wellcome Trust (U.K.), Grants 070060 and 071047.

[‡] The atomic coordinates and structure factors for glycosylation mutants tACE-G13 (codes 2iul and r2iulsf) and tACE-G1234 (codes 2iux and r2iuxsf) have been deposited in the RCSB Protein Data Bank, www.pdb.org.

* To whom correspondence should be addressed. E.D.S.: Division of Medical Biochemistry, Institute of Infectious Disease and Molecular Medicine, UCT Faculty Of Health Sciences, Anzio Road, Observatory 7925, South Africa; tel, +27-21406 6312; fax, +27-21406 6470; e-mail, sturrock@curie.uct.ac.za. K.R.A.: Department of Biology and Biochemistry, University of Bath, Claverton Down, Bath BA2 7AY, U.K.; tel, +44-1225-386238; fax, +44-1225-386779; e-mail, K.R.Acharya@bath.ac.uk.

[§] University of Cape Town.

[||] University of Bath.

[⊥] Present address: Department of Structural Biology and Genomics, IGBMC, 1 rue Laurent Fries, 67404 Illkirch Cedex, France.

[1] Abbreviations: ACE, angiotensin converting enzyme 1; sACE, somatic ACE; tACE, testis ACE; NB-DNJ, *N*-butyl-deoxynojirimycin; Ndom, N-terminal domain of sACE; ACE2, human angiotensin-converting enzyme 2; NMA, normal mode analysis; tACE-G13, mutant of human tACE lacking the N-terminal 36 residues and all but the first and third N-glycosylation sites; tACE-G1234, mutant of human tACE lacking the N-terminal 36 residues and all but the first four N-glycosylation sites; ACE2o, open unliganded structure of ACE2; ACE2c, closed liganded structure of ACE2; tACEo, model of human tACE based on ACE2o; tACEc, model of human tACE based on ACE2c; NAG, *N*-acetyl glucosamine.

10.1021/bi061146z CCC: $33.50 © 2006 American Chemical Society
Published on Web 10/03/2006

ynojirimycin (NB-DNJ), which enforces uniform glycosylation to the simple high-mannose oligosaccharide level (*12*). This approach was later also used in the crystallization and solution of the structure of the N domain of human sACE (Ndom) (*13*).

Another approach yielding promising initial results is the elimination of N-linked glycosylation sites by point mutation. It has been demonstrated that selected, but not all, N-linked glycosylation sites of human tACE can be removed, while maintaining functional integrity and crystal morphologies similar to the wild-type (*14*). This method bears further investigation since the extensive cocrystallization studies desirable for a structure-based drug design process would be prohibitively expensive given the cost of glucosidase-1 inhibitors such as NB-DNJ.

The crystal structures of tACE alone and bound to inhibitors lisinopril, captopril, and enalaprilat revealed a largely α-helical globular enzyme, divided down the middle by a deep cleft (*10*, *15*). This cleft is largely closed off from the external milieu, leaving only small pores less than 3 Å in diameter through which substrates could gain access to the active site. While the substrates of ACE are limited to short oligopeptides, even these would require an enlarging of these openings in order to enter the cleft and access the active site, and a flexibility or breathing motion has been suggested to allow substrate binding (*11*). A similar closed conformation was also observed for Ndom, the *Drosophila* homologue AnCE, and human homologue ACE2, all at least 40% homologous to tACE (*13*, *16*, *17*). However, in the case of ACE2, the structure determined without bound inhibitor was in an open conformation in which the sides of the active cleft hinged apart by ∼16° (*17*). Given the structural similarity between Ndom, tACE, and ACE2, and the need for movement to allow substrate entry, it seems likely that a similar hinge mechanism occurs in both domains of sACE during substrate binding. Such a hinge mechanism would also explain the large contribution of entropy that has been observed in the energetics of inhibitor binding by sACE, since closing of the active site would result in the association of numerous residues that would otherwise be bound to ordered solvent molecules when in an open conformation (*18*).

Here, we present the crystal structures of two glycosylation mutants of human tACE and show that the removal of intact glycan chains does not affect the three-dimensional structure. In addition, we draw together lines of evidence from normal mode analysis (NMA) and the crystal structures of ACE and ACE homologues that support the hypothesis of a conserved hinge mechanism for substrate entry.

MATERIALS AND METHODS

Protein Purification. Mutants tACE-G13 and tACE-G1234 were previously constructed and expressed in Chinese hamster ovary (CHO) cells (*14*). Their sequences differ from that of wild-type human tACE, in that they lack the N-terminal O-glycosylated region (residues1−36) and have had some of their N-glycosylation sites knocked out by Asn−Gln mutation (Figure 1). Specifically, tACE-G13 lacks all but the first and third sites, while tACE-G1234 lacks the fifth and sixth sites. The seventh potential glycosylation sequon was not mutated as it lies close to the cleavage site

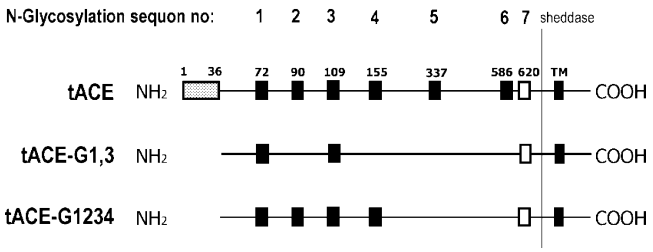

FIGURE 1: Schematic diagram of mutations introduced in tACE-G13 and tACE-G1234. Full-length wild-type tACE has an N-terminal O-glycosylated region (1−36), N-glycosylated sites 1−6 (filled squares), one unglycosylated site (open square), and a C-terminal transmembrane domain (TM). Both mutants lack the N-terminal 36 residues. Additionally, tACE-G13 lacks all but the first and third N-glycosylation sites, while tACE-G1234 lacks the fifth and sixth sites. Expressed protein is shed into the medium by cleavage between Arg 627 and Ser 628.

for shedding and has been shown to be unglycosylated (*12*). Soluble tACE-G13 and tACE-G1234 were expressed as previously described, and purified from harvested medium by lisinopril-Sepharose affinity chromatography (*5*, *14*, *19*).

ACE activity was determined by measuring the hydrolysis of hippuryl-histidyl-leucine (Sigma) in a fluorimetric assay (*20*). Protein concentration was determined by a Bradford protein assay (Bio-Rad protein microassay).

Crystallization, Data Collection, and Processing. Crystals of tACE-G13 and tACE-G1234 were grown as for the wild-type structure, published previously (*10*). Diffraction data to 2.9 Å for tACE-G13 were collected at 100 K at the in-house X-ray source in the Department of Biotechnology, University of the Western Cape [comprising a Rigaku RUH3R copper rotating-anode X-ray source operated at 40 kV, 22 mA; a Rigaku R-axis IV+ image plate camera; an X-stream 2000 low-temperature system; and an AXCO−PX50 glass capillary optic with a 0.1 mm focus]. Data for tACE-G1234 and an additional dataset for tACE-G13 were collected to 2.8 Å and 2.0 Å, respectively, on station PX14.1 of the synchrotron radiation source (Daresbury, U.K.) using a Quantum 4 CCD (Area Detection Systems, Poway, CA).

DENZO/SCALEPACK were used to process the 2.9 Å tACE-G13 data (*21*). Synchrotron data were processed and scaled using the HKL2000 software package, and data reduction was carried out using the CCP4 program TRUNCATE (HKL Research, Charlottesville, VA, refs *21*, *22*). The symmetry and systematic absences in all cases were consistent with the $P2_12_12_1$ space group with one protein molecule per asymmetric unit.

Structure Determination, Model Building, and Refinement. The tACE-G13 structure was solved twice, using independent methodology. The 2.9 Å dataset was solved by molecular replacement with *EPMR2.5* using protein atoms from the structure of minimally glycosylated wild-type tACE (PDB ID 1O8A) as a model (*23*). A correlation coefficient of 0.703 was obtained. *CNS* was used to perform energy minimization and *B*-factor refinement using bulk-solvent correction (*24*). 4.05% of reflections were reserved for R_{free} calculation. Small adjustments to side chains and loops were made, and carbohydrates and acetates added using *O 9.0.7*, against composite omit maps calculated by *CNS composite_omit_map* using torsion angle simulated annealing (starting temperature 1000 K; drop, 50 K per set; 5.0% omission) (*24*, *25*).

The 2.0 Å tACE-G13 dataset was solved by molecular replacement with *Refmac* through the *CCP4i* interface, again using the wild-type tACE structure as a model (*26*, *27*). 3% of reflections were kept aside for R_{free} calculation (*28*). Small alterations to side chains and a few loops were made, and carbohydrate chains and acetate molecules were added using *Coot* (*29*). Water picking was achieved with the implementation of the *ARP*/*wARP* module in *Refmac* and peak picking in *Coot* (*27*, *30*).

The structure of tACE-G1234 was solved by refinement of the 2.0 Å tACE-G13 structure with *CNS* (*24*). 8.4% of reflections were kept aside for R_{free} calculation (*28*). Small alterations to side chains and loops were made, and carbohydrates and acetate added using *Coot* (*29*). Water picking was done manually and using the peak picking option in *Coot* (*29*). Structure validation was carried out using *PROCHECK* and *SFCHECK* from the *CCP4* suite (*27*, *31*).

Rms deviations between structures were determined using the rms deviation function of *CNS* (*24*).

Identification of Hinge Regions in ACE2. The open and closed structures of ACE2 (ACE2o and ACE2c, PDB IDs 1R4L and 1R42, respectively) were compared in order to identify residues involved in the hinge-bending mechanism. Structure alignments and root-mean-square deviations for Cα atoms were computed using *ALIGN* (*32*). Hinge regions in ACE2 were defined as those residues to one side of which the Cα atoms of ACE2o deviated from those of ACE2c by >3 Å, and to the other side of which the structures were closely aligned. These residues were compared with the corresponding residues in tACE (PDB ID 1O8A), Ndom (PDB ID 2C6N), and AnCE (PDB ID 1J38), based on the sequence alignments of refs *33* and *13*.

Analysis of Temperature Factors. As an indication of structural order or disorder, the temperature factors (*B*-factors) of the 2.0 Å tACE-G13 structure were compared with those from ACE homologues. Structures of wild-type tACE, Ndom, human ACE2 (open and closed forms), and *Drosophila* AnCE were used for comparison (PDB entries 1O8A, 2C6N, 1R4L, 1R42, and 1J38, respectively). $\langle B \rangle$ and $\sigma(B)$ were calculated for all Cα atoms, and the *B*-factor of each individual Cα atom was expressed as a function of $\sigma(B)$.

Modeling of a Putative Open Form of tACE. Open and closed homology models of tACE (tACEo and tACEc) for NMA were built based on ACE2o and ACE2c. Amino acid sequences were aligned according to ref *17*, and 25 open and 25 closed models were generated using *MODELLER6v2* (*34*). *PROCHECK* was used to analyze the models. Based on their Ramachandran plots, one model was selected from each group as having the fewest residues falling into unfavorable regions. *SCWRL* was used to optimize the side-chain packing of these models (*35*).

Normal Mode Analysis. NMA of ACE2 and tACE was carried out according to the elastic network model, as implemented in the *elNémo* webserver (*36*, *37*). The block size was set to three residues, and the Cα−Cα linkage cutoff to 8 Å, being the recommended default values for these parameters.

The crystallographic structures of ACE2o, ACE2c (minus the collectrin-like domain) and unliganded wild-type tACE were analyzed, as well as models tACEo and tACEc. In each case, the 10 lowest-frequency normal modes were calculated, and the degree of cumulative overlap between the normal modes and the structural change was determined. Overlap is a measure of the degree of similarity between the direction of the observed conformational change and the one described by the normal mode in question, with a value of 1 indicating complete agreement between the normal mode and the observed change (*37*).

For the modes showing the highest overlap with the observed structural change, the amplitude of perturbation required for maximal agreement between the resulting structure and the comparison structure was determined, and perturbed models were generated accordingly. Cα−Cα distance fluctuations were calculated by *elNémo* and used as an indication of the residues involved in the motion described by the mode in question. The normal modes were also assessed in terms of collectivity, a measure of the proportion of atoms displaying large-amplitude displacements in the motion described by the mode (*36*).

RESULTS AND DISCUSSION

Structure of tACE Mutants. The crystal structure of tACE-G13 was determined to a final resolution of 2.0 Å with an *R* factor of 18.04% and an R_{free} of 22.01% (see Table 1). The quality of the map allowed rebuilding of residues 40 to 623, which includes an additional six C-terminal residues that were disordered in the wild-type structure. Two conformations were modeled for a few surface residues where $F_{obs} - F_{calc}$ maps clearly indicated a second conformation. The catalytic zinc ion and two chlorides could be modeled, as well as one *N*-acetyl-glucosamine residue at N72 and an ordered partial glycan chain at N109 (Figure 2). The structures solved independently using the 2.9 Å and 2.0 Å datasets are in close agreement (Table 2).

The structure of tACE-G1234 was determined to a final resolution of 2.8 Å, including residues 40 to 592. The catalytic zinc ion and both chloride residues were modeled, but ordered glycan density was only seen at two of the four glycosylation sites (N72 and N109), and a single *N*-acetyl-glucosamine residue was modeled at each of these residues.

Density for the 3_{10} helix 104−108 near the N109 glycosylation site and for loop 434−440 was weak in both mutants, and E436 could not be modeled. This pattern of loop disorder was also observed in the wild-type (*10*).

Like the wild-type structure, these mutants are largely α-helical, with a few short regions of β-strand. The molecules are ellipsoid in shape and divided down the middle by a deep cleft, in which the active site zinc ion is located. This cleft is closed off from the external milieu by the N-terminal two α-helices or lid-helices. Rms deviations computed between tACE-G13, tACE-G1234, and the crystal structures of unliganded and lisinopril-bound wild-type tACE revealed very close structural identity between these structures, especially around the active site zinc ion (Table 2). Each mutant structure contains two chloride ions occurring at the same sites as in wild-type tACE, and four of the five conserved water molecules proposed to represent an entrance pore for CL2 are also present in tACE-G13 (*38*, Figure 2A). The lower resolution of the tACE-G1234 data set prevented the detection of any water molecules at this site. The close structural agreement further indicates that the difference in glycosylation between tACE-G13, tACE-G1234, and wild-type tACE does not affect their structures in this crystal form.

Table 1: Data Collection and Model Refinement Statistics

	tACE-G13 (2.9Å)	tACE-G13 (2.0Å)	tACE-G1234
	Crystal Data		
space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
unit cell parameters (Å)			
a	59.81	56.63	56.44
b	85.16	84.72	85.15
c	135.58	134.47	133.42
molecules in asymmetric unit	1	1	1
	Diffraction Data Used in Refinement[a]		
resolution limits (Å)	50.0−2.90 (3.0−2.9)	50.0−2.0 (2.07−2.00)	50.0−2.8 (2.9−2.8)
no. of observations	60 751	515 005	541 512
redundancy	4.0	12.4	39.0
completeness (%)	89.9 (71.6)	96.2 (86.4)	83.6 (80.4)
R_{sym} (%)[b]	20.0 (41.3)	6.1 (21.4)	11.4 (33.7)
average $I/\sigma(I)$	5.88 (1.86)	26.8 (6.7)	7.7 (2.8)
	Refinement and Model Statistics		
no. of reflections used in refinement	14 338	40 225	13 572
R_{cryst}[c] [% reflections used]	21.62 [95.95]	18.04 [97.0]	20.06 [91.6]
R_{free}[d] [% reflections used]	24.29 [4.05]	22.01 [3.0]	23.69 [8.4]
mean *B*-factor (Å^2)			
protein	20.0	19.0	20.0
carbohydrate	46.0	50.4	58.6
water (no. of molecules)	14.8 (74)	25.0 (362)	15.3 (34)
rms deviations			
bond lengths (Å)	0.009	0.009	0.010
bond angles (deg)	1.57	1.12	1.46
Ramachandran plot % residues in			
most favored regions	88.5	94.8	91.6
additional allowed regions	10.5	5.0	8.2
generously allowed regions	1.0	0.2	0.2
disallowed regions	0.0	0.0	0.0

[a] Values in parentheses are for the last resolution shell. [b] $R_{sym} = \sum_h\sum_i|I_i(h) - \langle I(h)\rangle|/\sum_h\sum_i I_i(h)$, where $I_i(h)$ and $\langle I(h)\rangle$ are the *i*th 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*, respectively. [d] R_{free} is equal to R_{cryst} for a randomly selected 3% (tACE-G13 2.0 Å), 4.05% (tACE-G13 2.9 Å), or 8.4% (tACE-G1234) of reflections not used in the refinement.

N-Linked Glycans. Ordered N-linked glycan density was present only for sites 1 and 3 (N72- and N109-linked) in both mutants, with very little density being observed at sites 2 and 4 (N90- and N155-linked) in the tACE-G1234 structure, although this may be attributable to its lower resolution. The glycan chain at site 3 showed a particularly high degree of order in tACE-G13, allowing the positioning of two *N*-acetyl-glucosamine residues (NAG), two β-D-mannose residues, and one NAG-linked fucose residue (Figure 2B). O4 of the fucose residue is in hydrogen bonding contact (2.6 Å) with OE2 of E49 at the N-terminal end of α1, thus forming a glycan-mediated link between the N-terminus and the N-terminal end of α3 (Figure 2B). Glycan density for site 1, which lies in the loop between α1 and α2, was weaker, allowing the placement of only one NAG in both mutants. This NAG is in close proximity (3.2−4.5 Å) to the side chains of T74 and E76 at the C-terminal end of α1. It thus appears that these two more ordered glycan chains, lying as they do at either end of the lid helices, may play a role in stabilizing the interactions between the lid helices and the rest of the enzyme. This is consistent with the minimum requirement for the presence of an intact glycosylation site at either site 1 or site 3 in order for tACE to be properly expressed in an enzymatically active form (*14*).

Density in the Active Site. In both mutants, the active-site zinc ion is tetracoordinated by H383, H387, E411, and the carboxyl oxygen of an acetate molecule (Figure 2C). This acetate, also modeled in unliganded wild-type tACE, makes hydrogen-bonding contact with E384 of the zinc-binding motif, and superimposes onto the zinc-binding carboxyl group of the inhibitor lisinopril when aligned with the lisinopril-bound structure (PDB ID 1O8F). The active site of tACE-G13 contains an additional acetate molecule coordinated by K511, Q281, and Y520, that superimposes onto the carboxyl terminus of lisinopril. An actetate was also observed in this position in the unliganded N domain structure. In tACE-G1234, additional density at this site was modeled as an *N*-carboxyalanine moiety that was also present in the unliganded wild-type tACE structure.

The presence of additional density or acetate molecules in the active sites of both tACE mutants and the wild-type inhibitor-free structures of tACE and Ndom suggests that these may not be true unliganded conformations of ACE domains, but rather that the "unliganded" structures thus far observed are equivalent to another liganded state.

ACE2 Hinge Regions. To investigate potential domain hinging in ACE, the residues involved in the motion observed in ACE2 were identified and compared with the equivalent residues in tACE. The structural change between the closed and open forms of ACE2 (ACE2c and ACE2o) can be described as a hinge movement that opens up the active site. The lid helices α1 and α2 swivel to one side, and the cleft opens up more on one end than the other (Figure 3A). Looking down onto the active site cleft with the lid helices (α1 and α2) on top, the hinge axis stretches from under the N-terminus to the middle of the underside of the active site (Figure 3A).

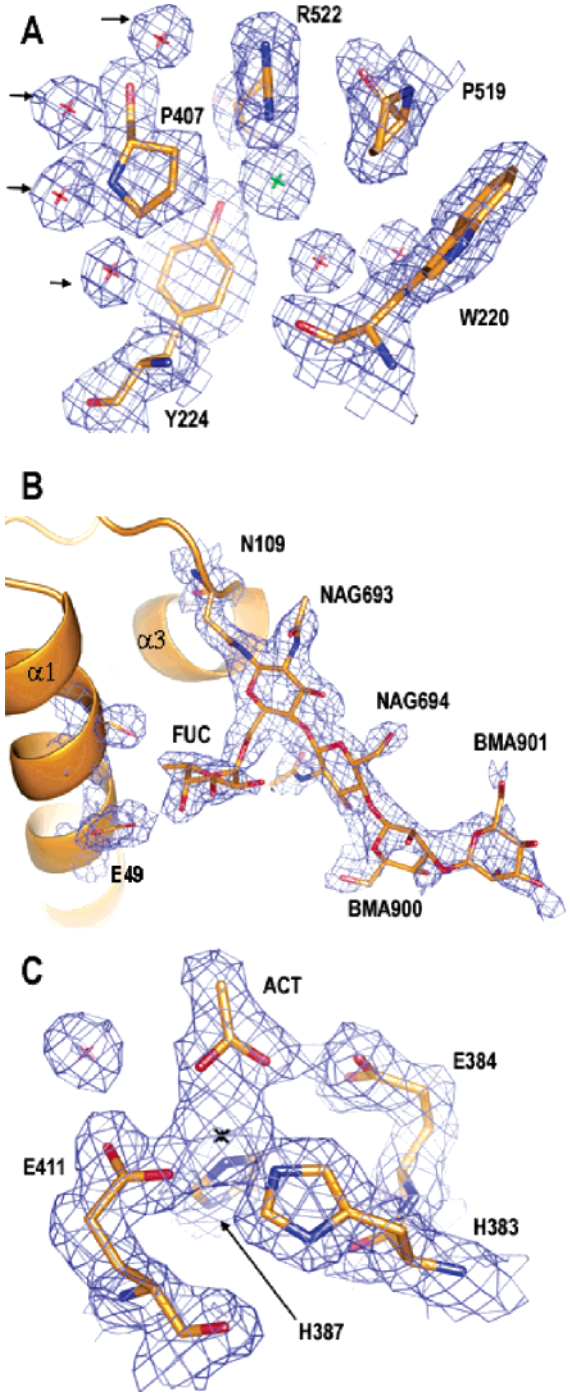

FIGURE 2: tACE-G13 in $2F_{obs} - F_{calc}$ map density to 2.0 Å, contoured at 1.0σ (blue mesh). (A) Residues at the CL2 binding site. Green star = CL2; red stars = water; arrows indicate putative CL2 pore waters. (B) Glycan residues at N109, with density for selected protein residues only. NAG = *N*-acetyl glucosamine; FUC = fucose; BMA = β-D-mannose; α1 and α3 = helices α1 and α3 in cartoon representation. (C) Zinc-binding residues at the active site. Black star = Zn; ACT = acetate. This figure was generated using *PYMOL 0.98* (DeLano Scientific).

Based on the structural alignment of ACE2o with ACE2c, six hinge regions were identified: 98−125, 296−297, 400−409, 434−439, 535−537, and 569−578 (tACE numbering; Figure 3A). These regions were compared with their equivalents in tACE, tACE-G13, tACE-G1234, Ndom, and *Drosophila* homologue AnCE in terms of sequence conservation and flexibility as evidenced by their crystallographic temperature factors (*B*-factors).

Based on their structural location these hinge regions can be divided into two groups. The first group comprises two loop regions (296−297 and 434−439; tACE numbering) that stretch one above the other, across the opening end of the active site cleft (Figure 3A). These residues are not "hinge" residues as such, but rather allow the active cleft to open by becoming elongated in the unliganded open state. The sequence of these residues is not conserved, but their temperature factors are high in all of the structures considered, indicating the functional conservation of a degree of flexibility (Figure 4; data for wild-type tACE and tACE-G1234 not shown). In the case of wild-type tACE the S435−S439 loop was disordered in the wild-type crystal structure and could not be modeled.

The second group comprises four regions (98−125, 400−409, 535−537, and 569−578; tACE numbering) all located along the hinge axis from the N-terminus side of the active site cleft to its underside (Figure 3A). This group could thus be described as the residues about which the hinge opens. Three of these regions (400−409, 535−537, and 569−578) show a high degree of sequence conservation, including several conserved glycines, alanines, and serines, and low temperature factors, indicative of a high degree of order in the crystal structures (Figure 4). The fourth (98−125) lies close to the surface between the α2 lid helix and α4, and has low sequence conservation and high temperature factors, indicating flexibility. 569−578 of this group lie between the conserved HEMGH of the zinc-binding motif (H383−H384 of tACE) and the additional conserved downstream zinc-binding Glu (E411 of tACE). Although hinging about this region does not alter the relative orientation of the zinc-coordinating residues, it does move the zinc ion away from the opposite wall of the active site cleft by ~3.8 Å, thus opening up the catalytic site.

Three of the six regions involved in hinging thus show low sequence-conservation and a high degree of disorder or flexibility in the crystal structures tACE, tACEG13, tACE-G1234, AnCE, ACE2, and Ndom as evidenced by their temperature factors. The other three regions display a high degree of order in the crystal structures and are highly conserved in the homologues considered. This indicates that the end points of the ACE2 hinge motion represent two stable conformations of these regions that have been captured under the crystallographic conditions. The presence of glycines, serines, and alanines in these regions may allow for more than one stable conformation. Moreover, their sequence conservation, together with the conservation of functional flexibility in the loop regions, suggests a conservation of the hinge mechanism among these homologues.

Conservation of Temperature Factors. The conservation of thermal stability or flexibility in these structures extends beyond the hinge regions, with the pattern of *B*-factors being conserved in tACE-G13, tACE-G1234, wild-type tACE, Ndom, ACE2, and AnCE structures (Figure 4; wild-type tACE and tACE-G1234 data not shown). Furthermore, residues having lower-than-average temperature factors (less than $-\sigma(B)$) are all located in the interior of subdomain II, surrounding the CL1 binding-site and including chloride and substrate ligands, whereas the residues having high temperature factors are all on the surface, and fall largely into subdomain I. The only residues in subdomain II that have high temperature factors are at the end of α4 (which flanks

Table 2: Structural Identity between tACE Glycosylation Mutants, Unliganded Wild-Type tACE (PDB ID 1O8A), and Lisinopril-Bound Wild-Type tACE (PDB ID 1O8F)

		rms deviations (Å)[a]					
structure	comparison structure	all protein atoms	main chain	side chain	Zn-binding residues	CL1-associated	CL2-associated
1O8F	1O8A	0.36	0.15	0.49	0.10	0.15	0.05
tACE-G13 (2.0 Å)	tACE-G13 (2.9 Å)	0.79	0.39	1.04	0.27	0.77	0.33
tACE-G1234	tACE-G13 (2.0 Å)	0.43	0.28	0.54	0.16	0.48	0.19
1O8F	tACE-G13 (2.0 Å)	0.53	0.69	0.27	0.14	0.40	0.14
1O8F	tACE-G1234	0.56	0.30	0.73	0.14	0.50	0.20

[a] Zn-binding residues: 383, 384, 387, and 411. CL1-associated residues 186, 278, 485, 486, 489, and 507. CL2-associated residues: 220, 224, 407, 519, 521, and 522.

Table 3: Normal Mode Analysis of tACE and ACE2. Statistics for the Mode Having the Highest Overlap with the Proposed Structural Transition for Each Structure Studied

structure analyzed	comparison structure	mode	overlap	collectivity	|amplitude| (DQ[a])	rmsd[b] (Å)
ACE2o	ACE2c	1	0.733	0.386	224	1.77
tACEo	tACEc	1	0.709	0.427	248	1.86
tACEo	wt tACE	1	0.643	0.427	245	2.19
ACE2c	ACE2o	2	0.117	0.132	85	3.43
tACEc	tACEo	6	0.133	0.500	96	3.62
wt tACE	tACEo	2	0.129	0.468	107	3.89

[a] The relative amplitude of perturbation along the direction of the normal mode concerned, required to achieve the best alignment with the comparison structure; arbitrary units. [b] Rms deviation for all Cα atoms between the comparison structure and a model perturbed according to the mode in question and the stated amplitude.

the lid helices), and in the C-terminal loop which was disordered in the wild-type tACE structure. Thus subdomain II, containing the zinc and chloride binding sites, has a greater degree of structural rigidity than subdomain I in all of the known structures of ACE homologues.

Since most of the residues involved in intermolecular contacts in tACE crystals lie on the surface of subdomain II, it could be argued that the higher thermal stability of this subdomain is an artifact of lattice packing. However, the other structures studied were crystallized in different space groups and under different conditions to tACE, indicating that this is more likely to be a genuine conservation, which may have functional significance.

Normal Mode Analysis Reveals Intrinsic Flexibility. To investigate the possibility of hinge-movement in tACE further, NMA of tACE and ACE2 was carried out. Since open structures of both were required for comparison, an open model of tACE (tACEo) was generated, based on the unliganded ACE2o, and a model of tACE (tACEc) based on ACE2c served as a control for errors introduced by the modeling process (Figure 3). Low frequency normal modes have been shown to depend more on overall shape or mass distribution than on atomic structure, with coarse-grained approaches such as that employed by *elNemo* being as effective as more detailed approaches at identifying biologically relevant modes (*39*, *40*). Thus despite the relatively low homology between tACE and ACE2 (40%), these models can be regarded as suitable for NMA.

NMA of both open structures tACEo and ACE2o yielded a single normal mode (mode 1, having the lowest frequency, in both cases) having a high degree of overlap with the putative or observed structural change (Table 3; Figure 5 A,E,F). In both cases, this mode has moderate collectivity

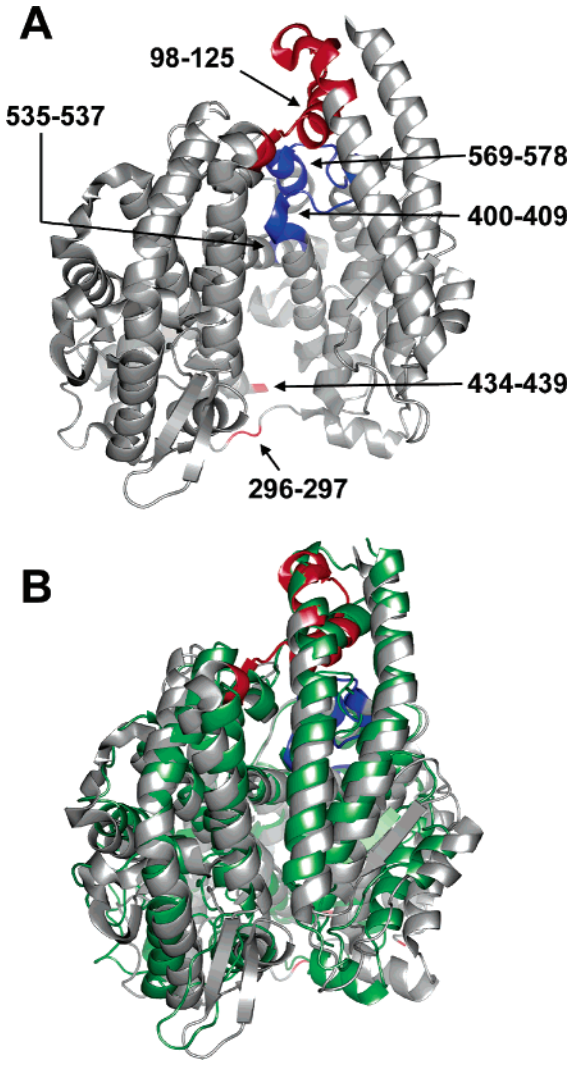

FIGURE 3: Hinging in the open model of tACE. (A) The open model of tACE (tACEo) used for normal mode analysis, showing ACE2 hinge residue equivalents. Hinge regions with low temperature factors and high sequence-conservation (400−409, 535−537, 569−578) are colored blue; those with high temperature factors and low sequence-conservation (98−125, 296−297, 434−439) are colored red. Note that S435−G438 were omitted from the model, as they were absent in the wild-type tACE structure. Deletion of the equivalent residues in ACE2 had no effect on the modes calculated (data not shown). (B) Alignment of the tACEo model perturbed in the direction of the lowest-frequency normal mode calculated (gray) with wild-type tACE (green). The amplitude of perturbation chosen is that which generated the best fit to the closed structure. Hinge residues in tACEo are colored as for the model in panel A. This figure was generated using *PYMOL 0.98* (DeLano Scientific).

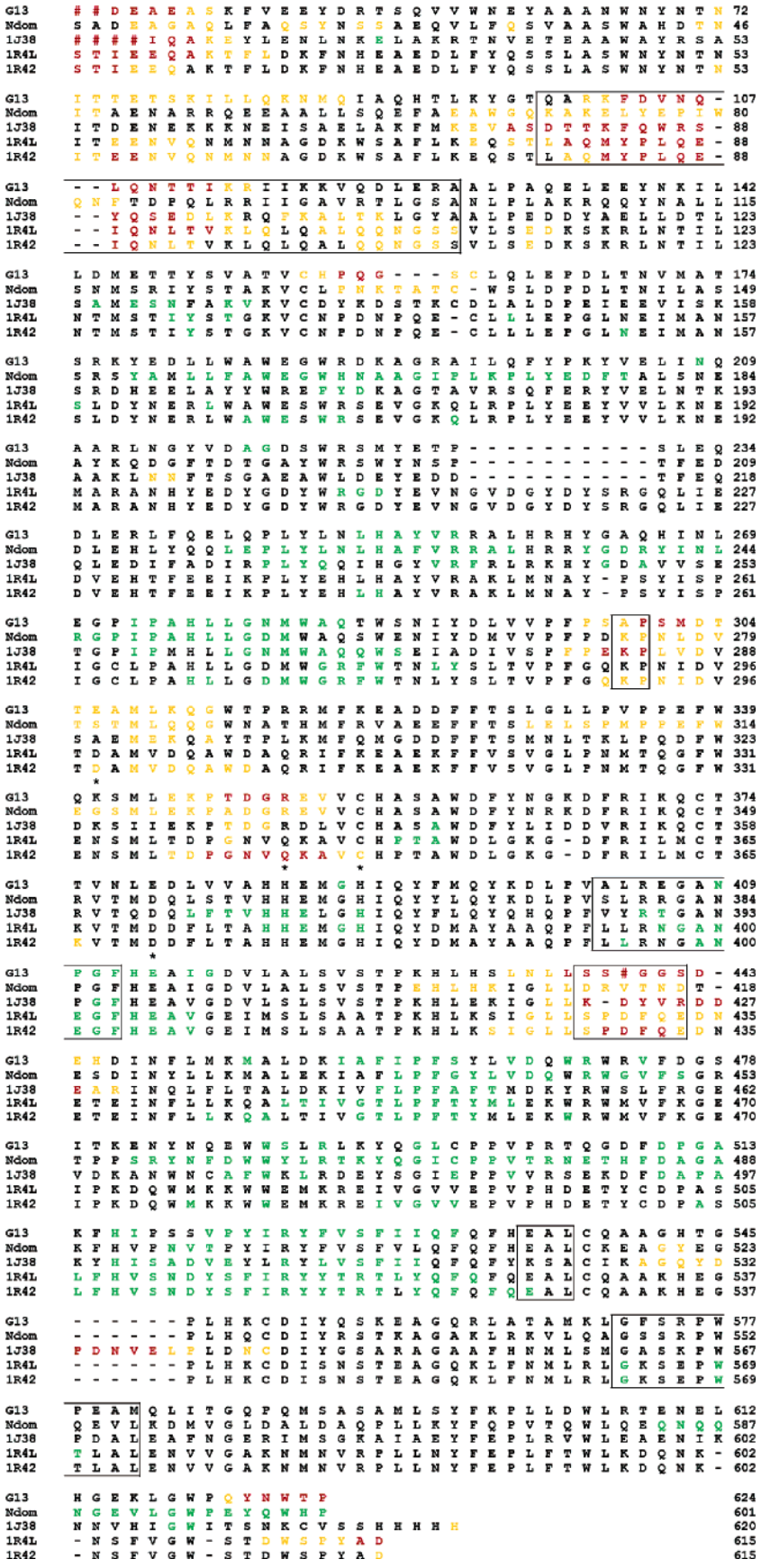

FIGURE 4: Temperature factor (*B*-factor) and sequence comparison between tACE-G13 (G13), the N domain of human sACE (Ndom, 2C6N), *Drosophila* AnCE (1J38), and the open (1R42) and closed (1R4L) structures of ACE2. Red hashes represent residues that were not modeled in the crystal structures due to disorder. Residues are colored according to their temperature factors: orange, *B*-factors greater than $\sigma(B)$; red, *B*-factors greater than $2\sigma(B)$; green, *B*-factors less than $-\sigma(B)$. Hinge regions in ACE2 are boxed; zinc ligands are marked with stars.

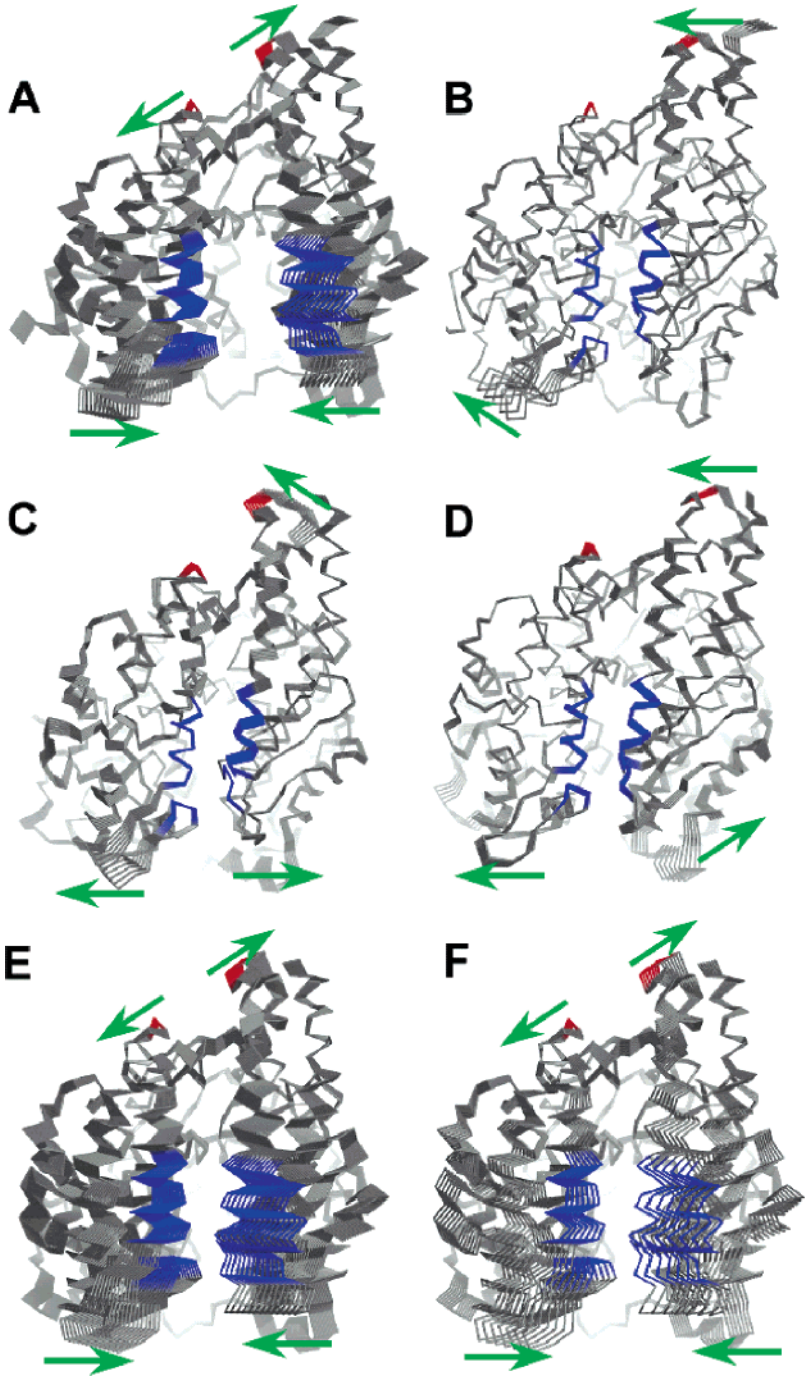

FIGURE 5: Normal mode analysis of tACE and ACE2: models perturbed according to the modes showing highest overlap with the putative structural change, the amplitude of perturbation being that which resulted in the closest alignment with the comparison structure. (A) ACE2o (vs ACE2c). (B) ACE2c (vs ACE2o). (C) wild-type tACE (vs tACEo). (D) tACEc (vs tACEo). (E) tACEo (vs wild-type tACE). (F) tACEo (vs tACEc). Colors are according to the residues showing the greatest relative displacements in ACE2o and tACEo: red residues move apart and blue residues move together as the cleft closes in the modes for these two structures. Green arrows indicate the major directions of movement in the mode in question. This figure was generated using *PYMOL 0.98* (DeLano Scientific).

and describes a closing of the active site cleft, with similar amplitudes of perturbation required for maximal overlap with the closed forms (Table 3). The residues on the underside of the active site cleft, including most of subdomain II, show little or no movement relative to neighboring residues. The motion corresponds closely to the hinging observed in the ACE2 crystal structures, as well as to that predicted by the tACEo model, as is evidenced by the low rms deviations between the resulting perturbed end-state models and the closed form comparison structures (Table 3).

It could be argued that the similarity in behavior of the normal modes of tACEo and ACE2o is an artifact due to the modeling of the open structure based on ACE2o. However, the close alignment of the perturbed model of tACEo with the wild-type tACE structure suggests that this model may represent a reasonable approximation of a real open form of tACE (Table 3; Figure 3B).

In contrast with these results, the closed structures (ACE2c, tACEc, and wild-type tACE) did not yield normal modes having high overlap with the structural change (Table 3;

Figure 5B,C,D). This is probably because the moving residues are close together in the starting structure, becoming linked in the force field so that large Cα−Cα fluctuations between them are not favored. This tendency of closed structures to yield normal modes that do not describe the structural transition as well as those of the corresponding open form has been documented for a number of cases (*40*). While the motions described by the normal modes showing the highest overlap do involve partial opening of the active site cleft, the amplitudes of displacement required for maximum agreement with the comparison structure are low, and the resulting structures do not align closely with the open forms (Table 3). However, since these results were similar for both wild-type tACE and ACE2c, they do not exclude the possibility that tACE might undergo a hinge movement. Rather, the fact that NMA of the wild-type structure of tACE does yield some normal modes having high collectivity is a further indication that some kind of concerted movement about the active site is likely to occur.

This evidence, taken together with the necessity of opening of the active cleft to allow substrate access, the entropically driven nature of substrate binding, the hinging of homologue ACE2, the presence of unexpected small molecules in the active sites of unliganded tACE-G13, tACE-G1234, wild-type tACE, and Ndom, and the conservation of the sequence or at least flexibility of the proposed hinge regions in these homologues, suggests that tACE and tACE homologues do have an open form and that this open form is similar to ACE2o. Moreover, since tACE is essentially identical to the C domain of sACE, this evidence points to a similar hinge motion in both domains of sACE.

CONCLUSIONS

Glycosylation mutants tACE-G13 and tACE-G1234 crystallize under the same conditions and in the same space group as wild-type tACE expressed in the presence of α-glucosidase-1 inhibitor NB-DNJ. The structures of the glycosylation mutants do not differ significantly from that of wild-type tACE, despite the presence of more complex glycan chains at the remaining glycosylation sites. It is thus possible to solve tACE structures without the use of expensive α-glucosidase-1 inhibitors, which in turn has important implications for future structural studies of ACE inhibitor binding. Elucidation of glycan chains at N72 and N109 revealed that interactions of glycan residues at these sites with adjacent protein residues may be important for the stabilization of the lid helices α1 and α2.

Based on analysis of temperature factors, subdomain II of tACE-G13, containing bound chloride ions and the zinc-binding site, displays a high degree structural rigidity while subdomain I appears to be more flexible. This domain stability is conserved in tACE-G1234, wild-type tACE, Ndom, ACE2, and AnCE and may have functional significance.

The residues involved in the hinge motion of ACE2 are conserved, at least at the level of functional flexibility, in tACE, Ndom, and AnCE. Normal mode analysis of the closed tACE structure and a modeled open form demonstrated that the intrinsic flexibility of this structure about the active site cleft is similar to that of ACE2. This suggests that hinging is a common mechanism for substrate entry in ACE homologues. This hypothesis is further supported by the observation that some kind of motion must occur to allow substrate access, the calorimetric evidence for a large entropic contribution to substrate binding, and the presence of unexpected small molecules in the active sites of the unliganded wild-type tACE, Ndom, and tACE glycosylation mutant structures (*10*, *16*, *18*).

REFERENCES

1. Skeggs, L. T., Jr., Kahn, J. R., and Shumway, N. P. (1956) The preparation and function of the hypertensin-converting enzyme, *J. Exp. Med. 103*, 295−299.
2. Acharya, K. R., Sturrock, E. D., Riordan, J. F., and Ehlers, M. R. (2003) ACE revisited: a new target for structure-based drug design, *Nat. Rev. Drug Discovery 2*, 891−902.
3. Zaman, M. A., Oparil, S., and Calhoun, D. A. (2002) Drugs targeting the renin-angiotensin-aldosterone system, *Nat. Rev. Drug Discovery 1*, 621−636.
4. Ehlers, M. R., and Riordan, J. F. (1989) Angiotensin-converting enzyme: new concepts concerning its biological role, *Biochemistry 28*, 5311−5318.
5. Ehlers, M. R., Fox, E. A., Strydom, D. J., and 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. U.S.A. 86*, 7741−7745.
6. 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, *Proc. Natl. Acad. Sci. U.S.A. 85*, 9386−9390.
7. Hubert, C., Houot, A. M., 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, *J. Biol. Chem. 266*, 15377−15383.
8. Ehlers, M. R., Chen, Y. N., and Riordan, J. F. (1992) The unique N-terminal sequence of testis angiotensin-converting enzyme is heavily O-glycosylated and unessential for activity or stability, *Biochem. Biophys. Res. Commun. 183*, 199−205.
9. Ehlers, M. R., Schwager, S. L., Scholle, R. R., Manji, G. A., Brandt, W. F., and Riordan, J. F. (1996) Proteolytic release of membrane-bound angiotensin-converting enzyme: role of the juxtamembrane stalk sequence, *Biochemistry 35*, 9549−9559.
10. Natesh, R., Schwager, S. L., Sturrock, E. D., and Acharya, K. R. (2003) Crystal structure of the human angiotensin-converting enzyme-lisinopril complex, *Nature 421*, 551−554.
11. Sturrock, E. D., Natesh, R., van Rooyen, J. M., and Acharya, K. R. (2004) Structure of angiotensin I-converting enzyme, *Cell. Mol. Life Sci. 61*, 2677−2686.
12. Yu, X. C., Sturrock, E. D., Wu, Z., Biemann, K., Ehlers, M. R., and Riordan, J. F. (1997) Identification of N-linked glycosylation sites in human testis angiotensin-converting enzyme and expression of an active deglycosylated form, *J. Biol. Chem. 272*, 3511−3519.
13. Corradi, H. R., Schwager, S. L., Nchinda, A. T., Sturrock, E. D., and Acharya, K. R. (2006) Crystal structure of the N domain of human somatic angiotensin I-converting enzyme provides a structural basis for domain-specific inhibitor design, *J. Mol. Biol. 357*, 964−974.
14. Gordon, K., Redelinghuys, P., Schwager, S. L., Ehlers, M. R., Papageorgiou, A. C., Natesh, R., Acharya, K. R., and Sturrock, E. D. (2003) Deglycosylation, processing and crystallization of human testis angiotensin-converting enzyme, *Biochem. J. 371*, 437−442.
15. Natesh, R., Schwager, S. L., Evans, H. R., Sturrock, E. D., and 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.
16. Kim, H. M., Shin, D. R., Yoo, O. J., Lee, H., and Lee, J. O. (2003) Crystal structure of Drosophila angiotensin I-converting enzyme bound to captopril and lisinopril, *FEBS Lett. 538*, 65−70.
17. Towler, P., Staker, B., Prasad, S. G., Menon, S., Tang, J., Parsons, T., Ryan, D., Fisher, M., Williams, D., Dales, N. A., Patane, M. A., and Pantoliano, M. W. (2004) ACE2 X-ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis, *J. Biol. Chem. 279*, 17996−18007.

18. Andujar-Sanchez, M., Camara-Artigas, A., and Jara-Perez, V. (2004) A calorimetric study of the binding of lisinopril, enalaprilat and captopril to angiotensin-converting enzyme, *Biophys. Chem. 111*, 183−189.
19. Ehlers, M. R., Chen, Y. N., and Riordan, J. F. (1991) Purification and characterization of recombinant human testis angiotensin-converting enzyme expressed in Chinese hamster ovary cells, *Protein Expression Purif. 2*, 1−9.
20. Friedland, J., and Silverstein, E. (1976) A sensitive fluorimetric assay for serum angiotensin-converting enzyme, *Am. J. Clin. Pathol. 66*, 416−424.
21. Otwinowski, Z., and Minor, W. (1997) [20] Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−326.
22. French, S., and Wilson, K. (1978). On the treatment of negative intensity observations. *Acta Crystallogr., Sect. A: Found. Crystallogr. 34*, 517−525.
23. Kissinger, C. R., Gehlhaar, D. K., and Fogel, D. B. (1999) Rapid automated molecular replacement by evolutionary search, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 55*, 484−491.
24. 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) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 54*, 905−921.
25. Jones, T. A. (1978) A graphics model building and refinement system for macromolecules, *J. Appl. Crystallogr. 11*, 268−272.
26. Murshudov, G. N. (1997). Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr. D53*, 240−255.
27. Collaborative Computational Project, N. 4. (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 50*, 760−763.
28. Brünger, A. T. (1992). Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures, *Nature 355*, 472−475.
29. Emsley, P., and Cowtan, K. (2004). Coot: Model building tools for molecular graphics, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60*, 2126−2132.
30. Perrakis, A., Morris, R. M., and Lamzin, V. S. (1999). Automated protein model building combined with iterative structure refinement, *Nat. Struct. Biol. 6*, 458−463.
31. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thronton, J. M. (1993). PROCHECK−A program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
32. Cohen, G. E. (1997) ALIGN: A program to superimpose protein coordinates, accounting for insertions and deletions, *J. Appl. Crystallogr. 30*, 1160−1161.
33. Guy, J. L., Jackson, R. M., Acharya, K. R., Sturrock, E. D., Hooper, N. M., and Turner, A. J. (2003) Angiotensin-converting enzyme-2 (ACE2): comparative modeling of the active site, specificity requirements, and chloride dependence, *Biochemistry 42*, 13185−13192.
34. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol. 234*, 779−815.
35. Canutescu, A. A., Shelenkov, A. A., and Dunbrack, R. L., Jr. (2003) A graph-theory algorithm for rapid protein side-chain prediction, *Protein Sci. 12*, 2001−2014.
36. Suhre, K., and Sanejouand, Y. H. (2004) ElNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement, *Nucleic Acids Res. 32*, W610−W614.
37. Tama, F., Gadea, F. X., Marques, O., and Sanejouand, Y. H. (2000) Building-block approach for determining low-frequency normal modes of macromolecules, *Proteins: Struct., Funct., Genet. 41*, 1−7.
38. Tzakos, A. G., Galanis, A. S., Spyroulias, G. A., Cordopatis, P., Manessi-Zoupa, E., and Gerothanassis, I. P. (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 Eng. 16*, 993−1003.
39. Lu, M., and Ma, J. (2005). The role of shape in determining molecular motions, *Biophys. J. 89*, 2385−2401.
40. Ma, J. (2005). Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes, *Structure 13*, 373−380.

BI061146Z