Dimerization and dynamics of angiotensin-I converting enzyme revealed by cryo-EM and MD simulations

Jordan M. Mancl[1], Xiaoyang Wu[1], Minglei Zhao[2], Wei-Jen Tang[1]

1. Ben-May Department for Cancer Research, The University of Chicago, Chicago, Illinois, United States

2. Department of Biochemistry and Molecular Biology, The University of Chicago, Illinois, United States

Keywords: Angiotensin-I converting enzyme, hypertension, amyloid peptide, cryo-EM, all-atom MD simulations.

Author contributions: J.M.M. and W.-J.T., designed the project. X. W. transduced and selected ACE-expression HEK293F cells. J.M.M. and W.-J. T. performed the expression and purification of ACE. J.M.M. performed cryo-EM grid preparation, data acquisition and processing assisted by M.Z. and overseen by W.-J. T., and M. Z. J.M.M. built and refined cryo-EM structural models and performed heterogeneity analysis. J.M.M. and W.-J. T. performed all-atom MD simulations and analysis. J.M.M. and W.-J.T. wrote the manuscript and all authors contributed to manuscript finalization.

Acknowledgement: We are grateful to Mahira Aragon for grid preparation using Chameleon, Kasahun Neselu for data collection, and Marc Auréle Gilles for assistance with RECOVAR analysis. This work was supported by the NIH grant GM121964 to W.-J. Tang. Some of this work was performed at the National Center for Cryo-EM Access and Training (NCCAT) and the Simons Electron Microscopy Center located at the New York Structural Biology Center, supported by the NIH Common Fund Transformative High Resolution Cryo-Electron Microscopy program (U24 GM129539,) and by grants from the Simons Foundation (SF349247) and NY State Assembly.

Abstract

Angiotensin-I converting enzyme (ACE) regulates the levels of disparate bioactive peptides, notably converting angiotensin-I to angiotensin-II and degrading amyloid beta. ACE is a heavily glycosylated dimer, containing 4 analogous catalytic sites, and exists in membrane bound and soluble (sACE) forms. ACE inhibition is a frontline, FDA-approved, therapy for cardiovascular diseases yet is associated with significant side effects, including higher rates of lung cancer. To date, structural studies have been confined to individual domains or partially denatured cryo-EM structures. Here we report the cryo-EM structure of the glycosylated full sACE dimer. We resolved four structural states at 2.99 to 3.65 Å resolution which are primarily differentiated by varying degrees of solvent accessibility to the active sites and reveal the full dimerization interface. We also employed all-atom molecular dynamics (MD) simulations and heterogeneity analysis in cryoSPARC, cryoDRGN, and RECOVAR to elucidate the conformational dynamics of sACE and identify key regions mediating conformational change. We identify differences in the mechanisms governing the conformational dynamics of individual domains that have implications for the design of domain-specific sACE modulators.

Introduction

Angiotensin-I converting enzyme (ACE) is an M2 clan zinc metalloprotease (EC3.4.15.1) located on the plasma membrane and in the extracellular milieu (1-3). ACE plays a crucial role in the renin-angiotensin system by converting angiotensin I to the potent vasoconstrictor angiotensin II, and in the kinin-kallikrein system by degrading bradykinin, a potent vasodilator. Inhibiting ACE is a primary approach for treating hypertension and is a proven therapy for heart failure, diabetic nephropathy, and other cardiovascular and renal dysfunctions (4). However, ACE inhibition has known side effects, e.g., dry cough and angioedema, and has been linked to higher rates of lung cancer (5, 6). This is partly because ACE also degrades a diverse range of peptides, resulting in complex physiological outcomes due to the interplay in altering levels of these peptides (4, 7-10). Furthermore, ACE degrades amyloid β (Aβ), which is associated with the progression of Alzheimer's diseases (11-14). Overexpression of ACE in myelomonocytes substantially reduced Aβ load and prevented Alzheimer's disease-like cognitive decline in mice (15). However, the brains of some Alzheimer's disease patients showed upregulation of the renin-angiotensin system which led to deleterious effects and ACE inhibition was beneficial (16-18). These studies indicate that global ACE inhibition could be either beneficial or harmful for individual Alzheimer's patients, and underscores a need for better understanding the molecular mechanisms governing ACE function and substrate selectivity to develop more selective, improved, ACE inhibitors and expand ACE-based therapies.

ACE is a 180 kDa heavily N-linked glycosylated type 1 transmembrane protein with two homologous catalytic domains residing at the N- and C-terminal ends (ACE-N and ACE-C) in its extracellular region (1, 3, 19). The ACE extracellular region can be released via proteolytic cleavage near the transmembrane region by ADAM (a disintegrin and metalloproteinase) family protease(s), generating soluble ACE (sACE) and ACE dimerization can regulate this process (20-22). ACE belongs to the cowrin (cowry like) family of zinc metallocarboxypeptidases because its catalytic domain is ellipsoidal and contains a long, deep, yet narrow catalytic cleft (23). ACE is known to cleave short peptides, such as angiotensin I (decapeptide) and bradykinin (nonapeptide) by removing the C-terminal two amino acid residues. Different from the funnelin family of metallocarboxypeptides that contains a funnel-like cavity, the catalytic cleft of ACE can

accommodate entire oligopeptides and ACE has been shown to cleave multiple discrete sites on Aβ outside the C-terminal end (13). Thus, it is more accurate to refer to ACE as an endopeptidase with dipeptidyl carboxylase activity.

Extensive crystallographic studies of ACE-N or ACE-C alone and in complex with various inhibitors and a catalytic product, angiotensin-II, have elucidated the overall structures of ACE catalytic domains and molecular details of the catalytic pockets (2, 24-28). The structures indicate that each domain is comprised of three sub-domains and offer insight for the substrate preferences around cleavage sites and the promiscuity of each domain toward short peptide substrates. Furthermore, they reveal the orientation of substrate binding sites around scissile bond, indicating that the entire oligopeptide substrates need to be engulfed into the catalytic chamber of ACE catalytic domain (24, 26, 27).

However, the catalytic pocket of ACE in the closed state is not large enough to accommodate the entrance of all its known substrates, particularly Aβ, to permit the observed cleavage pattern. Thus, similar to other Aβ-degrading proteases, such as insulin degrading enzyme (IDE), presequence protease (PreP), and neprilysin (NEP), ACE must undergo a large-scale open-closed transition to capture and degrade larger peptide substrates such as Aβ (29-31). Recent open state structures of sACE-N, sACE monomer, and a sACE-N dimer, along with molecular dynamics (MD) simulations of sACE-C, have begun to reveal ACE conformational dynamics, though they remain under-studied (32-34).

The extracellular domains of ACE are readily dimerized in the membrane-bound and soluble forms. Both the ACE-N and ACE-C domains within the ACE monomer are catalytically active and each has its own substrate preferences (7, 13, 35). For example, ACE-N can convert Aβ1-42 to Aβ1-40, which is less amyloid fibril-prone than Aβ1-42 but ACE-C cannot (13). Accumulating evidence support a complicated interplay between ACE-N and ACE-C within the ACE dimer. Enzyme kinetics analysis suggests negative cooperativity between two catalytic domains (36-38). However, ACE also exhibits positive synergy toward Aβ cleavage and allostery to enhance the activity of its binding partner, bradykinin receptor (13, 39). Recently, cryo-EM structures of soluble dimeric apo-sACE were reported (33); however, due to the poor resolution/denaturation of sACE-C, only full-

105 length sACE monomer or dimeric sACE-N was reported yet the conformations preclude
106 full-length sACE dimerization due to the severe steric crash between sACE-C domains.
107 Here, we present four discrete cryo-EM structures of full-length sACE that contain all four
108 catalytic domains (Fig. 1A). This is followed by cryo-EM heterogeneity analysis and all-
109 atom MD simulations. Together, our analyses provide structural insights into domain
110 interactions, open-closed transitions of the catalytic domain, and conformational
111 dynamics of soluble dimeric ACE.

112 **Results**

113 **Structures of the full-length sACE dimer determined by single particle cryo-EM**

114 Soluble human ACE (sACE), which includes the entire extracellular region of ACE,
115 was expressed in HEK293F cells and purified from culture supernatant using an ion
116 exchanger (Source Q) followed by size exclusion chromatography (Superdex S200).
117 Buffer conditions suitable for single particle cryo-EM analysis were identified and
118 optimized using differential scanning fluorometry (Figure S1, Table S1). Purified sACE
119 was then vitrified using a Vitrobot, and a dataset of ~3,600 micrographs was collected on
120 a 300 keV Titan Krios equipped with a Gatan K3 camera at the University of Chicago
121 Advanced Electron Microscopy Facility (Table S2). The data was processed in
122 cryoSPARC v4.4.0 (Figure S2). After filtering ~1.1 million picked particles to ~378K using
123 2D and 3D classification, three classes were identified (Figure 1, Figure S2). Of these,
124 ~268K particles mapped to a dominant 2-domain class reconstructed as an sACE-N
125 dimer, while ~21K particles mapped to a minor 3-domain class containing two sACE-N
126 and only one sACE-C, reminiscent of the recently published sACE cryo-EM structure (33).
127 The remaining 89K particles mapped to a full-length sACE dimer containing two sACE-N
128 and two sACE-C, which was refined to 3.65Å resolution (Figure 1C, S2, S3, Movie S1,
129 Table S2).

130 The partial denaturation of macromolecules during vitrification has been attributed
131 mainly to repeated exposure of the macromolecules to the air-water interface (40, 41).
132 The sACE-C domain was shown to be less stable than the sACE-N domain (42), so we
133 hypothesized that repeated exposure to the air-water interface could cause preferential
134 denaturation of sACE-C. We have previously showed that preferential denaturation of
135 human PreP occurred in one of its two homologous domains and that faster vitrification

using Chameleon significantly prevented such preferential denaturation (29). Therefore, we collaborated with the National Center for CryoEM Access and Training (NCCAT) to prepare sACE grids using Chameleon (43). Chameleon preparation reduced the time the sample spends on the grid before freezing by more than 10-fold. We collected a new dataset of ~18,000 micrographs and our 2D classification revealed a higher percentage of particles for four-domain classes than those prepared using the Vitrobot (Figure 1B). However, our 3D classification still showed the existence of structures of the sACE-N only dimer and those containing two sACE-N and one sACE-C (Figure S4). This indicates that the reduction in vitrification time using Chameleon reduced but did not prevent preferential denaturation.

From these data, we obtained a map for a consensus ACE structure that achieved a resolution of 2.8Å. However, while the core region showed excellent map quality, the map quality for the dynamic regions corresponding to the open-close transition of ACE catalytic domains was poor, suggesting multiple discrete conformational states exist. We thus performed 3D classification and reconstructed three distinct conformational states of full-length apo-ACE at resolutions of 2.99Å, 3.05Å, and 3.15Å (Figure 1C, S4, S5, Movie S2-4, Table S2). We refer to these ACE dimer structures by their resolutions, namely sACE-2.99, sACE-3.05, and sACE-3.15, along with the structure from data obtained using the Vitrobot, sACE-3.65. Given the rich structural information from more than 50 reported crystal structures of ACE and one recently reported cryo-EM study of ACE, below we focus on the new insights that our sACE dimer structures offer (2, 24, 26, 33, 44, 45).

The overall structures of sACE dimers and structural insight into dimerization

Soluble ACE has two homologous catalytic domains: sACE-N (amino acids 1-615) and sACE-C (amino acids 616-1202). Based on the comparison of various catalytic domains of our ACE dimer structures, we subdivide each catalytic domain of sACE into three subdomains: D1a, D1b, and D2 for sACE-N, and D3a, D3b, and D4 for sACE-C (Figure 1A). These three subdomains in either sACE-N or sACE-C form a catalytic chamber. The catalytic zinc ion is coordinated by the canonical HEXXH motif residing in the D1b and D3b subdomains (Figure 1A). Four sACE dimer structures manifest as pseudo-symmetric C2 homodimers across the dimer interface. The dominant symmetry-breaking features are the degrees of displacement between D1 and D2 or between D3 and D4, leading to changes in solvent accessibility to the catalytic pockets.

168 Based on the openness of the catalytic chamber, defined as the distance between
169 the D1/3a and D2/4 helices that gate access, we classified sACE-N into three
170 conformational states: open (O), intermediate open (I), and closed (C) (Figure 2A); sACE-
171 C is significantly less open than sACE-N in our structures, and is only classified into
172 intermediate (I) and closed (C) states (Figure 2B). The four sACE dimer structures differ
173 mostly by the degrees of openness of their four catalytic domains (Figure 2C, D). Our
174 sACE-2.99 and sACE-3.65 structures are globally similar, both contain sACE-N domains
175 in the open and intermediate states and sACE-C domains in the closed and intermediate
176 states. In sACE-3.05, both sACE-N domains adopt an open state, while the sACE-C
177 domains adopt an intermediate and closed state. In the sACE-3.15 structure, one subunit
178 has both sACE-N and sACE-C adopting a closed conformation, while the other subunit
179 has an open sACE-N domain and an intermediate sACE-C domain.

180 Our four cryo-EM structures of sACE provide the structural basis for sACE
181 dimerization. sACE dimerization is mediated by interfaces between D2 subdomains and
182 D4 subdomains of sACE monomers. The D2 and D4 subdomains generally have much
183 lower B-factors, and higher local resolution estimates, than the D1 and D3 subdomains
184 (Figure S3, S5, S6). The interaction of the D4-D4 subdomains is analogous to the D2-D2
185 interaction, yielding an RMSD of ~1 Å² when aligned (Figure 3A). The primary source of
186 variation is a general rigid body shift away from the interaction interface for non-interacting
187 regions in the D4-D4 region compared to the D2-D2 region. The observed D2-D2 interface
188 is consistent with one of three possible contacts between sACE-N in the non-
189 crystallographic symmetry mates from previous crystallographic analysis of sACE-N and
190 the previously published cryo-EM structure (Figure S7). In total, the interface between
191 sACE monomers buries a surface area of approximately 1,900 Å². Although sACE-N and
192 sACE-C are structurally analogous, their respective interfaces (sACE-N/N and sACE-
193 C/C) are not equivalent. The sACE-N/N interface is larger and stronger, burying a surface
194 area of approximately 1,150 Å², compared to a buried surface area of approximately 750
195 Å² for the sACE-C/C interface.

196 The interface between sACE-N is formed by a hydrophobic core, anchored by a
197 Y465 π-stacking interaction, and flanked by salt bridges formed by R453 and R458 as
198 well as E212 and E219 along with a collection of polar and van der Waal contacts, which
199 is consistent with the previous mutational studies of Y465 (Figure 3B) (46). The dimer

interface between two sACE-N domains is completed by glycan-glycan interactions between the glycans attached to N82, which is consistent with the role of glycosylation in sACE dimerization (Figure 3C) (47). However, the central residues that form the hydrophobic core of the interface between sACE-N are not present in the interface between sACE-C. Instead, the sACE-C interface is formed by a symmetric salt bridge between K1067 and E1060, side chain-backbone interactions, and van der Waals contacts (Figure 3D).

Conformational states of sACE catalytic domains

The comparison of our dimeric sACE cryo-EM structures reveals the conformational dynamics of sACE catalytic domains. The four sACE dimer structures represent different combinations of open, intermediate, and closed states of sACE catalytic domains (Figure 2). By aligning each sACE-N and sACE-C region in our structures to the D2/4 domain, it becomes clear that our structures represent a continuous gradient of structural heterogeneity, ranging from the most closed state to the most open state (Figure S8).

We performed a series of all-atom molecular dynamics (MD) simulations to explore the conformational dynamics of sACE. A total of 8 apo-sACE simulations were conducted, comprising four simulations with non-glycosylated sACE (>1.8 μs each) and four with sACE where 12 N-linked glycosylation sites that contained observable extra density were modeled with complex glycans (>2 μs each) (Movie S5, S6). Glycan orientations were randomized prior to each simulation to avoid bias. We did not observe a substantial difference in the global conformational dynamics between our sets of simulations with and without glycans. The domain motions of these simulations are summarized in Table S3.

RMSD analysis reveals that the dominant source of conformational change in our simulations is the open-close transition between the D1/3 and D2/4 domains (Figure S9). The D2/D2 and D4/D4 regions are the most stable, with the bulk of conformational changes localized to the D1/D3 domain. Within the D1/3 domain, we observed that the D1/3a motion was not always correlated with D1/3b or D2/D4, meaning that at a given time, D1/3a could move along with D1/3b, or with D2/D4, or independently. Consistent with our structural analysis, our MD simulations also indicate that sACE-N displays a much wider range of opening geometries than sACE-C.

Most of our MD simulations rapidly closed, and remained closed for the duration of the simulation, preventing an effective comparison of the open vs closed state dynamics. However, we were able to garner significant insight into the mechanism of open-close transition by integrating our MD simulations and cryo-EM structural analysis. Initially, we sought to understand the stable and mobile regions of sACE by calculating the Cα displacement for each residue among the four difference cryo-EM structures. As shown in Figure 4A, the D2 domain was found to be the most stable region of sACE while the D1 domain was the most mobile, highlighting the impact of the sACE-N open-close transition. sACE-C was, globally, very stable outside of surface exposed loops and the tip of the D3a subdomain. This analysis predicted the most mobile regions of sACE to be the ends of the D1a subdomain and a loop region with the D1b that closely associates with the tip of the D1a subdomain.

While the structural perturbations predicted by comparing our cryo-EM structures correlate well with the global RMSD analysis of our MD simulations (Figure S9), we performed further analysis to understand the finer details of sACE conformational dynamics. To best compare the cryo-EM structures with our MD simulations, we compared the calculated alpha carbon displacement of our cryo-EM structures to the root mean square fluctuation (RMSF) values for each residue in our MD simulations (Figure 4). We observed consistent, global agreement between our cryo-EM data and MD simulations, both with and without glycans present. This consistency is particularly noticeable in several key mobile regions (denoted 1-4 in Figure 4) that demonstrate high degrees of Cα displacement from our cryo-EM data along with correlating higher RMSF during subsets of our MD simulations (Figure 4A, B, D). The RMSF variation between stable and mobile regions in our MD simulations is dampened, relative to the cryo-EM data when the entire trajectories were used (Figure 4A, C, E). This is due to the open state regions closing relatively early in the simulation and remaining closed for the duration of the simulation. We then performed principal component analysis (PCA) (Figure S10) and cluster analysis (Figure S11) to compare our MD simulation data with our cryo-EM data. Both methods of analysis indicate that the conformational space sampled by our cryo-EM structures represents a fraction of the conformational space sampled by our MD simulations.

262 The dynamics observed in our MD simulations of sACE are highly complex, arising
263 from a combination of both interdomain and intradomain motions. To summarize these
264 simulations in simpler terms, individual sACE domains transition from a more open state
265 to a closed state without reopening, while sACE-N/N and sACE-C/C primarily move
266 relative to each other as rigid bodies. As outlined previously, PCA was employed to reduce
267 the complexity of these motions and to identify dominant trends in conformational
268 changes within the dataset. Visualizing the conformational changes captured by the
269 principal component vectors in the MD simulations (Movies S7, S8) revealed that the
270 majority of conformational variability within the MD data is driven by substantial
271 interdomain motion occurring between sACE-N/N and sACE-C/C. Consistent interdomain
272 motion was observed along all principal component vectors, regardless of the presence
273 or absence of intradomain motion. However, no principal components were identified that
274 exclusively described significant intradomain motion without associated interdomain
275 motion. The PCA further revealed that glycosylation significantly enhances the degree of
276 interdomain motion in sACE, as evidenced by the glycosylated simulation (Movie S8)
277 compared to the non-glycosylated simulation (Movie S7). This suggests that glycosylated
278 sACE exhibits greater global dynamics (Figure 4, S12). Notably, principal components 5
279 and 6 from the glycosylated MD simulations (Movie S8) highlight a striking difference,
280 where sACE-N/N and sACE-C/C appear to move independently and exhibit a decoupled
281 interaction. We postulate that this inter-domain motion may be the reason we observe
282 preferential denaturation of sACE-C during cryo-EM grid preparation. The relatively harsh
283 conditions that the protein is exposed to during the grid-making process may damage
284 specific conformations of particles and subsequently restrict the amount of information
285 that could be obtained solely from cryo-EM analysis. This highlights the importance of
286 integrating multiple approaches to investigate protein dynamics.

287 The MD simulations of glycosylated sACE provides additional structural insights
288 into the structure and functions of N-linked glycosylation of sACE. Glycosylation is shown
289 to affect the folding, secretion, and stability of ACE (22, 48, 49). Our MD simulation reveals
290 that the dynamics of 12 N-linked glycans cover a large surface area of sACE, shielding it
291 from environmental insults (Figure S13). Consistent with observations from our Coulomb
292 potential density maps and previously published data, we observed a substantial and
293 persistent inter-subunit interaction between glycans from N82 (Figure 3C). From RMSF

294 analysis, glycosylation was found to increase the overall dynamicism of sACE (Figure
295 S12). Furthermore, while there are N-linked glycans near the catalytic chamber,
296 particularly those at the tip of D1a and D3a, they were highly dynamic and did not occlude
297 access to the catalytic chamber. This is consistent with the observation that
298 deglycosylation only modestly increases the enzymatic activity of ACE (50-52).

299 The comparison of sACE-N in its open and closed states reveals the structural
300 basis for the open-to-closed transition of this domain. Our structures show that when the
301 bottom of the D1a subdomain (region of interest 2 in figure 4) moves away from the D2
302 subdomain, resulting in the open state, the top of the D1a subdomain (region of interest
303 1 in figure 4) moves closer to the D2 subdomain, as if the helix were balanced upon a
304 fulcrum (Figure 5A, B). Detailed structural analysis indicates that the fulcrum is primarily
305 formed by the salt bridge between two highly conserved residues: E66 in the D1a
306 subdomain and R108 in the D2 subdomain. To stabilize the open state, additional salt
307 bridges are formed, with D189 in the D2 subdomain pairing with K73 in all cases and
308 occasionally with R96 in the D1a subdomain (Figure 5A). The loss of this salt bridge in
309 the closed state is compensated for by the formation of a network of van der Waals
310 interactions between the D1a and D2 subdomains that stabilize the closed state. The
311 closed state of sACE-C is also stabilized by a network of van der Waals contacts between
312 the D3a and D4 subdomains. The transition between open and closed states is likely
313 influenced by the binding of substrates and reaction products during the ACE catalytic
314 cycle. Although a similar fulcrum mechanism is not evident in the sACE-C domain, the
315 connection between the D3a and D2 subdomains would act as a lever to pivot sACE-C
316 from the closed to the open state, and we speculate that this may form the structural basis
317 for allosteric communication between sACE-N and sACE-C.

318 On average, sACE-N is more open than sACE-C, likely due to multiple factors
319 (Figure 2D). In sACE-N, the D1a subdomain is at the N-terminal end of sACE, allowing
320 for a greater range of motion (Figure 5B), while the fulcrum and stabilization mechanisms
321 can maintain sACE-N in a more widely open state (Figure 5A). Conversely, in sACE-C,
322 the motion of the D3a subdomain is constrained by its connection to the D2 subdomain
323 (Figure 5C). Additionally, there are sequence differences at the interface between the
324 three subdomains of sACE-N and sACE-C. In the most closed state—the B chain sACE-
325 C domain of the sACE-3.05 structure—residues from all three subdomains, including

326 I649, V724, V926, and V927, form a hydrophobic latch that stabilizes the closed state
327 (Figure 5D). The hydrophobic latch in sACE-N is weaker because the residue
328 corresponding to V724 has been replaced with a threonine (T124) (Figure 5D). These
329 structural insights are supported by our RMSF analysis of our MD simulations and alpha
330 carbon displacement analysis of our cryo-EM structures, which reveal a correlation
331 between mobile regions and the "fulcrum" mechanism of opening for sACE-N (Figure 4,
332 5A). Furthermore, our slowest closing MD simulation revealed that the K73-D189
333 interaction was maintained for the entire time that sACE-N adopted an open
334 conformation, yet this interaction was lost when sACE-N transitioned to the closed state
335 (Figure 5E).

336 We also performed three MD simulations on the non-glycosylated sACE monomer
337 (>1 microsecond each) (Table S3). We observed the expected intra-subdomain open-
338 closed motions in sACE-N and sACE-C that were also observed in the simulations of ACE
339 dimer. We also observed that the motion of sACE-C relative to sACE-N was significantly
340 greater in our monomer simulations compared to our dimer simulations. This suggests
341 that dimerization places significant constraints upon the conformational dynamics of
342 sACE. Previous work has defined the conformational dynamics of sACE from cryo-EM
343 data where both sACE-C domains were not resolved (30). Given our observations from
344 the monomeric and dimeric MD simulations, we endeavored to understand the
345 conformational heterogeneity present within our cryo-EM data to better understand the
346 impact of dimerization on sACE dynamics.

347 **Cryo-EM heterogeneity analysis of sACE**

348 Cryo-EM structures are ensemble averages of many aligned particle images.
349 Recently, there have been many approaches developed to understand the particle
350 heterogeneity present within a cryo-EM structure as a proxy for molecular motion (53).
351 We employed three methods of heterogeneity analysis: cryoDRGN, RECOVAR, and 3D
352 Variability Analysis (3DVA) (54-56). CryoDRGN employs neural networks to resolve
353 structural heterogeneity in a nonlinear, multidimensional latent space, while RECOVAR
354 and 3DVA both employ linear subspace methods. 3DVA finds the linear subspace by an
355 alternating optimization scheme to capture image variations while RECOVAR utilizes a
356 regularized covariance estimator to identify subspaces that best capture the distribution
357 of states.

We conducted this analysis and observed global agreement between the results of both of our datasets. As we observed substantial particle damage in our Vitrobot-prepared dataset, we thought it unwise to utilize that dataset for in-depth analysis and therefore, we focus on the larger, and higher-resolution dataset collected at NCCAT below. For analysis, the entire particle set, prior to 3D classification and consisting of ~170k particles, was utilized. The particle alignments from the same C1 consensus refinement were used as input for each of the three methods as current methods to investigate particle heterogeneity are highly dependent upon initial pose assignment of individual particles. Globally, all three methods agreed with our ensemble structures, in that the dominant source of structural variability within our particle population could be described by the open-closed transition of individual domains, and that sACE-N displayed more variability than sACE-C.

Initially, we performed 3D variability analysis (3DVA) in cryoSPARC to provide a direct comparison to the previously published analysis of the partial ACE cryo-EM structure (57) (Figure 6A). Trajectories were reconstructed along the top 5 principal components of variance (Movie S9). Surprisingly, the trajectory describing the variance along the first principal component vector suggests a dominant interdomain motion, where the sACE-N/N region bends towards the sACE-C/C region, similar to what we observed for trajectories along the top principal component vectors in analysis of our MD simulations (Movie S7,8), yet this variance is not reflected in our ensemble structures. The trajectories along the remaining principal component vectors comprise varying combinations of individual domains undergoing the open-close transition. Consistent with our previous observations, the range of presumptive motion in these trajectories is greater in sACE-N than sACE-C. Previous 3DVA performed on the sACE monomer structure revealed 4 types of potential motion, termed bending, pivoting, breathing, and jumping (33). The breathing motion matches the open-close transition that dominates most of our principal component vectors, while the bending motion appears to correlate with the variance described by our top principal component vector. We did not observe any potential motions reminiscent of the jumping or pivoting motions in our 3DVA. Furthermore, the magnitude of the bending motion in our 3DVA is substantially less than previously reported for the partial dimeric sACE structure, which likely suffered from substantial denaturation (33). Together, this highlights a significant difference in the

390 conformational dynamics between our sACE dimer and previously reported cryo-EM
391 structures. While there is disagreement between the 3DVA results for the two datasets, it
392 should be noted that our MD simulations display a greater range of motion than either
393 cryo-EM dataset offers, and the varieties of motion described previously are clearly
394 observed within the PCA trajectories, especially when sACE is glycosylated (Movie S8).
395 As mentioned above, this type of cryo-EM analysis is sensitive to artifacts introduced by
396 particle damage and denaturation, underscoring the importance of preserving protein
397 integrity via sample preparation and the use of Chameleon to reduce the exposure to air-
398 water interface. This also suggests dimerization functions as a key constraint to the
399 conformational dynamics of sACE.

400 We then moved on to analyze our data using cryoDRGN, which revealed that the
401 data formed a single cluster across multiple dimensions (Figure 6B, S14). Trajectories
402 generated along five principal component vectors of structural variance were dominated
403 by the open-close transition, primarily in sACE-N (Movie S10). We observed little variance
404 in sACE-C and, unlike 3DVA, cryoDRGN did not reveal any substantial interdomain
405 motions. Faced with this discrepancy between the 3DVA and cryoDRGN results, we
406 employed a third method, RECOVAR, to further analyze our data. Similar to cryoDRGN,
407 RECOVAR revealed that our data formed a singular cluster across multiple dimensions
408 of analysis. Volumes representative of kmeans cluster reconstructions were similarly
409 consistent with our cryoDRGN results. RECOVAR indicated that the primary source of
410 structural variance could be explained by the open-close transition, primarily of the sACE-
411 N domains. With this in mind, we repeated the RECOVAR analysis with a focus mask
412 around the sACE-N/N region. We then estimated the conformational density by
413 deconvoluting the multidimensional latent space with a two-dimensional principal
414 component analysis. The resulting density plot revealed four clusters which corresponded
415 to different conformations of the sACE-N/N region (Figure 6C). The most populous cluster
416 corresponded to a state where both sACE-N subunits adopted an open conformation.
417 The second largest cluster, and the smallest cluster, corresponded to a state where one
418 sACE-N was open and the other was closed. The discrepancy in population size between
419 these clusters is likely due to bias in the initial particle orientation, rather than a subunit-
420 specific preference for the open state. As the O/C state and the C/O state are 180 degree
421 rotations of each other, particle assignment to either cluster is likely influenced by the

422 initial particle orientation of the complete dimer, and we currently lack the data to discern
423 any allosteric implication to the orientation assignment. The final cluster corresponded to
424 a state where both sACE-N subunits adopted the closed conformation. We then used
425 RECOVAR to generate a trajectory through the latent space to predict a potential
426 transition pathway from the closed-closed state to the open-open state (Movie S11). This
427 trajectory revealed that one subunit would open, followed by the second, and raised the
428 question of potential allostery or coordinated motion between the adjacent sACE-N
429 domains.

430 To investigate coordination between the sACE-N domains, we performed 3D
431 classification in cryoSPARC with a focus mask on a single sACE-N domain and requested
432 2 classes (Figure 6D). The particles were split roughly evenly between the two classes,
433 one resulting in an open state and the other being closed. We then used these classes
434 for another round of 3D classification with a focus mask on the opposite sACE-N domain
435 and again asked for 2 classes. Again, these particles were split roughly evenly among the
436 classes, leaving us with four similarly sized classes adopting the same conformations as
437 the clusters revealed by RECOVAR.

438 **Discussion**

439 Our cryo-EM and MD simulations analysis of the extracellular regions of dimeric
440 ACE (sACE) offers insights into the stability of sACE catalytic domains. We show that the
441 sACE dimer interface consists of three major components, protein-protein interfaces
442 between sACE-N, glycan-glycan interaction between sACE-N, and protein-protein
443 interface between sACE-C. The interface between sACE-N, in general, is 50% larger and
444 has higher numbers of hydrogen bonds and salt bridges than that between sACE-C. The
445 necessity for the stronger interaction between sACE-N is likely due to the fact that sACE-
446 N is distal to the transmembrane helix so that the additional reinforcement is required.
447 The weaker interaction and lack of contribution from glycans in the interface between
448 sACE-C could also contribute to the observed preferential denaturation of sACE-C
449 despite the high structural similarity between sACE-N and sACE-C. The denaturation is
450 likely due to the exposure to air-water interface during vitrification, a common issue in
451 single particle cryo-EM analysis (40, 41). We have partially overcome the denaturation
452 issue of sACE-C by optimizing the buffer pH to increase the melting temperature and
453 using the faster plunging time during vitrification afforded by Chameleon to reduce the

454 exposure to air-water interface (Figure 1). However, we did not completely prevent the
455 preferential denature of sACE-C during vitrification.

456 ACE is part of a family of proteases capable of effectively degrading Aβ. This family
457 includes M16 metalloproteases like IDE and PreP, and M13 metalloproteases such as
458 neprilysin and endothelin converting enzyme. Integrative structural analysis has
459 demonstrated how IDE employs conformational dynamics, charge, and secondary
460 structure complementarity to selectively capture peptide substrates into its sizable
461 catalytic chamber, subsequently unfolding and degrading amyloidogenic peptides such
462 as insulin and Aβ, without degrading structurally similar yet non-amyloidogenic peptides
463 (30, 31, 58). The other mentioned enzymes also belong to this group of chamber-
464 containing proteases, known as cryptidases (29, 30, 59). Similar to cryptidases, structural
465 analysis reveals that ACE must undergo an open-closed transition to capture, engulf, and
466 degrade its oligopeptide substrates, such as angiotensin 1 and bradykinin.

467 Our apo-sACE structures and MD simulations identify a fulcrum mechanism that
468 governs the open-closed transition of sACE-N, while the interplay between sACE-N and
469 the D3a domain likely plays a crucial role in the corresponding transition in sACE-C.
470 Analogous to how the binding of MB60, a PreP inhibitor, to the hydrophobic pocket
471 triggers the open to closed transition, factors such as binding of substrates and ACE
472 inhibitors may alter the balance, leading to the regulation of ACE activity (29). We also
473 observe that the catalytic pocket of each sACE catalytic domain is not large enough to
474 engulf the entire Aβ. Additionally, the distances between the domains (sACE-N to sACE-
475 C, sACE-N to sACE-N, or sACE-C to sACE-C) are too large for these domains to bind Aβ
476 simultaneously (60). Therefore, ACE likely employs a different mechanism from IDE and
477 PreP to selectively bind and degrade Aβ. Future integrative structural studies await to
478 reveal how ACE selectively recognizes and degrades a diverse array of substrates.

479 Globally, we observed that sACE-N is both more open and more dynamic than
480 sACE-C in the absence of substrate. We speculate that the increased dynamicism of
481 sACE-N may be due to a weakening of the hydrophobic latch region relative to sACE-C,
482 which renders the closed state more thermodynamically favorable in sACE-C, along with
483 increased stabilization of the sACE-N open state through the fulcrum mechanism
484 described above. These factors result in sACE-N demonstrating a preference for the open
485 state, while sACE-C appears to prefer the closed state, in the absence of substrate. The

486 preference of sACE-C towards the closed state may explain why it displays a greater
487 catalytic activity compared to sACE-N for substrates such as angiotensin-I (32). The
488 difference in dynamics between sACE-N and sACE-C also offers new pathways forward
489 in the development of domain-specific inhibitors to better modulate ACE function. The
490 fulcrum-associated region centered on the K73-D189 interaction stands out as a
491 promising new druggable target to modulate the openness of sACE-N independent of
492 sACE-C. Additionally, given the range in size of ACE substrates, i.e. amyloid β (42
493 residues) vs bradykinin/angiotensin-I (9/10 residues), it is likely that the openness of
494 individual ACE domains functions as a key mediating step in substrate recognition and
495 altering the open-to-close dynamics of ACE domains may prove an effective strategy to
496 modulate ACE activity towards specific subsets of substrates. These insights come with
497 the caveat that they are derived from the apo-state dynamics of sACE and it remains
498 unknown how sACE dynamics are influenced by the presence of substrate. Studies in the
499 presence of substrate will be especially vital in exploring potential mechanisms of ACE
500 allostery.

501 The interplay between ACE-N and ACE-C within the ACE dimer has complicated
502 consequences for ACE function, showing negative cooperativity for enzymatic activities
503 while exhibiting positive synergy in Aβ cleavage by ACE-N and activity at the bradykinin
504 receptor (13, 36-39). Understanding the conformational dynamics in both inter- and
505 intradomain motions of the extracellular regions of the ACE dimer holds promise for
506 designing ACE modulators to improve and broaden the therapeutic use of ACE
507 modulation. Our cryo-EM heterogeneity analysis of sACE alone has revealed coordinated
508 yet complicated patten of open-closed transitions across four catalytic domains using
509 cryoDRGN and cryoSPARC 3DVA (Figure 6A,B, Movies S9,10). RECOVAR analysis
510 further illustrates the distribution of open-closed states within the sACE-N dimer and the
511 potential pathways for such transitions (Figure 6C). Additionally, focused classification
512 using cryoSPARC suggests no bias in the conformational state of sACE-N within a dimer,
513 irrespective of the state of its counterpart (Figure 6D). Altogether, these findings establish
514 a framework for understanding how the complex open-closed transitions within the four
515 catalytic domains of the sACE dimer contribute to its allosteric behavior (13, 36-39).

516 3DVA, cryoDRGN, and RECOVAR all appeared to display varying degrees of
517 coordinated motion between sACE domains. However, this type of analysis lacks an

518 explicit time component, so it is not possible to draw concrete conclusions regarding
519 coordinated motion. Likewise, while MD simulations contain an explicit time factor, most
520 of our MD simulations closed quickly, and did not reopen. While we observed an apparent
521 association between intradomain and interdomain motion in our PCA (movies S7, S8),
522 the analysis removes time dependency. Furthermore, the aforementioned allostery of
523 ACE was observed under the condition where substrate is present (13, 36, 38).
524 Complicating matters further, this conformational heterogeneity appears to stem partially
525 from subtle yet significant conformational differences between the two chains within an
526 ACE dimer and the inability of cryoSPARC to accurately align these conformationally
527 distinct chains, likely due to particle pseudo-symmetry (Figure S15). Furthermore,
528 CryoDRGN and RECOVAR revealed smaller magnitudes of potential interdomain
529 motions than cryoSPARC 3DVA and they were often accompanied by various
530 combinations of individual domains opening and/or closing. This suggests that the
531 apparent discrepancy in the magnitude of inferred interdomain motion could be a result
532 of the different theoretical approaches of dimensionality reduction and trajectory
533 generation employed by the respective methods of analysis and underscores the
534 importance of incorporating multiple methods of analysis in the absence of a singular
535 "gold-standard" of heterogeneity analysis. Conversely, it is possible that interpretation of
536 any potential interdomain motion is complicated by the preferential denaturation of sACE-
537 C. Principal component analysis of our MD simulations revealed the dominant source of
538 motion to be alterations in the position of sACE-C/C relative to sACE-N/N. We speculate
539 that many of these relative orientations cannot tolerate the rigors of the cryo-EM grid-
540 making process and may be the source of the observed preferential denaturation of
541 sACE-C. Future studies will need to focus on cryo-EM analysis of ACE in the presence of
542 substrates, ideally using data in which partial denaturation is substantially minimized.

543 Single particle cryo-EM holds promise for revealing the conformational dynamics
544 of biological macromolecules from embedded heterogeneity. Consequently, cryo-EM
545 heterogeneity analysis is an active and rapidly evolving field providing insights into the
546 conformational dynamics and allostery of macromolecules (53). However, many
547 challenges exist in using cryo-EM heterogeneity analysis due, in part, to the low signal-
548 to-noise ratio in observed micrographs. Often, these analyses address questions difficult
549 or impossible to tackle experimentally, leaving no ground truth for validation. We

550 encounter such challenges in our heterogeneity analysis as well. Consistent with our
551 single particle cryo-EM structural analysis, we observed the expected intradomain open-
552 closed transition using cryoDRGN, cryoSPARC 3DVA, RECOVAR, and all-atom MD
553 simulation. Surprisingly, we did not observe such motion using cryoSPARC 3DFlex, a
554 neural network-based method analyzing our cryo-EM data of sACE (54). Central to the
555 working of cryoSPARC 3DFlex is the generation of a tetrahedral mesh used to calculate
556 deformations within the particle population. Proper generation of the mesh is critical for
557 obtaining useful results and must often be determined empirically. Despite several
558 attempts, we were unable to obtain results from 3DFlex comparable to what we observed
559 with our other methods. Even using the results from our 3DVA as prior input to 3DFlex,
560 the largest conformational change we observed was a slight wiggling at the bottom of the
561 D3a subdomain (Movie S12). The authors of 3DFlex note that 3DFlex struggles to model
562 intricate motions, and the implementation of custom tetrahedral meshes currently requires
563 a non-cyclical fusion strategy between mesh segments. Given these limitations, and the
564 complexity of sACE conformational dynamics, it appears that sACE, as a system, is not
565 well-suited to analysis via 3DFlex in its current implementation. Similarly, MDSPACE
566 holds tremendous promise as a method for investigating conformational dynamics from
567 cryo-EM data (61). MDSPACE integrates cryo-EM particle data with short MD simulations
568 to fit atomic models into each particle image through an iterative process which extracts
569 dynamic information. However, the lack of GPU-enabled processing for MDSPACE
570 requires either a dedicated a computational setup that diverges from most other cryo-EM
571 software, or access to a CPU-based supercomputer, which severely limits the
572 accessibility of such software. Despite these challenges, both 3DFlex and MDSPACE use
573 promising approaches to study protein conformational dynamics. We look forward to
574 exploring effective methods to incorporate these strategies into our future research.

575 Advances in single-particle cryo-EM analysis have enabled detailed structural
576 examination of the sACE dimer. Beyond its implications for future drug discovery efforts
577 aimed at enhancing ACE-based therapies, sACE can serve as a valuable model of a
578 multi-domain glycoprotein for studying methods to preserve sample integrity, particularly
579 due to the high propensity of sACE-C to denature during vitrification. Furthermore,
580 considering the aforementioned challenges and caveats, we propose that, due to its size,

581 biomedical importance, and the complexity of its motions, sACE also represents an
582 excellent model system for evaluating and improving cryo-EM heterogeneity analysis.

583 **Experimental procedures**

584 **Expression of human sACE**: Human 293T and 293F cells were maintained in DMEM +
585 10% fetal calf serum (FCS). The lentiviral plasmid for the expression of the extracellular
586 domain of human ACE (sACE, aa 1-1235) that contains a 6-histidine tag at its C-terminus
587 was made by Vectorbuilder and co-transfected with helper plasmids to human 293T cells
588 for lentiviral production. Human 293F cells were infected with the resulting virus and
589 selected for stable sACE expression. 293F that stably expressed sACE grew poorly in
590 Free-style 293 media (ThermoFisher 12338018). Thus, for ACE production, 293F cells
591 that stably expressed sACE were first grown in DMEM+10% FCS at 37°C and gradually
592 shifted into Free-style 293 media by reducing FCS by first in Free-style 293 media + 2%
593 FCS and then to media without FCS. These cells were then shifted to and maintained in
594 30°C for 3-5 days for ACE production to reduce proteolysis.

595 **Purification of soluble human ACE:** Culture supernatant was centrifuged at 3,000xg for
596 5 min, diluted 10-fold with 25 mM Tris-HCl pH 8, filtered, and run over a Source Q column
597 pre-equilibrated with 25 mM Tris-HCl pH 8. Column was washed to baseline with Tris-HCl
598 pH 8 and bound protein was eluted with a linear gradient of NaCl to 1 M. sACE containing
599 fractions were pooled and EDTA was added to 10 mM to strip the catalytic zinc ion.
600 Sample was concentrated and run over a Superdex S200 SEC column pre-equilibrated
601 with the desired cryo-EM buffer.

602 **Differential scanning fluorometry (DSF)**: To optimize the grid making conditions,
603 differential scanning fluorimetry was applied to screen 96 buffers conditions. DSF was
604 performed with 3 mg/ml sACE and 10X Sypro Orange in 20 µl buffers in a Thermo Fisher
605 Step ONE RT-PCR thermocycler. Using the melting temperature (Tm) and slope as
606 selective criteria, 25 mM citrate pH 5.5, 25 mM sodium phosphate pH 5.5, 150 mM NaCl,
607 10 mM EDTA was identified as the optimal condition for the highest Tm and optimal slope,
608 which was used for all subsequent cryo-EM experiments.

609 **Cryo-EM data collection and processing:** Peak fractions were collected off the SEC
610 column and used to immediately prepare grids, resulting in a protein concentration of ~4.5
611 mg/ml. Quantifoil R1.2/1.3 Cu200 grids were glow-discharged in air and grids were

612 prepared by Vitrobot mark 3. 3.5 μl of sample was applied to the grid in the chamber with
613 100% humidity, 25°C. Excess sample was blotted for 1 sec, blot force 2, and grids were
614 plunge frozen in liquid ethane. All images were acquired using a Titan Krios microscope
615 (FEI) operated at 300KeV with a Gatan K3 direct electron detector. A dataset of ~3,600
616 micrographs was collected at The University of Chicago Advanced Electron Microscopy
617 facility. The dataset was processed using cryoSPARC (Figure S2) (62). Briefly, a four-
618 domain sACE dimer model was generated using sACE-N dimer (PDB 7Q4D) and sACE
619 monomer (PDB 7Q3Y) and adjusting the distance between ACE-C domains to avoid
620 steric clashes as a template for particle picking (33). After template picking, 1.1 million
621 particles were extracted with a box size of 360 pixels. Several rounds of 2D classification
622 were used to remove junk particles. ~577K particles were processed in successive rounds
623 of *ab initio* reconstruction and heterogeneous refinement. sACE grids were also prepared
624 using Chameleon and collected a dataset of ~19,000 micrographs at the National Center
625 for Cryo-EM Access and Training. For those, Quantifoil Active R 1.6/0.9 Cu/Rh 300 were
626 coated with gold using a Safematic CCU-010 Compact Coating Unit. Briefly, grids were
627 placed nanowire side down on a glass slide, gold disks were sputtered with a 35 nm target
628 thickness at 1.5 x 10^{-2} Torr, and a process current of 30 mA. Grids were glow discharged
629 with air at 12mA for 80 secs using the internal Chameleon glow discharger. The grids
630 were plunged at 160 to 180 milliseconds. All images were acquired using a Titan Krios
631 microscope (FEI) operated at 300KeV with a Gatan K3 direct electron detector (Gatan).
632 Images were automatically acquired using Leginon (63) using collection parameters as
633 shown in Table S2. This dataset was processed in cryoSPARC (Figure S3). Briefly,
634 templates were generated from our full-length ACE structure described above using
635 vitrobot-prepared grids. Micrographs with a CTF fit resolution worse than 4Å were
636 removed from the dataset, leaving ~12,000 micrographs for template picking, which
637 yielded 10.7 million particles. The particle set was trimmed to ~2.8 million based on quality
638 metrics, particles were extracted with a box size of 256 pixels and run through 2 rounds
639 of 2D classification, reducing the particle number to ~660,000. 2D classes were
640 dominated by 4 domain views, and yielded a much greater diversity of views, indicating
641 that Chameleon preparation produced a dramatic improvement in particle quality. Initial
642 3D homogeneous refinement reached 3.6 Å, matching the previous final structure. Four
643 maps from 3D classes that contains all four catalytic domains of ACE, one from particles
644 derived from grids prepared by using vitrobot and three from those prepared by using

645 Chameleon, were used for structure building using Chimera and Coot and the structure
646 was refined using Phenix real space refinement. Structural analysis was performed using
647 PYMOL and Chimera.

648 **Heterogeneity analysis**: All heterogeneity analysis was performed on the full final
649 particle set, encompassing all of the NCCAT structures. Initially, cryoSPARC 3DVA was
650 performed on both the NCCAT- and UChicago-collected datasets and results were found
651 to be consistent, so the NCCAT-collected dataset was used for further analysis, given its
652 higher particle number and quality. Input particle poses for each method of heterogeneity
653 analysis were derived from the same C1 refinement job in cryoSPARC, which reached
654 2.88 Å. 3DVA was used to solve the top 5 modes of variance with a filter resolution of 6
655 Å. Particles were downsampled to a box size of 128 pixels in cryoDRGN to speed up
656 processing. The same downsampled particle set was used for both cryoDRGN and
657 RECOVAR. CryoDRGN was run with default parameters and the analysis was performed
658 on the top 5 principal components for direct comparison to cryoSPARC 3DVA. RECOVAR
659 pipeline.py was initially run on the full sACE structure. Reconstructions representative of
660 kmeans cluster centers revealed the dominant source of structural variability to be the
661 open-close transition in sACE-N. As a result, RECOVAR was run again with a focus mask
662 on sACE-N/N and a zdim=4. Conformational density was estimated with a 2-dimensional
663 PCA, and an index value of 3 was chosen for further analysis. The points in latent space
664 representing the top 12 stable conformational states were estimated from the top 20% of
665 particles and representative volumes were generated. The generated states populated
666 four clusters and the 12 states were simplified to 4 distinct conformations where sACE-
667 N/N adopted the open-open (OO), open-closed (OC), closed-open (CO), and closed-
668 closed (CC) states (Figure 6C). We then generated a trajectory through the latent space
669 using the OO and CC representative volumes as endpoints (Movie S11).

670 **All-atom MD simulations and analysis**: The sACE-3.65 structure was used to initiate
671 all MD simulations. For the glycosylated sACE simulations, N-linked glycans were built
672 onto all 12 known N-linked glycosylation sites (residues 9, 25, 45, 82, 117, 289, 416, 480,
673 648, 666, 685, 731) with the complex glycan BGLC-BGLC-BMAN-AMAN-BGLC-ANE5-
674 AMAN-BGLC-BGAL-ANE5-AFUC using CHARMM-GUI (64). Simulations were prepared
675 in QwikMD or CHARMM-GUI and run at pH 7, 150 mM NaCl, 310 K, 1 atm with periodic
676 boundary conditions in NAMD3.0 using the University of Chicago Beagle3 partition (65).

Explicit solvent was described with the TIP3P model (66). Data was analyzed in VMD 1.9.4 (67) using a combination of built-in plugins and custom TCL scripts, as well as custom Python scripts. Analysis scripts are available upon reasonable request.

680 **Figure + figure legends:**

681 Figure 1: Cryo-EM analysis of human sACE dimer. (A) Diagram for the key features of
682 domain on primary sequence of human ACE. The construct used in this study contains
683 the first 1235 residues, which we refer to as the soluble region of ACE (sACE) however,
684 the first 29 residues comprise a signal. By convention, sACE is labeled based on the
685 mature, processed peptide (68). D1a and D3a domains, also called "lid" encompasses
686 residues 1-98 and aa 616-696, respectively. D1b and D3b domains each have two
687 discrete segments; residues 263-436 and 496-574 for D1b and residues 868-1031 and
688 1091-1171 for D3b. D2 encompasses three discrete segments, residues 99-262, 437-
689 495, and 575-615 while D4 domain encompasses residues 697-867, 1032-1090, and
690 1172-1202. sp = signal peptide. asterisk = zinc binding motif. Colored by sub-domain:
691 D1a blue, D1b gold, D2 magenta, D3a cyan, D3b green, D4 purple. (B) 2D classification
692 of human sACE particles from grids made by vitrobot (360 pixels box size) and
693 Chameleon (256 pixels box size). Clear four domain classes visible in the Chameleon-
694 derived classification are boxed in red, similar views are lacking in the vitrobot dataset.
695 White scale bar measures 100 Å. (C) Full-length sACE 3D volumes, colored by sub-
696 domain as in (A). Chain B sub-domains are depicted as lighter tones of their chain A
697 counterparts. Glycan density is shown in gray. See Figure S2 for vitrobot-prepared data
698 processing details, Figure S4 for Chameleon-prepared data processing details and Table
699 S2 for data refinement statistics.

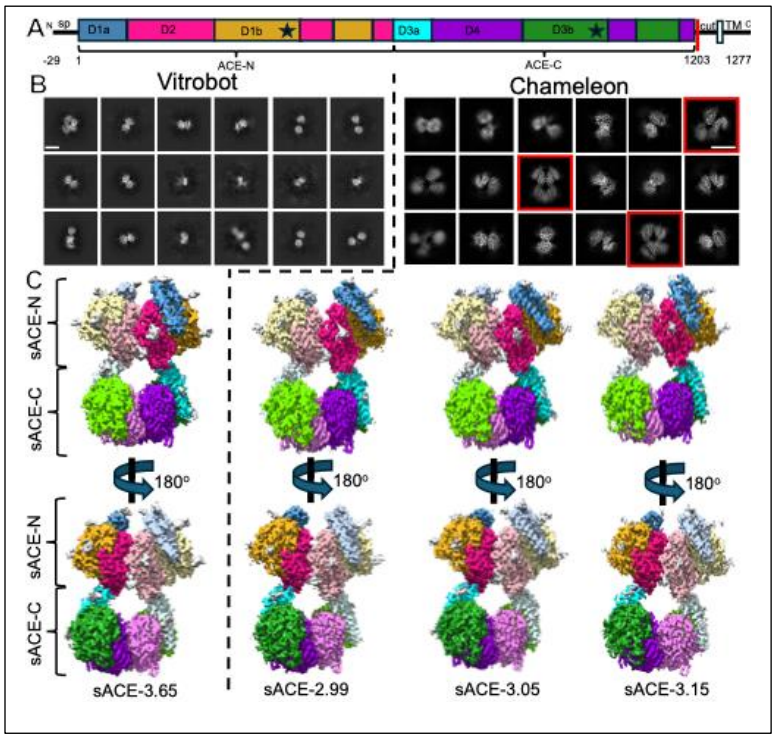
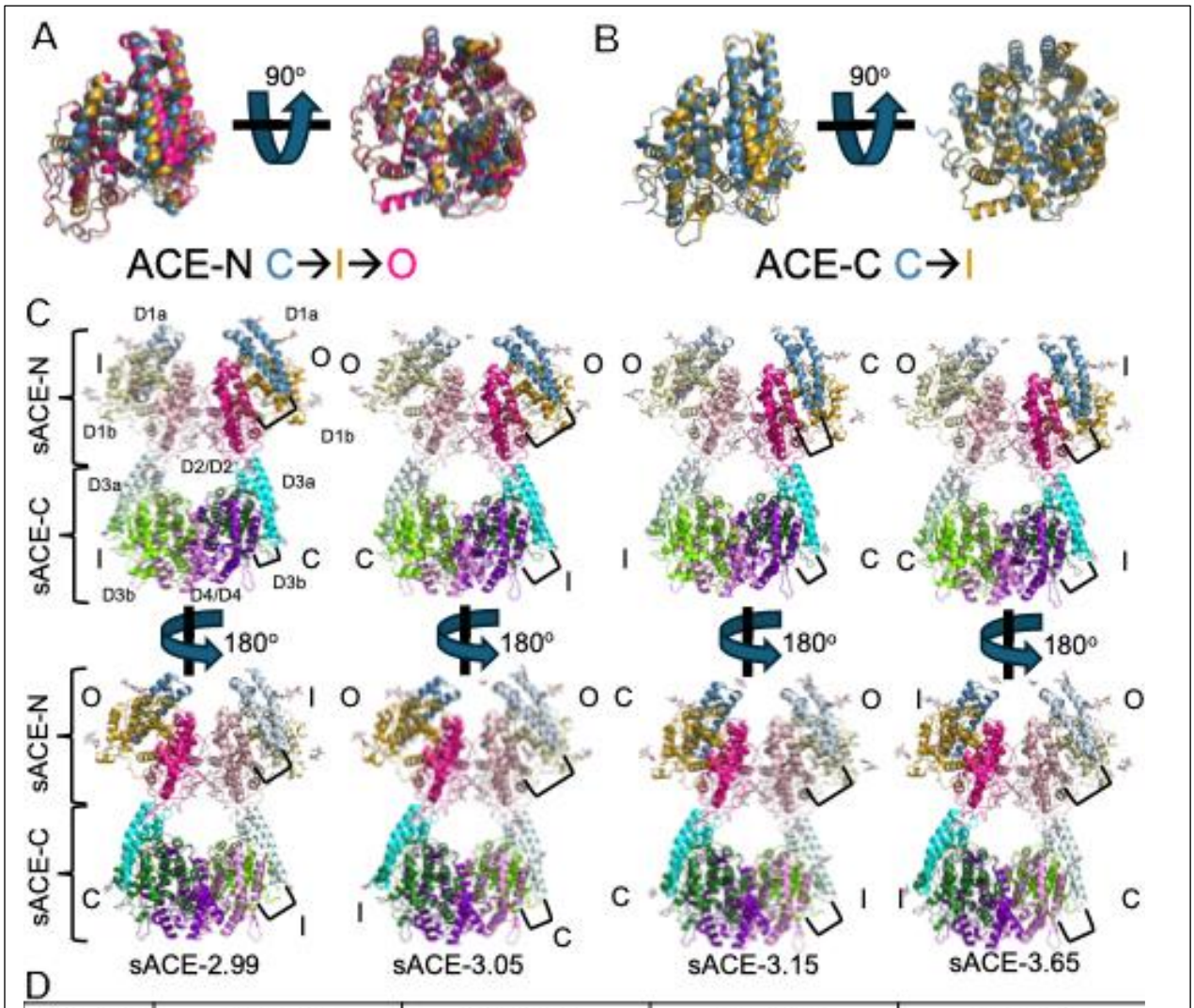


Figure 2: Overall structure of human sACE. (A) Overlay comparing sACE-N states highlighting the structure differences between the closed “C”, intermediate “I”, and open “O” states. (B) Overlay comparing sACE-C states highlighting the structure differences between the closed “C”, and intermediate “I” states. We define the state based on the distance between the edge of the D2/4 domain bordering the catalytic cleft (residues 121-126 or 721-726) and the tip of the D1/3a region (residues 41-51 or 647-657): closed <15 Å, intermediate >15 Å and <19 Å, open >19 Å. (C) Overall dimer comparisons. Black bracket depicts the D1/3a-D2/4 distance measurement used to define the state of each domain, as described above. (D) Table of openness measurements for each domain per structure.



D

Opening distance (Å)	sACE-2.99		sACE-3.05		sACE-3.15		sACE-3.65	
	Chain A	Chain B	Chain A	Chain B	Chain A	Chain B	Chain A	Chain B
ACE-N	23.4 (O)	17.7 (I)	22.5 (O)	22.1 (O)	14.1 (C)	20.2 (O)	16.2 (I)	20.6 (O)
ACE-C	14.4 (C)	18.0 (I)	17.9 (I)	13.9 (C)	13.7 (C)	17.0 (I)	15.9(I)	14.4 (C)

Figure 3: sACE dimerization interfaces. (A) Overlay comparing sACE-N/N interface (blue) and sACE-C/C (gold) interfaces. Interfaces adopt the same secondary structure but interacting residues vary between them. (B) Residue-specific interactions at the sACE-N/N interface, see text for details. (C) Unsharpened Coulomb potential density map (blue) showing density corresponding to glycan-glycan interaction from N82 as part of sACE-N/N interface. Sharpened map is shown in magenta for reference. (D) Residue-specific interactions in the sACE-C/C interface, see text for details.

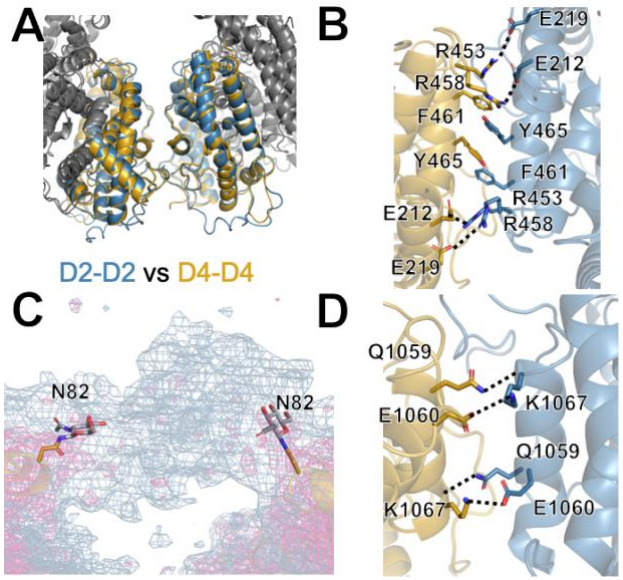


Figure 4: Comparison of molecular dynamics simulations with cryo-EM data. (A) Alpha carbon displacement values for each residue were calculated by comparing each of our cryo-EM structures against one another in a pairwise manner and averaged. Individual residues are colored by sub-domain as defined in figure 1. Alpha carbon displacement values were also mapped onto the structure of sACE to better visualize the mobile regions. Residues are colored by degree of displace from blue (no displacement) to red (high displacement, values in Angstrom, as indicated). Numbers denote regions of interest: 1, the top of the D1a sub-domain, 2, the bottom of the D1a sub-domain, 3, a flexible loop region within the D1b that interacts with the bottom of the D1a sub-domain near the catalytic cleft, 4, the bottom of the D3A sub-domain. (B) Alpha carbon RMSF values for a representative 100 ns subset from our non-glycosylated MD simulations that demonstrates the simulated dynamics correlate well with displacement calculated from the cryo-EM structures. (C) Alpha carbon RMSF values for the entirety of our non-glycosylated MD simulations. Mobile regions agree with predictions from the cryo-EM structures, yet the magnitude is dampened due to the protein rapidly transitioning from an open state to a closed state and remaining closed for most of the simulation time. (D) Alpha carbon RMSF values for a representative 100 ns subset from our glycosylated MD

simulations. (E) Alpha carbon RMSF values for the entirety of our glycosylated MD simulations.

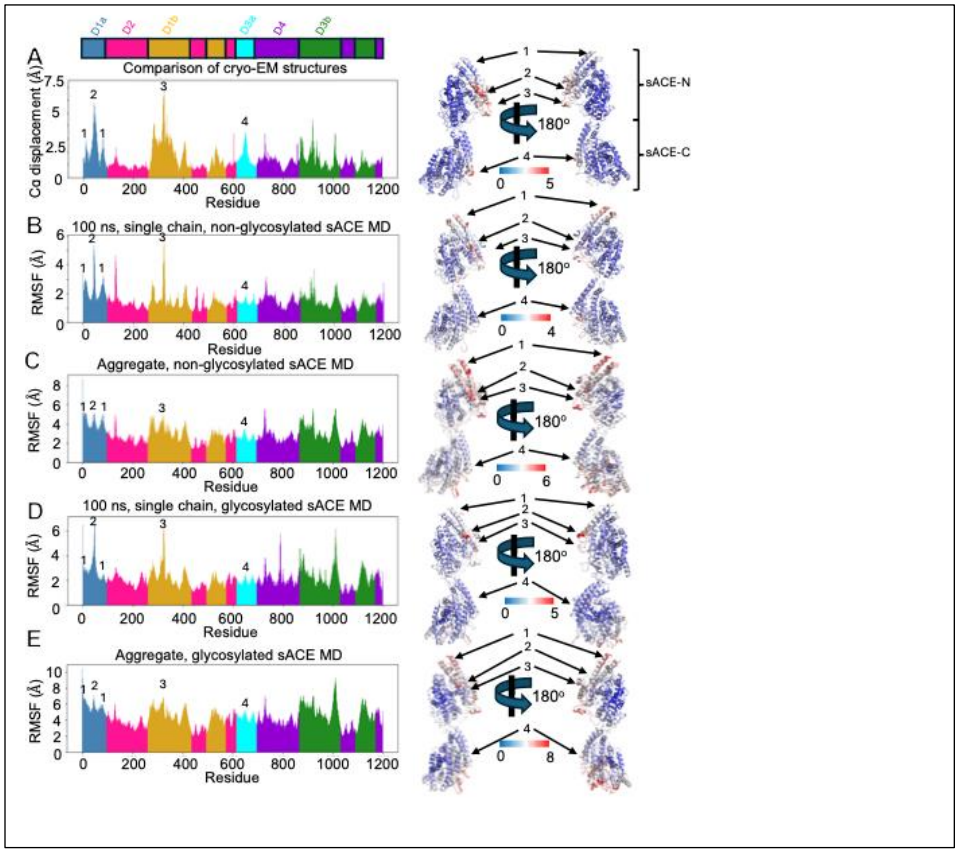


Figure 5: Structural mechanism of the sACE open/close transition. (A) sACE-N overlay comparing the open (left panel) and closed (right panel) states in detail. The open state is stabilized by interaction between residues in the D1a (blue) and D2 (magenta) regions that, notably K73-D189. In the closed state, the K73-D189 interaction is broken. (B) Overlay of the sACE-N open (dark shades, sub-domains colored as above) and closed (light shades, sub-domains colored as above) states showing the range of motion. The D1a region rotates about a fulcrum region described in (A), while the D1b region moves as a rigid body. Front view arrows depict the “fulcrum” motion of the D1a subdomain, with the top and bottom of the subdomain moving in opposite directions. Side view arrows depict rigid body motion of the D1a and D1b subdomains moving together. (C) Overlay of sACE-C closed (light shades, sub-domains colored as above) and intermediate (dark shades, sub-domains colored as above) states. Unlike sACE-N, the “top” of the D3a region is constrained by its connection to sACE-N and largely immobile. The primary source of opening is only the motion of the D3a tip. We did not observe any open state

754 structures of sACE-C, suggesting a smaller range of motion relative to sACE-N. Front
755 view arrow depicts motion of D3a subdomain. Only the bottom of the subdomain moves,
756 in contrast to the “fulcrum” motion observed in the D1a subdomain. Side view depicts the
757 rigid body motion of the D3a and D3b subdomains moving together. (D) Comparison of
758 the hydrophobic “latch” region formed in the closed state between residues of the D1/3a,
759 D1/3b, and D2/D4 domains. V724 in sACE-C has been replaced by T124 in sACE-N,
760 suggesting that the closed state in sACE-N may be less stabilized than the sACE-C
761 closed state. (E) Example all-atom MD simulation tracking the openness of one sACE-N
762 region (black line) and this distance between K73 and D189 (red line). These residues
763 form a salt bridge early in the simulation when sACE-N is open (left inset) but the
764 interaction breaks as sACE-N transitions to the closed state (right inset). Distance
765 measurements for MD simulations were consistently greater than distance values in our
766 static structures and cannot be directly compared to Figure 2.

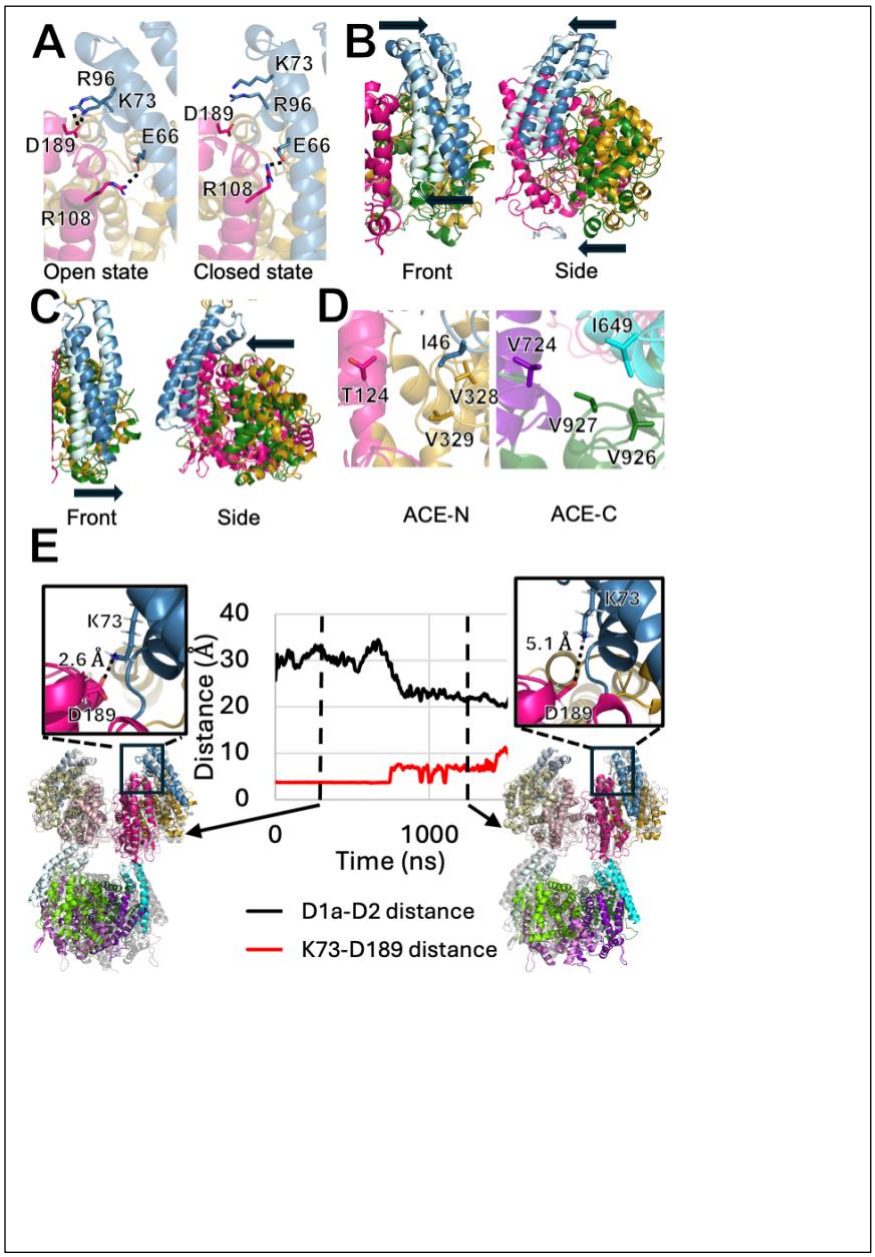
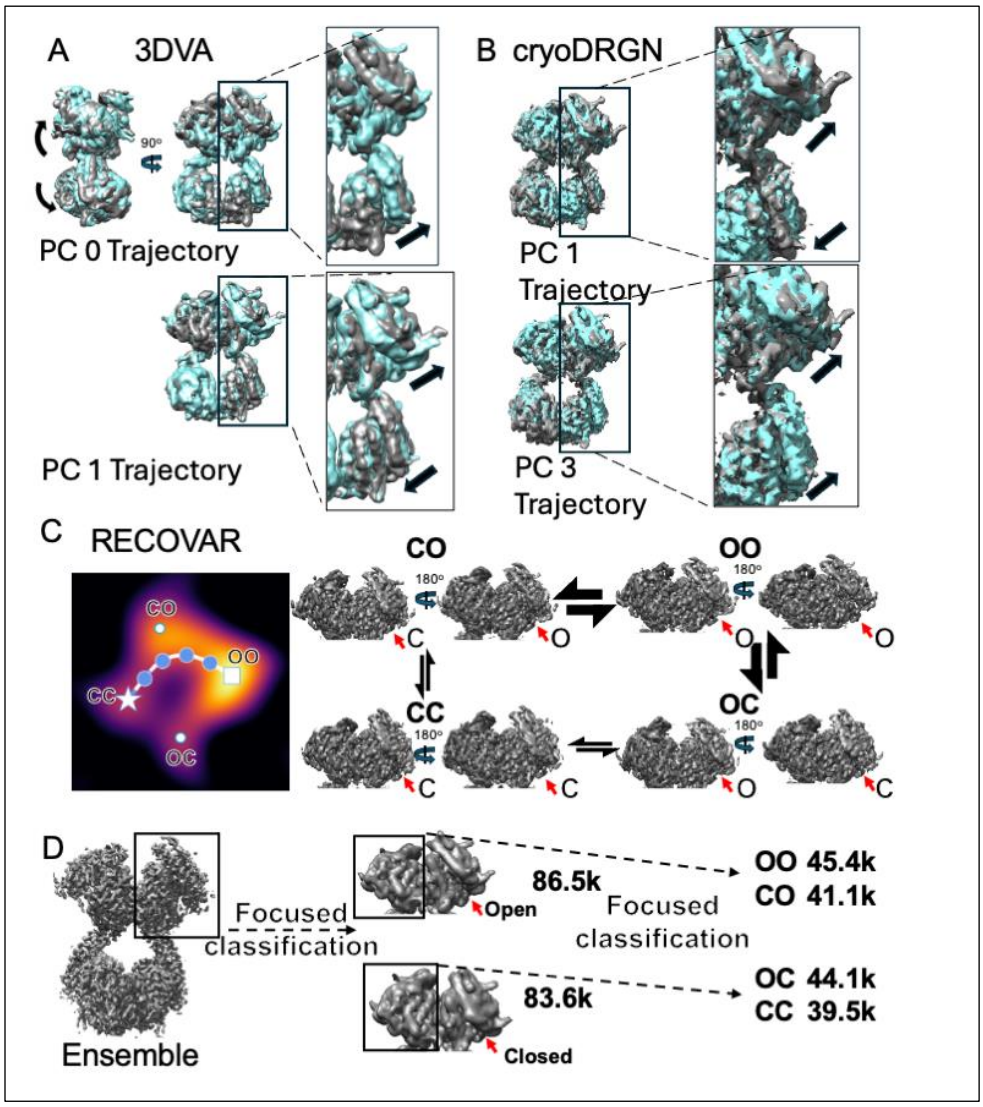


Figure 6: Cryo-EM heterogeneity analysis. (A) Visualization of the structural changes revealed by cryoSPARC 3DVA trajectories calculated along two principal components (PCs) of structural variance. Starting states are showing in cyan, ending states in gray. PC 0 reveals a large, inter-domain bending motion accompanied by the open/close transition in sACE-C. PC 1 and the remaining PCs are dominated by the open/close transition of individual regions. See Table S3, and Movie S9 for additional details. Arrows depict generalized motions. (B) Visualization of the structural changes revealed by the cryoDRGN trajectories calculated along two PCs of structural variance. Starting states are shown in cyan, ending states in gray. PCs are dominated by the open/close transition of individual regions. See Table S3, and Movie S10 for additional details. Arrows depict

generalized motions. (C) Analysis in RECOVAR with a focus mask on sACE-N/N reveals that particles adopt roughly four clusters within the latent space (heat map of particle density) corresponding to the open-open (OO, white square), open-closed (OC, white dot), closed-open (CO, white dot) and closed-closed (CC, white star) states of sACE-N/N. A trajectory estimating the path through latent space corresponding to the structural transition from sACE-N/N CC state to the OO state (blue points) suggests that individual sACE regions transition at different rates, as indicated by the size of the transition arrows between states. See Movie S11 for trajectory. (D) Focused 3D classification was performed in cryoSPARC to explore evidence of coordinated motion between sACE-N regions. 3D classification focusing on one sACE-N region revealed 2 roughly equal classes of particles: open and closed. Subsequent 3D classification focused on the other sACE-N region again revealed 2 roughly equal classes, suggesting the lack of coordinated motion between sACE-N regions in the absence of substrate.



Data availability:

Unprocessed micrographs, particle stacks, and associated pose/CTF information have been deposited to EMPIAR under the accession numbers: EMPIAR-12181 (Vitrobot-prepared grids) and EMPIAR-12484 (Chameleon-prepared grids)

EM structures and maps have been deposited to the EMDB and PDB, respectively, under the following accession numbers:

sACE-2.99: 9D5S and EMD-46581

sACE-3.05: 9D5M and EMD-46579

sACE-3.15: 9D55 and EMD-46574

sACE-3.65: 9CLX and EMD-45733

References:

1. K. E. Bernstein *et al.*, A Modern Understanding of the Traditional and Nontraditional Biological Functions of Angiotensin-Converting Enzyme. *Pharmacol Rev* **65**, 1-46 (2013).
2. L. Lubbe, G. E. Cozier, D. Oosthuizen, K. R. Acharya, E. D. Sturrock, ACE2 and ACE: structure-based insights into mechanism, regulation and receptor recognition by SARS-CoV. *Clin Sci* **134**, 2851-2871 (2020).
3. A. J. Turner, N. M. Hooper, The angiotensin-converting enzyme gene family: genomics and pharmacology. *Trends Pharmacol Sci* **23**, 177-183 (2002).
4. L. L. Herman, S. A. Padala, I. Ahmed, K. Bashir, "Angiotensin-Converting Enzyme Inhibitors (ACEI)" in StatPearls. (Treasure Island (FL) ineligible companies. Disclosure: Sandeep Padala declares no relevant financial relationships with ineligible companies. Disclosure: Intisar Ahmed declares no relevant financial relationships with ineligible companies. Disclosure: Khalid Bashir declares no relevant financial relationships with ineligible companies., 2024).
5. Z. C. Wu *et al.*, Association between angiotensin-converting enzyme inhibitors and the risk of lung cancer: a systematic review and meta-analysis. *Brit J Cancer* **128**, 168-176 (2023).
6. T. K. Yao *et al.*, Association between angiotensin-converting enzyme inhibitor-induced cough and the risk of lung cancer: a Mendelian randomization study. *Front Pharmacol* **14** (2023).
7. M. Semis, G. B. Gugiu, E. A. Bernstein, K. E. Bernstein, M. Kalkum, The Plethora of Angiotensin-Converting Enzyme-Processed Peptides in Mouse Plasma. *Anal Chem* **91**, 6440-6453 (2019).
8. W. Y. Zheng *et al.*, Small molecule angiotensin converting enzyme inhibitors: A medicinal chemistry perspective. *Front Pharmacol* **13** (2022).
9. K. E. Bernstein *et al.*, Angiotensin-converting enzyme in innate and adaptive immunity. *Nat Rev Nephrol* **14**, 325-336 (2018).
10. R. Danziger *et al.*, The effects of enhancing angiotensin converting enzyme in myelomonocytes on ameliorating Alzheimer's-related disease and preserving cognition. *Front Physiol* **14**, 1179315 (2023).
11. H. Hampel *et al.*, The Amyloid-β Pathway in Alzheimer's Disease. *Mol Psychiatr* **26**, 5481-5503 (2021).
12. M. L. Hemming, D. J. Selkoe, Amyloid beta-protein is degraded by cellular angiotensin-converting enzyme (ACE) and elevated by an ACE inhibitor. *J Biol Chem* **280**, 37644-37650 (2005).
13. K. Zou *et al.*, A beta 42-to-A beta 40-and Angiotensin-converting Activities in Different Domains of Angiotensin-converting Enzyme. *J Biol Chem* **284**, 31914-31920 (2009).
14. Y. Zhang, H. Chen, R. Li, K. Sterling, W. Song, Amyloid beta-based therapy for Alzheimer's disease: challenges, successes and future. *Signal Transduct Target Ther* **8**, 248 (2023).
15. K. E. Bernstein *et al.*, Angiotensin-converting enzyme overexpression in myelomonocytes prevents Alzheimer's-like cognitive decline. *J Clin Invest* **124**, 1000-1012 (2014).
16. F. Gouveia *et al.*, Targeting brain Renin-Angiotensin System for the prevention and treatment of Alzheimer's disease: Past, present and future. *Ageing Res Rev* **77** (2022).
17. J. W. Wright, L. H. Kawas, J. W. Harding, A Role for the Brain RAS in Alzheimer's and Parkinson's Diseases. *Front Endocrinol (Lausanne)* **4**, 158 (2013).
18. W. Q. Qiu *et al.*, Angiotensin Converting Enzyme Inhibitors and the Reduced Risk of Alzheimer's Disease in the Absence of Apolipoprotein E4 Allele. *J Alzheimers Dis* **37**, 421-428 (2013).
19. J. F. Riordan, Angiotensin-I-converting enzyme and its relatives. *Genome Biol* **4**, 225 (2003).
20. W. R. English, P. Corvol, G. Murphy, LPS activates ADAM9 dependent shedding of ACE from endothelial cells. *Biochem Bioph Res Co* **421**, 70-75 (2012).
21. N. M. Hooper, A. J. Turner, The search for α-secretase and its potential as a therapeutic approach to Alzheimer's disease. *Curr Med Chem* **9**, 1107-1119 (2002).

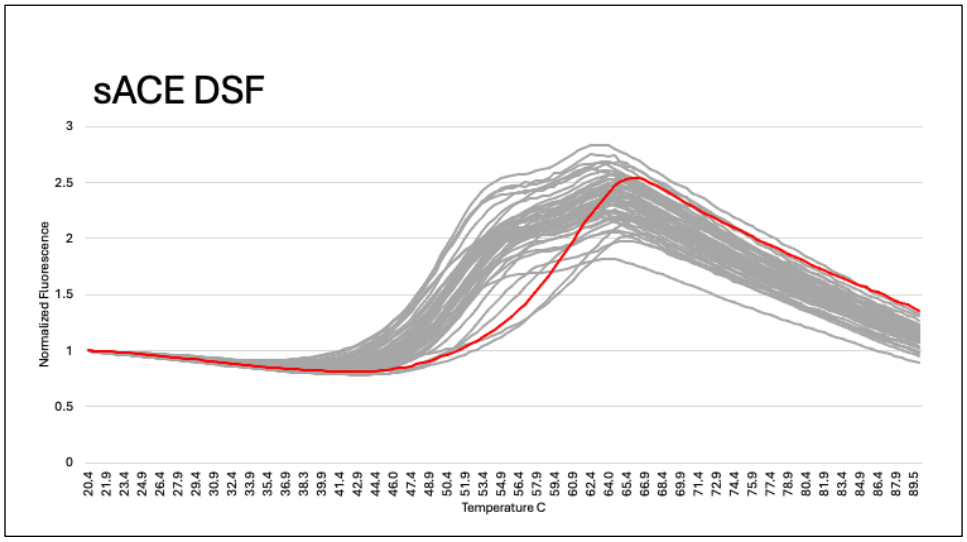
22. O. A. Kost *et al.*, Epitope-dependent blocking of the angiotensin-converting enzyme dimerization by monoclonal antibodies to the N-terminal domain of ACE: possible link of ACE dimerization and shedding from the cell surface. *Biochemistry* **42**, 6965-6976 (2003).
23. F. X. Gomis-Ruth, Structure and Mechanism of Metallocarboxypeptidases. *Critical Reviews in Biochemistry and Molecular Biology* **43**, 319-345 (2008).
24. K. R. Acharya, K. S. Gregory, E. D. Sturrock, Advances in the structural basis for angiotensin-1 converting enzyme (ACE) inhibitors. *Bioscience Rep* **44** (2024).
25. H. R. Corradi, S. L. Schwager, A. T. Nchinda, E. D. Sturrock, K. R. Acharya, 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 (2006).
26. G. Masuyer, S. L. U. Schwager, E. D. Sturrock, R. E. Isaac, K. R. Acharya, Molecular recognition and regulation of human angiotensin-I converting enzyme (ACE) activity by natural inhibitory peptides. *Sci Rep* **2**, 717 (2012).
27. R. Natesh, S. L. Schwager, E. D. Sturrock, K. R. Acharya, Crystal structure of the human angiotensin-converting enzyme-lisinopril complex. *Nature* **421**, 551-554 (2003).
28. K. R. Acharya, E. D. Sturrock, J. F. Riordan, M. R. Ehlers, Ace revisited: a new target for structure-based drug design. *Nat Rev Drug Discov* **2**, 891-902 (2003).
29. W. G. Liang *et al.*, Structural basis for the mechanisms of human presequence protease conformational switch and substrate recognition. *Nat Commun* **13**, 1833 (2022).
30. E. Malito, R. E. Hulse, W. J. Tang, Amyloid β-degrading cryptidases: insulin degrading enzyme, presequence peptidase, and neprilysin. *Cell Mol. Life Sci.* **65**, 2574-2585 (2008).
31. Z. N. Zhang *et al.*, Ensemble cryoEM elucidates the mechanism of insulin capture and degradation by human insulin degrading enzyme. *Elife* **7**, e33572 (2018).
32. G. E. Cozier, L. Lubbe, E. D. Sturrock, K. R. Acharya, Angiotensin-converting enzyme open for business: structural insights into the subdomain dynamics. *FEBS J* **288**, 2238-2256 (2021).
33. L. Lubbe, B. T. Sewell, J. D. Woodward, E. D. Sturrock, Cryo-EM reveals mechanisms of angiotensin I-converting enzyme allostery and dimerization. *EMBO J* **41**, e110550 (2022).
34. T. T. V. Phan, S. Y. Heo, W. K. Jung, M. Yi, Spontaneous Hinge-Bending Motions of Angiotensin I Converting Enzyme: Role in Activation and Inhibition. *Molecules* **25**, 1288 (2020).
35. K. E. Bernstein *et al.*, Different in vivo functions of the two catalytic domains of angiotensin-converting enzyme (ACE). *Curr Opin Pharmacol* **11**, 105-111 (2011).
36. P. V. Binevski, E. A. Sizova, V. F. Pozdnev, O. A. Kost, Evidence for the negative cooperativity of the two active sites within bovine somatic angiotensin-converting enzyme. *Febs Lett* **550**, 84-88 (2003).
37. G. I. Rice, D. A. Thomas, P. J. Grant, A. J. Turner, N. M. Hooper, Evaluation of angiotensin-converting enzyme (ACE), its homologue ACE2 and neprilysin in angiotensin peptide metabolism. *Biochem J* **383**, 45-51 (2004).
38. O. E. Skirgello, P. V. Binevski, V. F. Pozdnev, O. A. Kost, Kinetic probes for inter-domain co-operation in human somatic angiotensin-converting enzyme. *Biochem J* **391**, 641-647 (2005).
39. E. G. Erdos, F. Tan, R. A. Skidgel, Angiotensin I-converting enzyme inhibitors are allosteric enhancers of kinin B1 and B2 receptor function. *Hypertension* **55**, 214-220 (2010).
40. R. M. Glaeser, Proteins, interfaces, and cryo-EM grids. *Curr Opin Colloid In* **34**, 1-8 (2018).
41. A. J. Noble *et al.*, Reducing effects of particle adsorption to the air-water interface in cryo-EM. *Nat Methods* **15**, 793-795 (2018).
42. S. Voronov, N. Zueva, V. Orlov, A. Arutyunyan, O. Kost, Temperature-induced selective death of the C-domain within angiotensin-converting enzyme molecule. *Febs Lett* **522**, 77-82 (2002).
43. M. C. Darrow, J. P. Moore, R. J. Walker, K. Doering, R. S. King, Chameleon: Next Generation Sample Preparation for CryoEM based on Spotiton. *Microscopy and Microanalysis* **25**, 994-995 (2019).

44. H. R. Corradi, S. L. U. Schwager, A. T. Nchinda, E. D. Sturrock, K. R. Acharya, Crystal structure of the N domain of human somatic angiotensin I-converting enzyme provides a structural basis for domain-specific inhibitor design. *Journal of Molecular Biology* **357**, 964-974 (2006).
45. R. Natesh, S. L. U. Schwager, E. D. Sturrock, K. R. Acharya, Crystal structure of the human angiotensin-converting enzyme-lisinopril complex. *Nature* **421**, 551-554 (2003).
46. S. M. Danilov *et al.*, An angiotensin I-converting enzyme mutation (Y465D) causes a dramatic increase in blood ACE via accelerated ACE shedding. *PLoS One* **6**, e25952 (2011).
47. O. A. Kost, N. V. Bovin, E. E. Chemodanova, V. V. Nasonov, T. A. Orth, New feature of angiotensin-converting enzyme: carbohydrate-recognizing domain. *J Mol Recognit* **13**, 360-369 (2000).
48. H. G. O'Neill, P. Redelinghuys, S. L. U. Schwager, E. D. Sturrock, The role of glycosylation and domain interactions in the thermal stability of human angiotensin-converting enzyme. *Biol Chem* **389**, 1153-1161 (2008).
49. R. Sadhukhan, I. Sen, Different glycosylation requirements for the synthesis of enzymatically active angiotensin-converting enzyme in mammalian cells and yeast. *J Biol Chem* **271**, 6429-6434 (1996).
50. B. Baudin, N. Alves, A. Pilon, B. Beneteau-Burnat, J. Giboudeau, Structural and biological roles of glycosylations in pulmonary angiotensin I-converting enzyme. *Glycobiology* **7**, 565-570 (1997).
51. T. Orth, S. Voronov, P. Binevski, W. Saenger, O. Kost, Glycosylation of bovine pulmonary angiotensin-converting enzyme modulates its catalytic properties. *Febs Lett* **431**, 255-258 (1998).
52. X. C. Yu *et al.*, 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 (1997).
53. W. S. Tang, E. D. Zhong, S. M. Hanson, E. H. Thiede, P. Cossio, Conformational heterogeneity and probability distributions from single-particle cryo-electron microscopy. *Curr Opin Struc Biol* **81** (2023).
54. A. Punjani, D. J. Fleet, 3DFlex: determining structure and motion of flexible proteins from cryo-EM. *Nature Methods* **20**, 860-870 (2023).
55. M. A. Gilles, A. Singer, Cryo-EM Heterogeneity Analysis using Regularized Covariance Estimation and Kernel Regression. *bioRxiv* 10.1101/2023.10.28.564422 (2024).
56. E. D. Zhong, T. Bepler, B. Berger, J. H. Davis, CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. *Nature Methods* **18**, 176-+ (2021).
57. A. Punjani, D. J. Fleet, 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *J Struct Biol* **213** (2021).
58. W. J. Tang, Targeting Insulin-Degrading Enzyme to Treat Type 2 Diabetes Mellitus. *Trends Endocrinol Metab* **27**, 24-34 (2016).
59. W. G. Liang, J. M. Mancl, M. Zhao, W. J. Tang, Structural analysis of Mycobacterium tuberculosis M13 metalloprotease Zmp1 open states. *Structure* **29**, 709-720 (2021).
60. G. F. Chen *et al.*, Amyloid beta: structure, biology and structure-based therapeutic development. *Acta Pharmacol Sin* **38**, 1205-1235 (2017).
61. R. Vuillemot *et al.*, MDSPACE: Extracting Continuous Conformational Landscapes from Cryo-EM Single Particle Datasets Using 3D-to-2D Flexible Fitting based on Molecular Dynamics Simulation. *Journal of Molecular Biology* **435** (2023).
62. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296 (2017).
63. C. Suloway *et al.*, Automated molecular microscopy: The new Leginon system. *J Struct Biol* **151**, 41-60 (2005).

64. J. Lee *et al.*, CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J Chem Theory Comput* **12**, 405-413 (2016).
65. J. C. Phillips *et al.*, Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J Chem Phys* **153** (2020).
66. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of Simple Potential Functions for Simulating Liquid Water. *J Chem Phys* **79**, 926-935 (1983).
67. W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-38, 27-38 (1996).
68. F. Soubrier *et al.*, Two putative active centers in human angiotensin I-converting enzyme revealed by molecular cloning. *Proc Natl Acad Sci U S A* **85**, 9386-9390 (1988).

963 **Supplemental figures**

964 Figure S1: DSF optimization of sACE grid-making conditions. 48 buffer conditions were
965 screened for sACE stability in the presence and absence of EDTA (see table S1). Citrate-
966 phosphate buffer pH 5.5 (red) was chosen for the best combination of stability
967 (temperature shift right) and homogeneity (steepness of transition slope). Two other
968 citrate-phosphate conditions, pH 5 and pH 4.5 gave better stability, but a more gradual
969 transition slope, and so were not chosen.



970

971 Figure S2: Processing workflow for dataset using Vitrobot-prepared grids. 3,653
972 micrographs were collected on a 300 keV Titan Krios at the University of Chicago. A full-
973 length model of sACE was manually generated by combining available PDB structures of
974 ACE-N and ACE-C. The resulting PDB file was converted to an EM map in Chimera and
975 imported into cryoSPARC to generate particle picking templates. Using these templates,
976 1,171,502 particles were picked and extracted. Successive rounds of 2D classification
977 were used to trim the particle number down to 577,123 particles. *Ab initio* reconstruction
978 followed by heterogeneous refinement yielded a full sACE dimer map (214,893 particles),
979 a map representing only the N/N dimer (268,471 particles), and 93,759 particles that
980 mapped to an uninterpretable junk class. The particles mapping to the full length sACE
981 structure were then transferred to PCA-based 3D classification. 103,507 particles
982 mapped to low resolution partial structures that could not clearly be assigned structural
983 significance and were discarded as junk. Of the remaining particles, 20,533 particles
984 mapped to a three-domain sACE structure. The remaining 88,930 particles mapped to a

full-length sACE structure that was refined to 3.65 Å (sACE-3.65). Scale bar in representative micrograph shows 200 Å, scale bars in 2D classes show 100 Å.

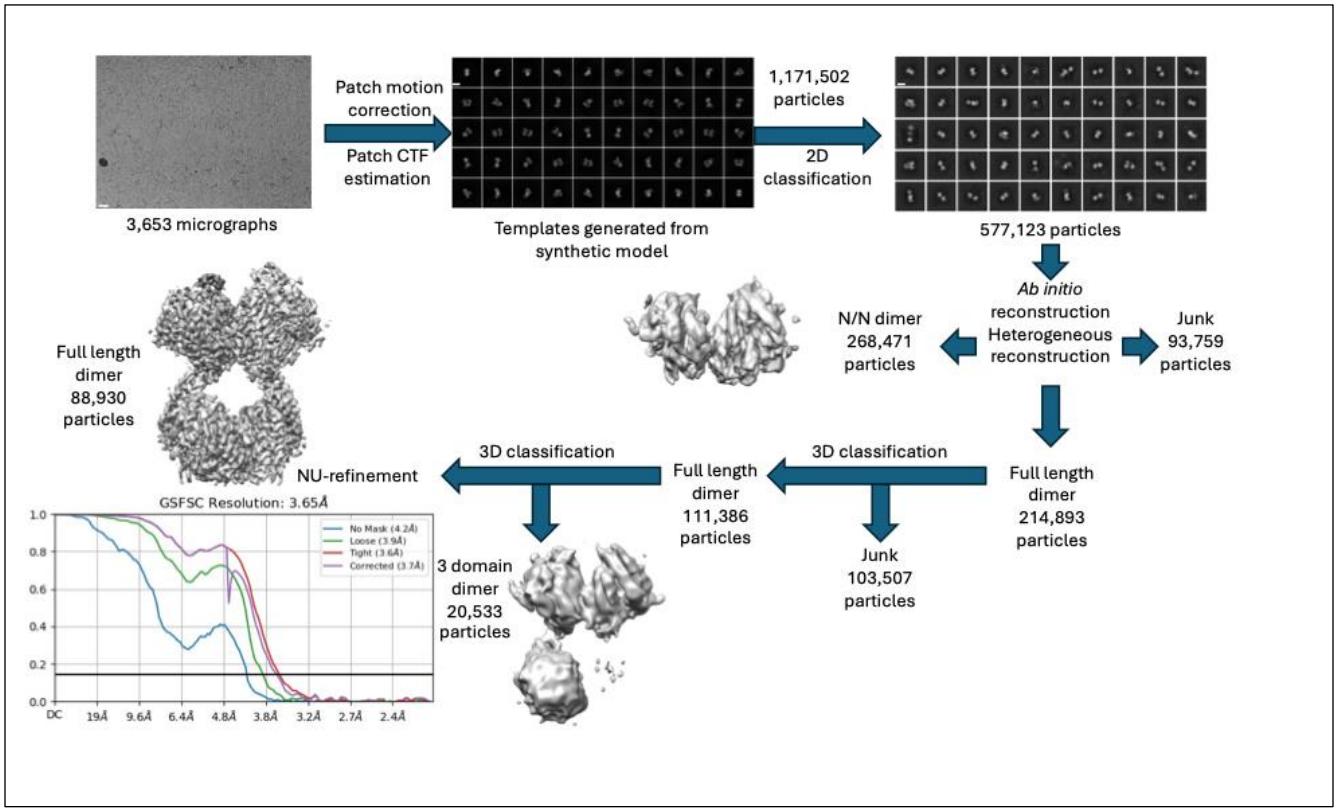


Figure S3: Assessment of map quality from Vitrobot prepared grids. (A) Angular distribution. (B) Local resolution colored as blue-cyan-green-yellow-red spectrum from 3-5 Å. (C) Example residue fits to density map. (D) 3D-FSC results.

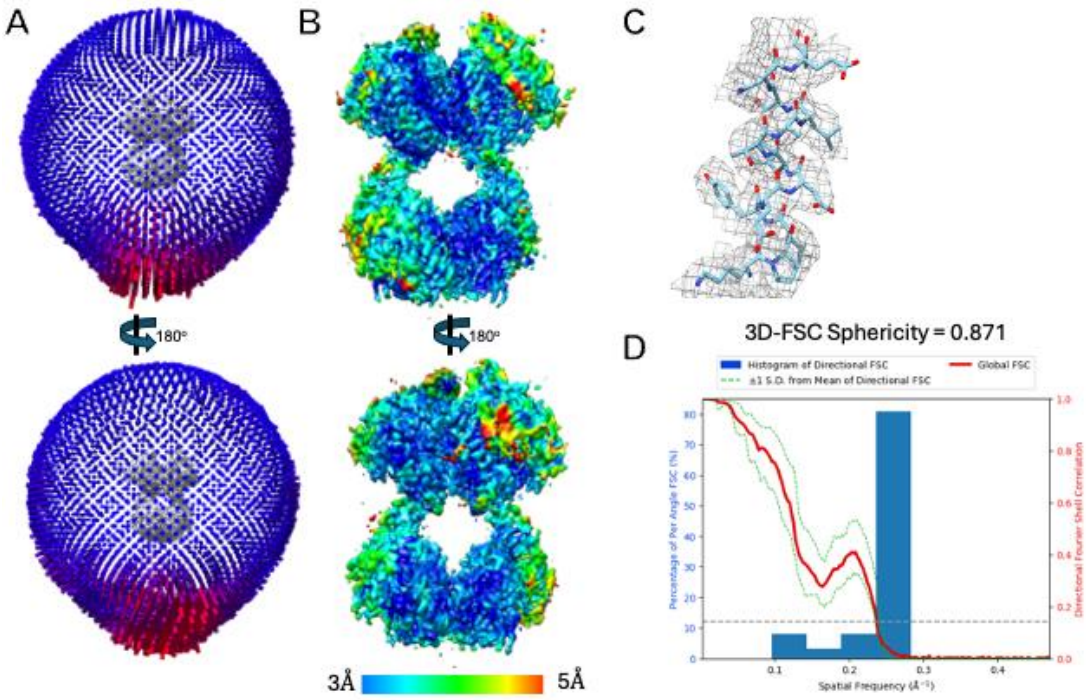


Figure S4: Processing workflow for dataset using Chameleon-prepared grids. 18,281 micrographs were collected at the National Center for CryoEM Access and Training. After motion correction and CTF estimation, the dataset was trimmed to include only micrographs with a max CTF value lower than 4 Angstroms. The sACE-3.65 structure was used to generate 2D templates for particle picking. 10,708,232 particles were originally picked. This number was trimmed to 2,580,990 based on quality metrics and visual observation. The resulting particles were used in successive rounds of 2D classification to eliminate junk particles, reducing the particle count to 662,219. *Ab initio* reconstruction followed by heterogeneous refinement revealed four classes: 165,395 particles mapped to a class we interpret as the sACE-C/C dimer, 140,038 particles mapped to a class we interpret as the sACE-N/N dimer, 50,647 particles mapped to a 3-domain sACE class, and the remaining 306,139 particles mapped to a full-length sACE class. PCA-based 3D classification was used to eliminate junk particles, and the remaining ~170k particles were refined with C2 symmetry to 2.8 Angstroms. The D1 domains in the C2 reconstruction displayed poor density, likely resulting from molecular motion, so a final round of 3D classification and refinement with C1 symmetry was performed to finalize three states of full-length sACE. Scale bar in representative micrograph shows 150 Å, scale bar in 2D classes show 100 Å.

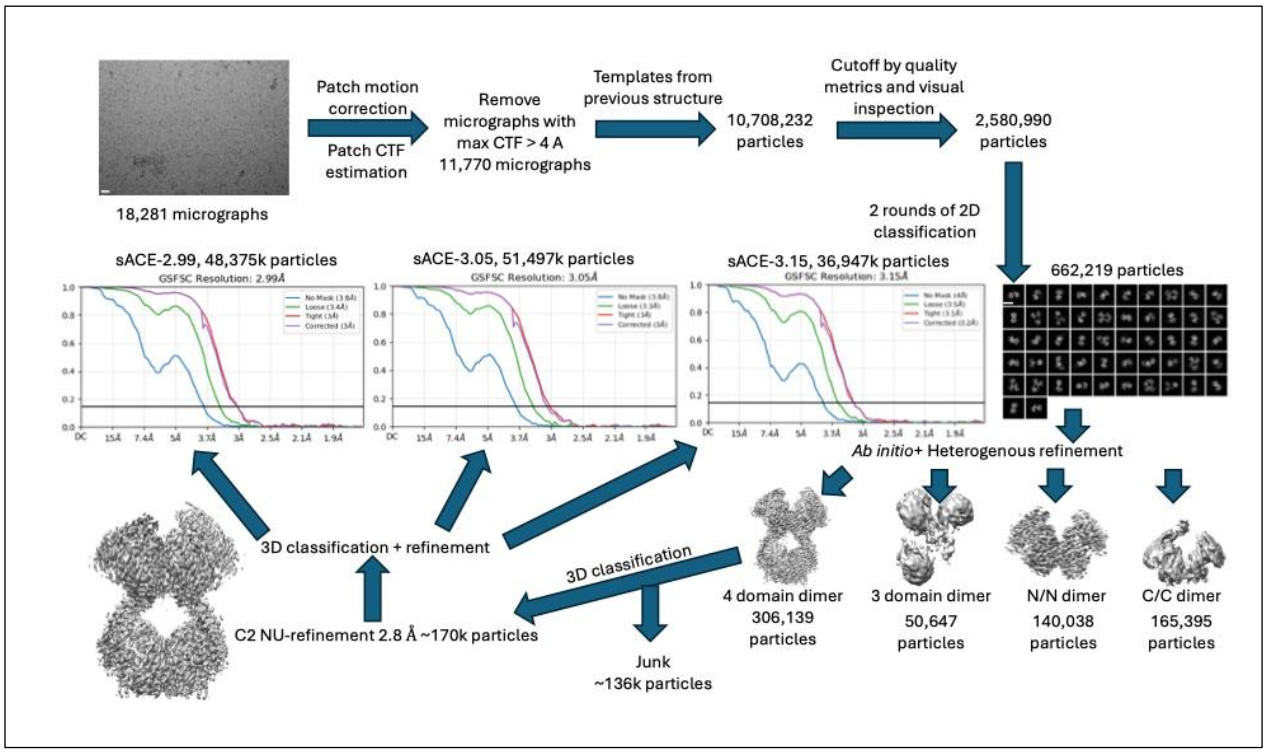


Figure S5: Assessment of map quality from Chameleon prepared grids. Angular distribution, local resolution, example density fit and 3D-FSC curves for (A) sACE-3.05, (B) sACE-3.15, and (C) sACE-2.99.

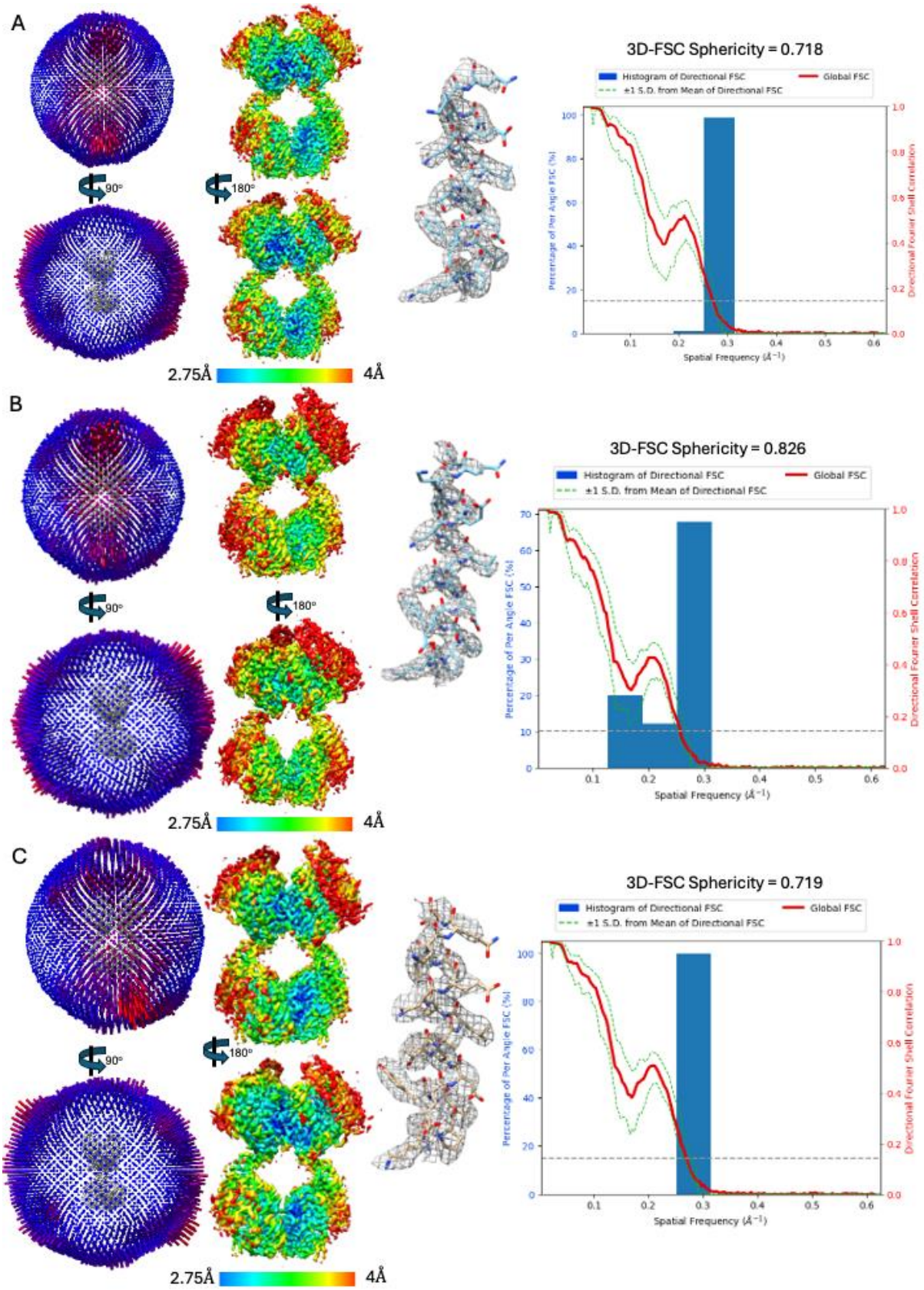

A
3D-FSC Sphericity = 0.718
Histogram of Directional FSC
Global FSC
±1 S.D. from Mean of Directional FSC
Percentage of Per Angle FSC (%)
Directional Fourier Shell Correlation
Spatial Frequency (Å⁻¹)
90°
180°
2.75Å
4Å
B
3D-FSC Sphericity = 0.826
Histogram of Directional FSC
Global FSC
±1 S.D. from Mean of Directional FSC
Percentage of Per Angle FSC (%)
Directional Fourier Shell Correlation
Spatial Frequency (Å⁻¹)
90°
180°
2.75Å
4Å
C
3D-FSC Sphericity = 0.719
Histogram of Directional FSC
Global FSC
±1 S.D. from Mean of Directional FSC
Percentage of Per Angle FSC (%)
Directional Fourier Shell Correlation
Spatial Frequency (Å⁻¹)
90°
180°
2.75Å
4Å

Figure S6: B factor comparison of all structures. Full-length sACE structures colored by B factor. Scale ranges from a B factor of 10 (dark blue) to 225 (red). Protein is depicted as cartoon, glycans are shown as sticks.

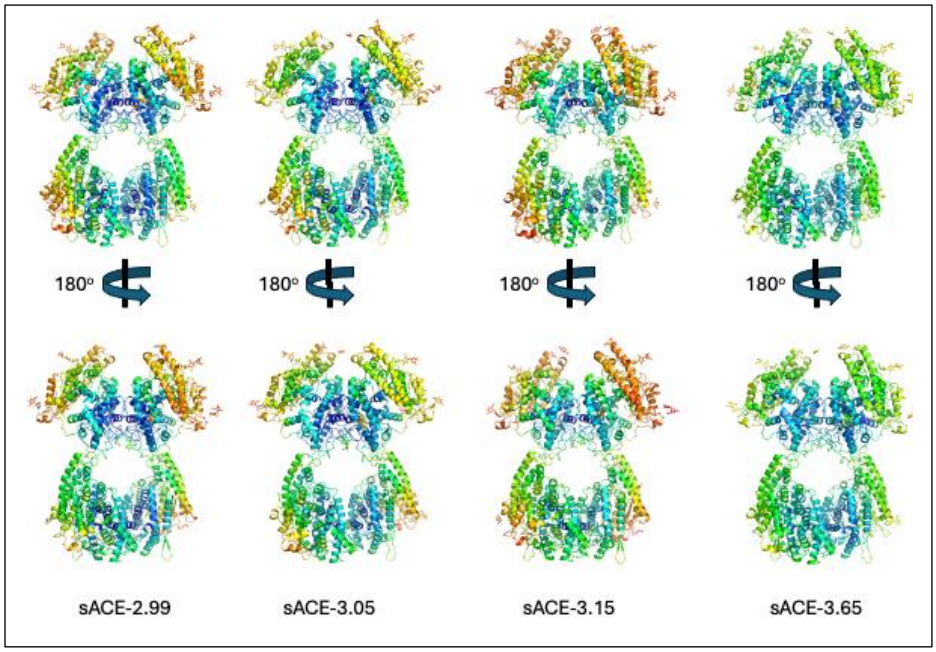
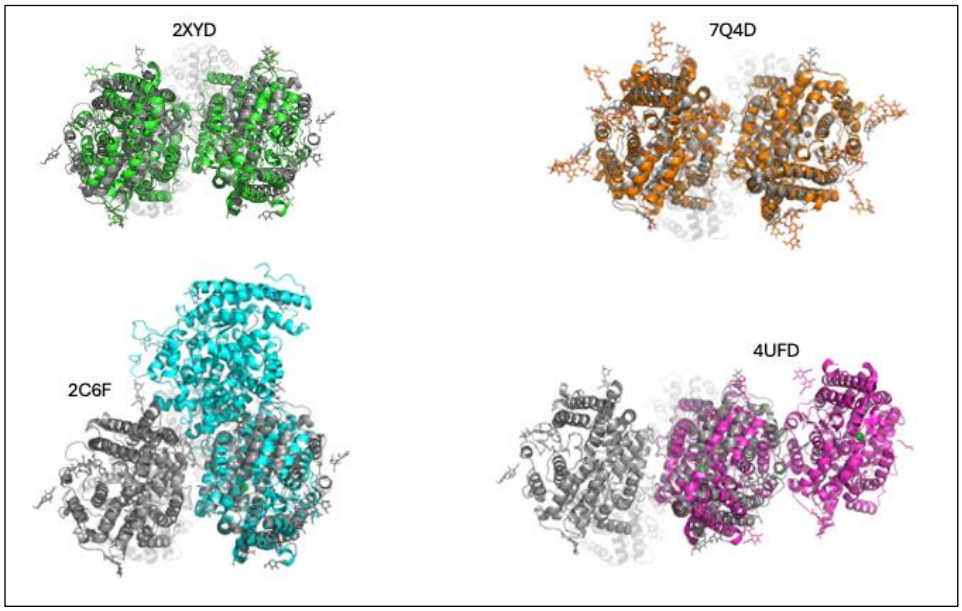
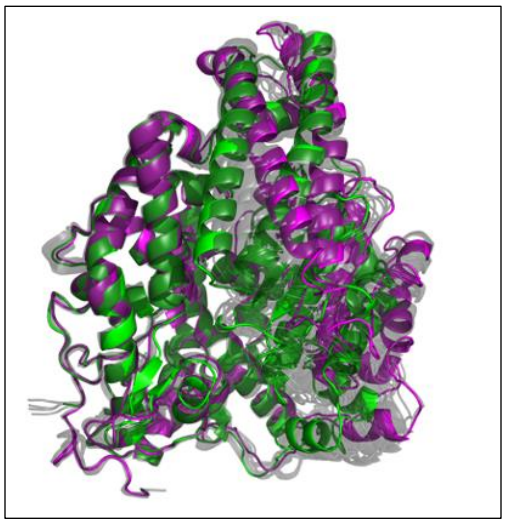


Figure S7: sACE-N/N dimers compared to previously published structures. Previously identified sACE-N/N dimers 2XYD (green, which is similar to structures of following pdb codes: 3nxq, 4bxk, 4bzs, 4ca6, 4ufa, 6f9r, 6f9v, 6h5x, 6qs1, 6tt1, 6tt3, 6tt4, 7q25, 7q26, 7q27, 7q29, 7z6z), 2C6F (cyan, which is similar to structures of following pdb codes: 2c6n, 5amb, 5amc), 4UFB (magenta, which is similar to structures of following pdb codes: 5am8, 5am9, 5ama, 6en5, 6en6, 6zpq, 6zpt), and 7Q4D (orange) aligned to sACE-2.99 (gray).



1027 Figure S8: sACE exhibits a continuous gradient of structural heterogeneity. All 16
1028 individual domains from our sACE structures were aligned to their respective D2/4
1029 regions, revealing a continues gradient of “openness” (gray) from the most closed
1030 conformation (green) to the most open conformation (magenta).



1031

1032 Figure S9: RMSD analysis of non-glycosylated MD simulations. RMSD values were
1033 calculated for the D1/3a (red), D1/3b (gray), and D2/4 (orange) regions and compared to
1034 the RMSD values for the global domain (blue) to identify the primary source of structural
1035 variance in the non-glycosylated sACE MD simulations. RMSD was calculated using
1036 backbone atoms only. For each panel, the Y-axis is RMSD value, the X-axis is simulation
1037 frame. Every frame represents a time step of 0.1 ns.

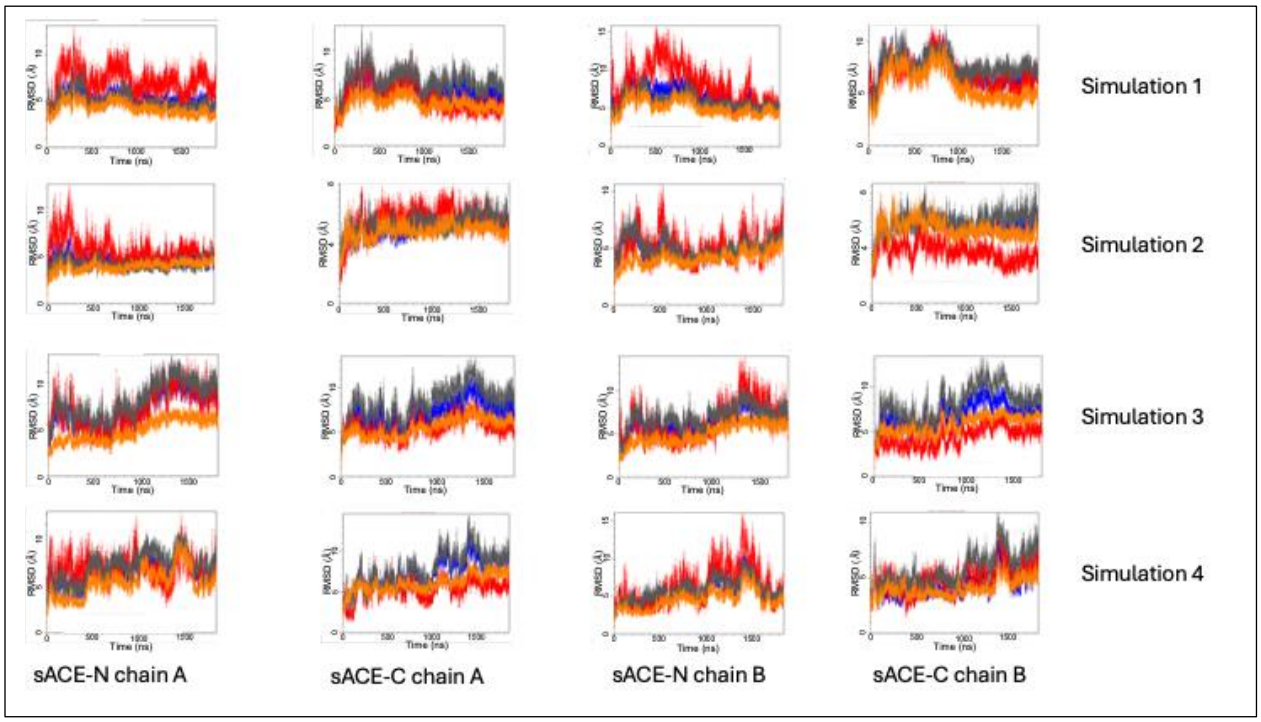
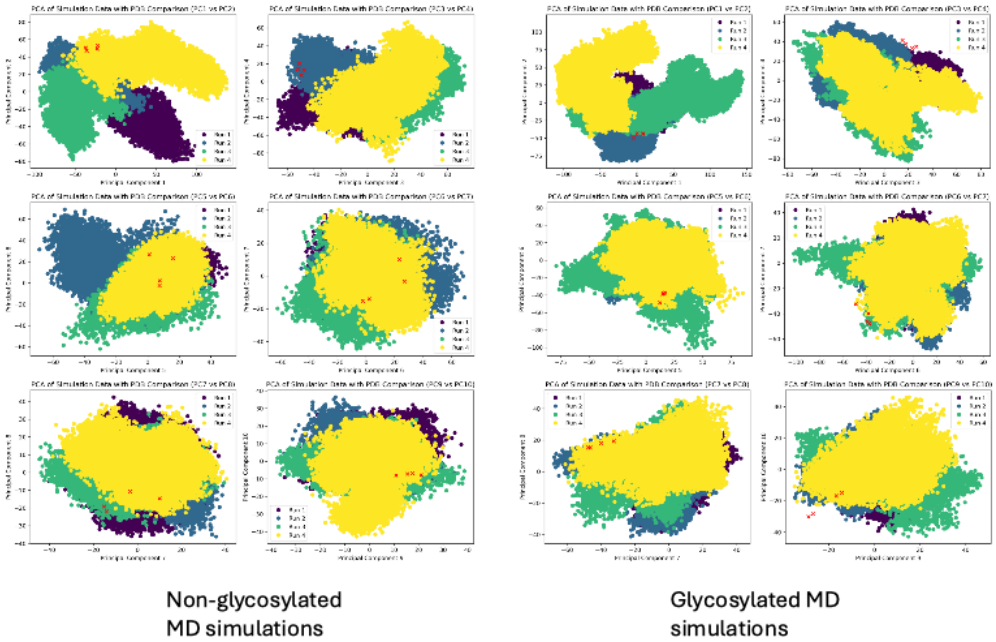


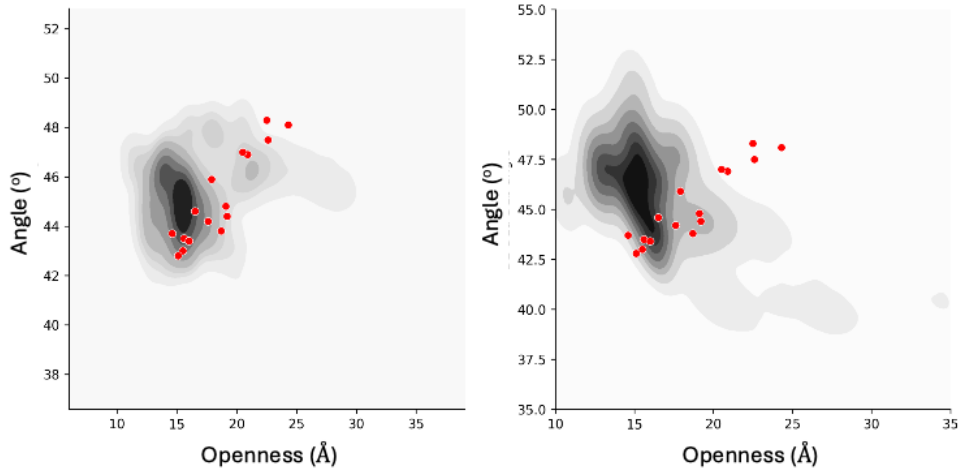
Figure S10: Principal Component Analysis. PCA was performed using the alpha carbon coordinates for every residue in every frame of our MD simulations along with each of our cryo-EM structures. Red X's denote data from cryo-EM structures. MD simulation data is colored by run as indicated. Non-glycosylated MD simulation data is shown on the left, glycosylated MD simulation data is shown on the right. The conformational space explored in our MD simulations is significantly greater than the range of conformations observed in our cryo-EM structures, as expected. The most significant source of conformational variation in our MD simulations relative to our cryo-EM structures is the reposition of sACE-C/C relative to sACE-N/N (Movies S7,8).



Non-glycosylated MD simulations

Glycosylated MD simulations

Figure S11: Clustering Analysis. The conformational space sampled by each domain in our non-glycosylated (left) and glycosylated (right) MD simulations was assessed by tracking the angle formed by the centers of mass of the D2-D1a-D1b (or D4-D3a-D3b for sACE-C) relative to the openness of the domain (D1/3a-D2/4 measurement as shown in figure 2) for each frame of our simulations and plotted as density. The corresponding measurements from our cryo-EM structures are shown in red.



Non-glycosylated

Glycosylated

Figure S12: RMSF comparison of glycosylated vs non-glycosylated MD simulations. RMSF values for each residue over the course of all glycosylated and non-glycosylated MD simulations. The RMSF values for each residue in the non-glycosylated simulations were subtracted from the values for each residue in the glycosylated simulations and plotted. Chain A residues are shown in blue, chain B residues are shown in orange.

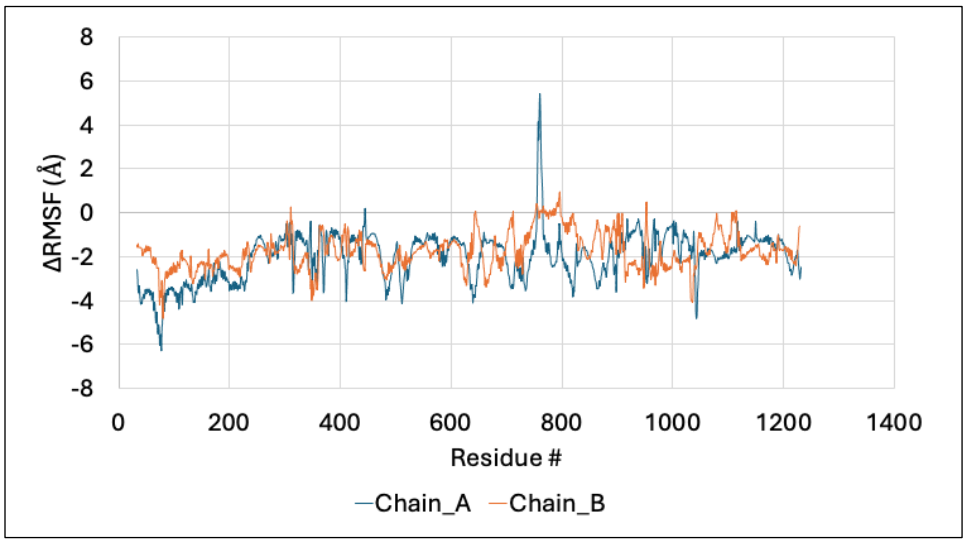


Figure S13: Shielding of glycans in MD simulations. 200 representative frames of glycan orientations (light blue) were overlaid the sACE structure (pink shade by chain) to visualize the amount of protein surface insulated from solvent access by glycans.

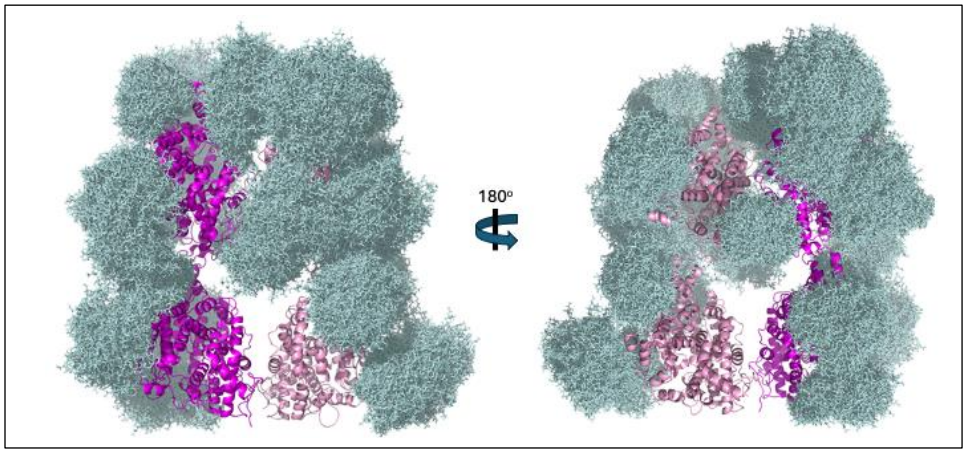


Figure S14: CryoDRGN clustering of particles. Top panel shows the deconvolution of the latent space along specified principal component vectors and the path of each trajectory generated as movies (see Movie S10). Bottom panels show particle clustering resulting from UMAP deconvolution of the latent space with the path of each trajectory generated.

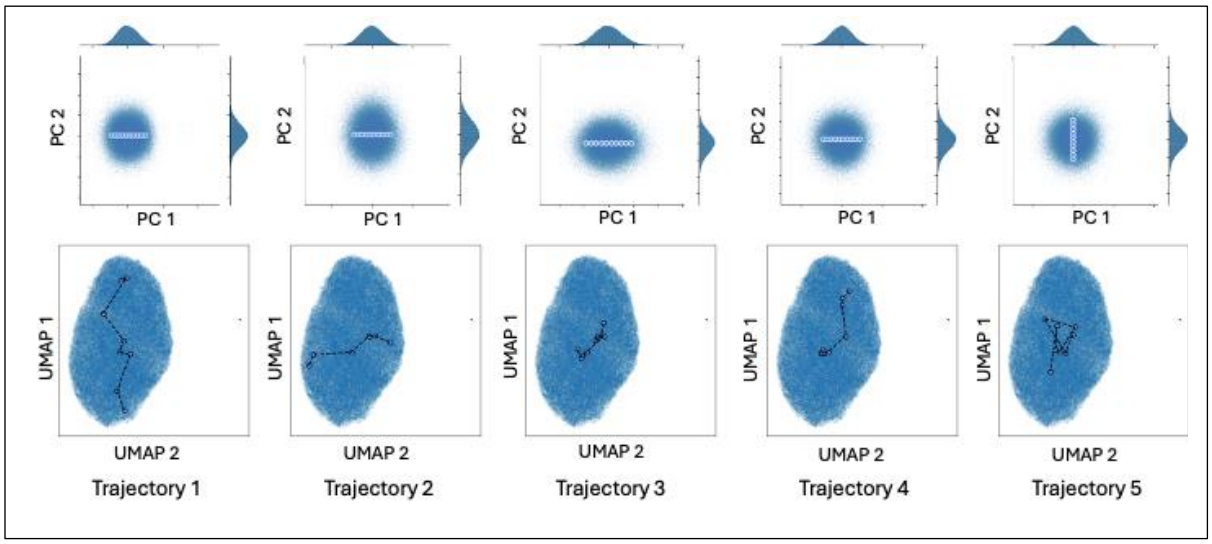
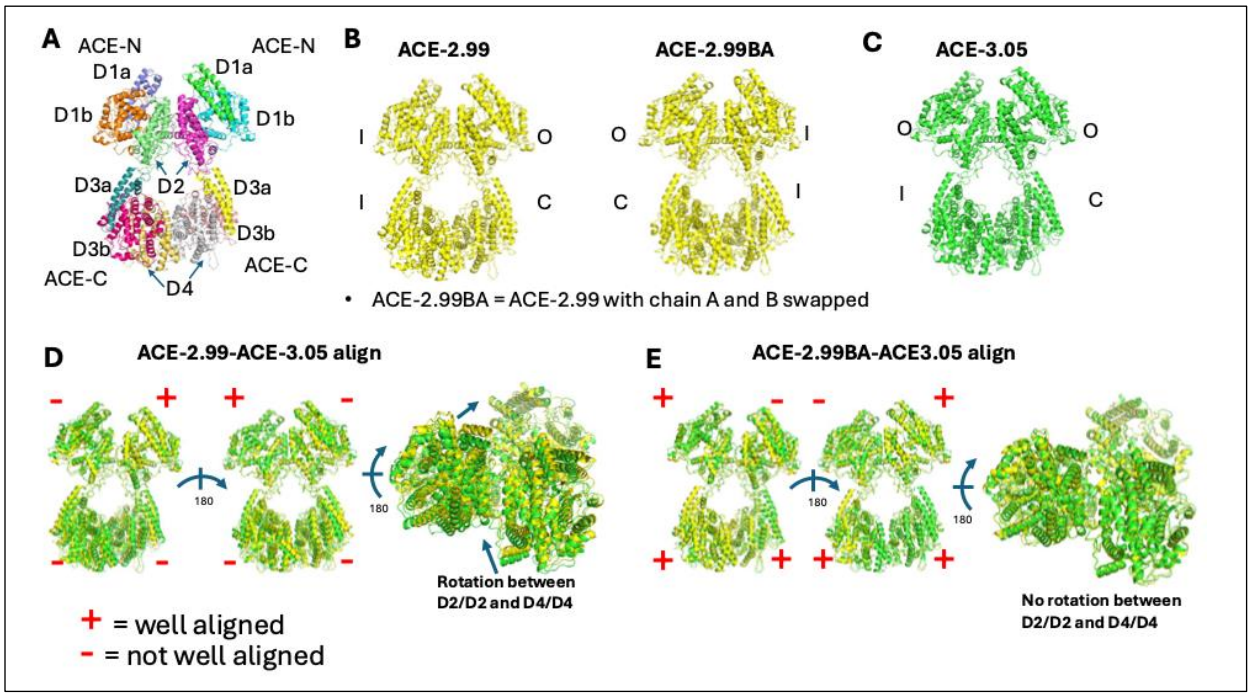


Figure S15: Artificially observed interdomain motion caused by the subtle yet noticeable differences between the two chains within the sACE dimer and the choice of chain pairs between sACE dimer structures. A. Overall structure of the sACE dimer and domain assignment. B. Ribbon representation of sACE-2.99 structures where the assignment of chain A and chain B within the sACE dimer is swapped. C. Ribbon representation of the sACE-3.05 structure. D. Alignment of the sACE-2.99 structure with sACE-3.05. When sACE-2.99 A/B chains are aligned with sACE-3.05 A/B chains, one domain is well-aligned but the other three are not. There is an ~5° rotation of D4/D4 relative to D2/D2 from the alignment. E. The alignment of two structures that are the same as D except chain A and chain B in sACE-2.99 is swapped. When sACE-2.99 B/A chains are aligned with sACE-3.05 A/B chains, three domains are well-aligned and the remaining one depicts the open-closed transition of sACE catalytic domain. In this case, D2/D2/D4/D4 of the two structures are well aligned and there is no rotation.



Supplemental tables

Table S1: List of buffers used in DSF optimization of grid-making conditions.

Table S2: CryoEM processing statistics for both datasets

Table S3: Summary of MD simulations

Table S4: Summary of heterogeneity analysis

Supplemental movies

Movie S1: Overview of sACE-3.65 structure. Coulomb potential density map colored by subdomain as Figure 1, followed by cartoon-depicted model of sACE-3.65 structure (colored by sub-domain) fit into its Coulomb potential map (gray), and then cartoon-depicted model of sACE-3.65 structure (colored by sub-domain).

Movie S2: Overview of sACE-2.99 structure. Coulomb potential density map colored by subdomain as Figure 1, followed by cartoon-depicted model of sACE-2.99 structure (colored by sub-domain) fit into its Coulomb potential map (gray), and then cartoon-depicted model of sACE-2.99 structure (colored by sub-domain).

Movie S3: Overview of sACE-3.05 structure. Coulomb potential density map colored by subdomain as Figure 1, followed by cartoon-depicted model of sACE-3.05 structure (colored by sub-domain) fit into its Coulomb potential map (gray), and then cartoon-depicted model of sACE-3.05 structure (colored by sub-domain).

Movie S4: Overview of sACE-3.15 structure. Coulomb potential density map colored by subdomain as Figure 1 followed by cartoon-depicted model of sACE-3.15 structure

(colored by sub-domain) fit into its Coulomb potential map (gray), and then cartoon-depicted model of sACE-3.15 structure (colored by sub-domain).

Movie S5: Non-glycosylated sACE MD simulation. One microsecond excerpt from a representative sACE all-atom MD simulation without modeled glycans. Protein shown as cartoon and colored by subdomain. Solvent molecules hidden for clarity.

Movie S6: Glycosylated sACE MD simulation. One microsecond excerpt from a representative sACE all-atom MD simulation with modeled glycans. Protein shown as cartoon and colored by subdomain. Glycans shown as gray sticks . Solvent molecules hidden for clarity.

Movie S7: Non-glycosylated PCA trajectories were generated as linear interpolations representing the conformational changes described along each of the top 10 principal components from our analysis (See figure S10). Protein is colored by subdomain.

Movie S8: Glycosylated PCA trajectories were generated as linear interpolations representing the conformational changes described along each of the top 10 principal components from our analysis (See figure S11). Protein is colored by subdomain. Modeled glycans are not shown for clarity, and due to the fact that PCA was performed using only alpha carbons of protein residues.

Movie S9: Cryosparc 3D variability analysis. Trajectory movies representing structural variability along the top 5 principal component vectors of variance. Trajectories depict a volume series of the representative structure generated by averaging particles in every tenth percentile along the principal component vector.

Movie S10: CryoDRGN heterogeneity analysis. Trajectory movies representing structural variability along the top 5 principal component vectors of variance. Trajectories depict a volume series of the representative structure generated by averaging particles in every tenth percentile along the principal component vector. See Figure S14 for the path each trajectory follows through PCA and UMAP space.

Movie S11: RECOVAR trajectory. RECOVAR was used to generate a trajectory through the deconvoluted multidimensional latent space described in Figure 5C, from the sACE-N/N focus masked analysis. The path was calculated from the latent point representing the closed-closed state to the latent point representing the open-open state. Representative structures were generated at six equidistant points along the path and stitched together as a volume series to generate the trajectory.

Movie S12: 3DFlex trajectory. The most dynamic result from CryoSPARC 3DFlex is shown. Results were generated using a default tetrahedral mesh and inputting the modes from 3DVA. Despite this, the motions observed in 3DVA (see movie S9) were not observed in 3DFlex. The most significant motions observed was a slight opening of sACE-C (red arrow) and “rocking” of sACE-N/N.