Characterization and modulation of human insulin degrading enzyme conformational dynamics to control enzyme activity

Jordan M. Mancl[1], Wenguang G. Liang[1,3], Nicholas L. Bayhi[2], Hui Wei[3], Bridget Carragher[3,5], Clinton S. Potter[3,5] and Wei-Jen Tang[1]*

1. Ben-May Institute for Cancer Research, The University of Chicago, 929 East 57th Street, Chicago, Illinois 60637, USA

2. Biophysics Science Graduate Program, The University of Chicago, 929 East 57th street, Chicago, Illinois 60637, USA

3. Simons Electron Microscopy Center, New York Structural Biology Center, 89 Convent Ave, New York, New York, 10027, USA

4. Current affiliation - Bayer HealthCarePharmaceuticals LLC

5. Current affiliation - Chan Zuckerberg Institute for Advanced Biological Imaging

6. Current affiliation - Rutgers University

* Corresponding author

Email: wtang@bsd.uchicago.edu

Author Contributions: W.-J.T., J.M.M., W.G.L. designed the project. W.G.L. performed protein purification, cryoEM grid preparation, data acquisition and processing assisted by H.W. and overseen by W.-J.T., B.C. and C.S.P. J.M.M. and W.G.L. built and refined structural models. J.M.M. performed heterogeneity analysis of cryoEM structures, MD simulations, and mutant characterizations. N.L.B performed the initial MD simulations. J.M.M. W.-J. T. wrote the manuscript and all authors contributed to manuscript finalization.

Acknowledgement: We would like to thank our colleagues at the University of Chicago: Adam Antoszewski and Aaron Dinner for providing us with the data from their insulin unfolding simulations, as well as Rong Shen for general guidance with the MD simulations, and Jessica Martinez Martinez for assistance with cloning the IDE R668A construct. We would also like to thank James Byrnes at the LiX beamline at Brookhaven National Laboratory for collecting SEC-SAXS data. Computer resources came, in part, from an allocation on the Beagle3 computing cluster at the University of Chicago. This work was supported by the NIH grant GM121964 to W.-J. Tang. The Simons Electron Microscopy Center is supported by a grant from the Simons Foundation (SF349247). Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, under contract No. DE-AC02-06CH11357. The LiX beamline is supported by NIH grants P30GM133893 & S10 OD012331, DOE grants KP1605010 & DE-SC0012704.

Competing Interest Statement: None.

Keywords (5 max): amyloid peptides, insulin degrading enzyme, proteostasis, diabetes, cryo-electron microscopy

Abstract

Insulin degrading enzyme (IDE) is a dimeric 110 kDa M16A zinc metalloprotease that degrades amyloidogenic peptides diverse in shape and sequence, including insulin, amylin, and amyloid-β, to prevent toxic amyloid fibril formation. IDE has a hollow catalytic chamber formed by four homologous subdomains organized into two ~55 kDa N- and C- domains (IDE-N and IDE-C, respectively), in which peptides bind, unfold, and are repositioned for proteolysis. IDE is known to transition between a closed state, poised for catalysis, and an open state, able to release cleavage products and bind new substrate. Here, we present five cryoEM structures of the IDE dimer at 3.0-4.1 Å resolution, obtained in the presence of a sub-saturating concentration of insulin. Analysis of the heterogeneity within the particle populations comprising these structures combined with all-atom molecular dynamics (MD) simulations permitted a comprehensive characterization of IDE conformational dynamics. Our analysis identified the structural basis and key residues for these dynamics that were not revealed by IDE static structures. Notably arginine-668 serves as a molecular latch mediating the open-close transition and facilitates key protein motions through charge-swapping interactions at the IDE-N/C interface. Our size-exclusion chromatography-coupled small-angle X-ray scattering and enzymatic assays of an arginine-668 to alanine mutant indicate a profound alteration of conformational dynamics and catalytic activity. Taken together, this work highlights the power of integrating experimental and computational methodologies to understand protein dynamics, offers the molecular basis of unfoldase activity of IDE, and provides a new path forward towards the development of substrate-specific modulators of IDE activity.

Introduction

Protein homeostasis (a.k.a. proteostasis) is maintained by three primary mechanisms: chaperones, ubiquitination/proteasome, and autophagy (1, 2). Disruptions in proteostasis can lead to the accumulation of amyloid fibrils and subsequent human diseases (3–5). As a result, many proteases have evolved to specifically target amyloidogenic peptides (6, 7). Among those, the cryptidase family, which includes the M16 metalloproteases insulin degrading enzyme (IDE) use an internal catalytic chamber, or "crypt", to capture and selectively degrade the monomeric form of amyloid peptides to control the formation of amyloid fibrils (6, 8). IDE effectively degrades various bioactive peptides, including amyloid-β (Aβ), a peptide associated with the progress of Alzheimer's disease, and three blood glucose-regulating hormones: insulin, amylin, and glucagon (9). Consequently, defects in IDE alter the progression of type 2 diabetes mellitus and Alzheimer's disease in animal models and are linked to these diseases in humans (9–17). IDE is a promising therapeutic target, as its inhibition improves glucose tolerance, yet progress has been hampered by the contrary actions of its diverse substrate pool (9, 18, 19). For example, IDE overexpression has been shown to reduce Aβ loads, in mice, however, this reduction is accompanied by hyperglycemia from a similar reduction in insulin levels (20). Therefore, a deeper understanding of the molecular mechanisms underlying IDE catalysis is required to facilitate the development of future therapeutics.

Integrative structural approaches have provided key insights into how IDE selectively degrades amyloidogenic peptides (6, 9, 21). IDE is a dimeric 110 kDa metalloprotease. Each protomer is comprised of four homologous subdomains organized into two ~55 kDa domains, IDE-N and IDE-C, connected by a short flexible linker (22, 23). Our previous cryoEM analysis revealed the major conformational states of IDE. In the absence of substrate, each protomer was found to adopt an open (O) or partial open (pO) state which are primarily differentiated by the displacement of IDE-N relative to IDE-C, the combination of which led to three dimeric conformations (O/O, O/pO, pO/pO) (21). It was observed that substrate binding induced both protomers of IDE to fully close, referred to as the partial closed (pC, pC/pC for the dimer) state, as the cryoEM closed state was found to be slightly more open than previously solved crystal structures, likely influenced by the constraints of the crystal lattice (22, 24). Only the O state permits substrate access to the catalytic chamber, yet the catalytic cleft is stabilized only in the pC state with substrate bound, requiring IDE to undergo a substantial open-close transition during catalysis, the specifics of which are unknown (21).

It has been established that many IDE substrates must be unfolded prior to cleavage but, as IDE substrates are highly diverse in size and structure, the mechanistic basis of this unfoldase activity has proven elusive. We hypothesized that the unfoldase activity of IDE is mediated by the relative motions of IDE-N and IDE-C. To test this hypothesis, we solved cryoEM structures of IDE with a 2:1 IDE:insulin ratio and used recently developed computational approaches to deconstruct the particle heterogeneity within our structures to infer conformational dynamics. Integration of this experimental data with all-atom molecular dynamics (MD) simulations allowed us to identify and modulate the molecular interactions governing key conformational dynamics to manipulate the enzymatic activity of IDE *in vitro*.

Results

CryoEM structures of IDE in the presence of a sub-saturating concentration of insulin

We have previously used cryoEM to solve the structure of IDE in the presence of a 5-fold molar excess of insulin, resulting in a dominant pC/pC state with both substrate binding chambers yielding density corresponding to unfolded insulin (21). To explore the mechanism of IDE-directed substrate unfolding under more physiologically relevant conditions, here we investigated IDE-insulin interactions with an IDE:insulin molar ratio of 2:1. We collected a dataset of ~7,600 micrographs on a Titan Krios operated at 300 keV, from which 7.2 million particles were picked for processing in RELION. Following an established workflow, we generated 5 structures: three previously observed structures with improved resolution (pC/pC, O/O, and O/pO at 3.0 Å, 3.8 Å, and 4.1 Å resolution) and two novel states (O/pC and pO/pC at 3.4 Å and 3.3 Å resolution, respectively) (Fig. 1A, S1, Table 1). Interestingly, the pC/pC was still found to be dominant, despite IDE and insulin being present at a 2:1 molar ratio, consistent with the allostery of IDE (21,24).

The individual protomers of IDE adopt the same three conformations, O, pO, and pC states reported previously but there are significant differences in the O and pO states (Fig. 1B) (21). While the pC/pC state is nearly identical to the two states reported previously (RMSD of ~0.6 Å; PDB ID: 6BFC, 6B3Q) (21), the O and pO states appear more closed when insulin is present, suggesting that the presence of insulin promotes the open to closed transition. Specifically, the pO and O subunits are 3-5 Å more closed based on the distance between the centers-of-mass (COM) of the IDE D1 and D4 subdomains, while the dihedral angle formed by the COM of the D1-D2-D3-D4 subdomains is reduced by 5-10^{o} (Fig. 1B). Alignment of the new O/O and O/pO states to their previously solved counterparts reveals little difference, with a global RMSD of 1.316 Å (O/O, PDB ID: 6B7Y) and 1.405 Å (O/pO, PDB ID: 6BF8), respectively (21). The primary source of variation between the different states remains the degree of opening between the two domains of each subunit, as a global alignment of IDE-N and IDE-C across all structures reveal the domains primarily act as rigid bodies (RMSD <1.4 Å). Consistent with previous observations, the door region, which contains the catalytic zinc binding site, exhibited higher B-factors in the O states than in either the pO or pC state (21) (Fig. S2).

All five cryoEM structures presented here contain four subunits that have clear density present at the catalytic cleft and exosite indicative of bound substrate, thus adopting the pC state. Previous structures have modeled the same chain of insulin in both the exosite and catalytic cleft, but the best models we could build into our density suggest that the chain bound to the exosite is not the same chain positioned for cleavage at the catalytic cleft (Fig. 1C). The preponderance of data suggests it to be unlikely that insulin adopts a preferred orientation within the catalytic chamber, rather the catalytic chamber can accommodate insulin in four different orientations, with either the A or B chain bound to the exosite and catalytic cleft in *cis* and *trans* orientations. This promiscuity of binding has been theorized as a reason for why there is no chain preference for the initial insulin cleavage event (25). Crystal structures of insulin are highly compact and globular, with little spatial separation between the N- and C- termini of each chain, and the IDE cleavage

sites are located within α-helices (26). In its crystallized state, insulin cannot interact with both the exosite and catalytic cleft unless it unfolds, at least partially. Recent work used MD simulations to study the conformational dynamics of insulin in solution and identified several major "elements of disorder" to describe the partially unfolded intermediate structures they observed (27). We docked these structures of insulin into the closed catalytic chamber of IDE and found that the chamber was able to easily accommodate the positioning of either chain near the exosite and catalytic cleft in *cis* and *trans* orientations (Fig. S3).

Structural heterogeneity of cryoEM structures

Numerous structures of IDE exist with subunits adopting either an open or closed conformation, yet we lack an understanding of how IDE transitions from an open to a closed state. The simplest explanation of this transition, based on structural data, would be a direct, rigid body translation of IDE-N relative to IDE-C, and this is the predominant model within the field (21, 22). To better understand what this transition would look like, we measured and plotted the changes in the D1-D4 COM distance versus the changes in the D1-D2-D3-D4 COM dihedral angle for all available cryoEM structures of IDE (Fig. S4). We found that in the absence of substrate, the transition pathway produced a linear relationship between states. However, when our structures generated in the presence of insulin were analyzed, the linear relationship no longer held (Fig. S4), suggesting that the simple linear transition typically presumed for the open-close transition between distinct IDE states does not accurately depict the complexity of IDE dynamics.

Recently, several approaches have been developed to understand the conformational heterogeneity present within cryoEM data. We employed multibody analysis in RELION and 3D variability analysis (3DVA) in cryoSPARC to investigate the conformational heterogeneity within our particle populations (28, 29). Of these, the range of motion predicted by multibody analysis most closely matched the reported structures while 3DVA exhibited consistent results across a smaller magnitude of transition (Fig. S4-S6). Therefore, we focused on multibody analysis below. Multibody analysis models the structural heterogeneity in our data as the result of motions of independent, user-defined rigid bodies. As discussed above, the primary source of structural variation in our structures is the positioning of IDE-N relative to IDE-C. We see essentially no structural changes within the domains, suggesting the domains behave as rigid bodies, an assertion further supported by our all-atom molecular dynamics (MD) simulations (Fig. 2A). Multibody analysis defines the principal components of variance along discrete eigenvectors, which we interpret as proxies for the dominant components of molecular motion. We quantified the change in D1-D4 COM distance and D1-D2-D3-D4 COM dihedral of representative structures along the heterogeneity gradients of the top 9 eigenvectors for each of our IDE structures, representing ~70-90% of the total structural heterogeneity per structure. The results indicated that the particles comprising our cryoEM structures sampled a significantly greater conformational space than would be expected from analysis of the ensemble structures alone (Fig. S5). Most intriguingly, the conformational changes displayed an unexpectedly high degree of change in the D1-D2-D3-D4 COM dihedral angle, indicating a significant rotation of IDE-N relative to IDE-C (Fig. 2B,C, Fig. S5). Integration of the multibody results across all states of IDE suggest two dominant components of structural variance. First is a translation-dominant

conformational change wherein IDE-N swings toward or away from IDE-C as if mediated by a hinge formed by the interdomain linker region of IDE (Fig. 2D, Movie S1). This conformational change closely resembles our understanding of the IDE open-closed transition inferred from analysis of the ensemble structures. The second component is a rotation-dominant conformational change wherein IDE-N rotates orthogonal to the plane of the dimer as if it were being screwed into or ground against IDE-C (Fig. 2E, Movie S2). Importantly, these dominant components of structural variance were also observed when the particle populations were analyzed with cryoSPARC's 3DVA implementation, albeit over a smaller magnitude of conformational change (Fig. S6). These dominant components also correlate well with the lowest frequency modes revealed by normal mode analysis (Movies S3, S4). Interestingly, while the presence of insulin was found to substantially influence the consensus reconstructions of IDE, we found that the presence or absence of insulin had no significant effect on the principal components of structural variance for IDE, consistent with the current understanding that enzyme conformational changes are "hardwired" into the structure (30, 31).

Molecular basis of the open-close transition revealed by all-atom MD simulations

Analysis of the particle heterogeneity comprising cryoEM structures provides excellent information about the conformational space sampled by the protein of interest. However, such analysis lacks a temporal component, and thus only implies motions. To overcome this barrier and examine how the conformation of IDE changes over time, we performed six replicate all-atom MD simulations with explicit solvent. Simulations were started from the O/pO state (PDB: 7RZG), with the missing loops modeled in to generate continuous peptide chains and run for one microsecond. By starting with the O/pO structure, we investigated the dynamics of both the open and closed subunits simultaneously. We observed that the conformational space sampled by our simulations correlated well with the conformational space sampled by our cryoEM particle population, as revealed by multibody analysis (Fig. S7). In 5 of our 6 simulations, the open subunit closed quickly, typically in <200 ns (Fig. 3A). Analysis of the open-close transition revealed two key findings. First, the open subunits did not close to a singular structure. We define closing as a subunit reaching a similar D1-D4 COM distance as the pO or pC states, but we observed distinct differences among the subunits indicative of IDE-N rotation relative to IDE-C. These differences are in line with our second key finding, that the conformational changes associated with the open-close transition did not follow the same general pathway among the simulations, as we observed a high degree of variability upon closing (Fig. 3B). Despite this variability, R668 consistently stood out as one of the first IDE-C residues to form new interactions with IDE-N, either D309 or E381 (Fig. 3C). In most simulations, formation of this interaction preceded a large decrease in the D1-D4 COM distance (Fig. 3D, Fig. S8). Interestingly, these R668 interactions are not present in the ensemble cryoEM or crystal structures of IDE. We hypothesize that the biochemical properties and spatial localization of R668 enable it to essentially reach out and grab onto the N-domain, initiating the formation of a complex hydrogen bonding network that drives the open-close transition towards completion. Previously published hydrogen-deuterium exchange mass spectrometry (HDX-MS) data indicates that R668 is stabilized under conditions that promote IDE closing, suggesting that R668 may function as a sort of molecular latch (Fig. 3E)(21).

To further probe the importance of R668, we ran 6 additional simulations with IDE carrying a R668 to alanine (R668A) point mutation under identical conditions as described previously. These simulations displayed significantly altered dynamics when compared to the wild-type simulations. Notably, the R668A mutant was found to close more slowly, and many simulations did not reach the D1-D4 COM distance typified by a closed state structure, instead stabilizing with a significantly larger D1-D4 COM distance than we had previously observed in our wild type (WT) simulations or our experimental structures (Fig. 3F). We also observed that the R668A mutant preferentially sampled a subset of conformational space rarely explored in our WT simulations and displayed significantly greater variation in the D1-D2-D3-D4 dihedral angle (Fig. 3G). As a result of this increased rotational exploration, several of our R668A simulations closed to a conformation where IDE-N is offset and rotated relative to IDE-C when compared to the experimentally determined structures (Fig. S9). This altered conformation results in a structure where IDE-N and IDE-C are fully engaged, but the offset produces a deceivingly high D1-D4 COM distance. With this in mind, we sought to better understand the conformational dynamics of IDE in the closed (pO) state.

R668A mediates IDE conformational dynamics *in vitro*

The R668A mutation was found to significantly alter the conformational dynamics of IDE's open-close transition in our simulations (Fig. 3F, G), so we expressed and purified the R668A mutant to determine if the changes predicted from our MD simulations resulted in altered enzymatic activity relative to WT. Previous work has established that IDE adopts a dominant O/pO state in solution (21). Consistent with our MD simulations, size-exclusion chromatography revealed that the R668A construct eluted slightly earlier than WT IDE, suggesting that the R668A mutation induces a greater proportion of molecules to adopt a more open conformation, possibly shifting to a dominant O/O state (Fig. 4A). Next, we compared the enzymatic activity of the R668A and WT constructs using the fluorogenic substrate (7-methoxycourmarin-4-yl)acetyl-RPPGFSAFK(2,4-dinitrophenyl)-OH (substrate V), a bradykinin mimetic which has previously been used to characterize the enzymatic activity of IDE and related enzymes (24, 32). We found that WT IDE has activity consistent with previously published data (24) and the R668A mutation produced a ~5-fold decrease in activity (Fig. 4B). ATP is known to promote IDE degradation of small substrates, i.e. bradykinin, but not larger substrates, i.e. insulin (33). Consistent with this observation, the addition of ATP increased the activity of both WT and R668A constructs (Fig. 4B). We also tested if the R668A mutation would alter the degradation of larger, well-folded substrates, using insulin in a direct competition assay. We found that, after accounting for the previously observed 5x decrease in activity, insulin yielded an apparent K_i of ~8 nM for WT IDE, but this value was reduced to ~52 nM for the R668A construct (Fig. 4C). We then performed size exclusion chromatography-coupled small-angle X-ray scattering (SEC-SAXS) for both constructs to ascertain if the observed differences in enzymatic activity could be explained by altered biophysical properties. The WT scattering profile produced an R_g of 49.4 +/- 0.9 Å, consistent with previous SEC-SAXS results (21)(Fig. 4D, E). The R668A scattering profile produced an R_g of 54.7 +/- 0.4 Å, consistent with our SEC data that R668A mutant has slightly larger hydrodynamic radius (Fig. 4D). Together, our data indicates

that R668A mutation can profoundly affect the conformational dynamics governing the open-closed transition, which leads to the altered substrate binding and catalysis.

Charge-swapping at the IDE-N/C interface mediates conformational dynamics in the closed state

The WT pO subunit remained closed throughout all simulations yet was found to initially relax to a slightly more open state within the first 200 ns, roughly the same time frame it took the WT O subunit to close (Fig. 4F) and, despite not opening, demonstrated a range of D1-D2-D3-D4 dihedral variation similar to that of the O subunits (Fig. 4E). Interestingly, the conformational variation within our simulations increased greatly after this relaxation point was reached (Fig. 4F). This relaxed conformational state consistently displayed an altered IDE-N/C interface compared to the experimentally determined structures that was replicated by the open WT subunits upon closing, leading to a nearly symmetric pO/pO state (Fig. S11). Biochemically, this relaxed IDE-N/C interface makes more sense than the interface observed experimentally. It has previously been demonstrated that the experimental constraints of crystallographic structure determination force IDE to adopt a closed conformation and determination of the open state can only be accomplished once those restraints are removed, i.e. cryoEM (21). These results suggest that both ensemble methods of structure determination may impact the conformation of IDE, however slightly. The IDE-N/C interface is littered with charged residues, yet few interactions are observed in the ensemble structures (Fig. 5A), whereas our MD simulations reveal a complex hydrogen bonding network (Fig. 5B). We observed an extensive network at the D2-D3 interface, with sporadic patches of interactions between D1 and D4 (Fig. S12). Within the D2-D3 network R668 again stands out as a notably key residue. In addition to forming hydrogen bonds with D309 and E381, we observed R668 form π-π interactions with R311 (Fig. 5C). Arginine-mediated π-interactions have been well characterized in other systems for their ability to stabilize interaction interfaces and facilitate conformational changes (34, 35).

The IDE-N/C interactions we observe in our simulations are not restricted to discrete cognate pairs; rather, we observed that residues within the network periodically swap among several interaction partners. This “charge-swapping” phenomenon allows IDE-N to rotate or slide against IDE-C to adopt multiple conformations and maintain favorable interacting contacts (Fig. 5D). While multiple, concurrent, instances of charge-swapping were associated with rapid changes in the D1-D2-D3-D4 dihedral angle, singular events of charge-swapping did not appear to be significant enough to stimulate large-scale conformational change (Fig. 5E). Instead, substantial conformational changes appear to be caused by multiple charge-swapping events occurring in conjunction. This may explain why we observed no noticeable change in the conformational dynamics of the pO state subunits when comparing the WT and R668A constructs. Consurf analysis reveals that most of the residues forming these interaction networks are highly conserved among IDE homologs (Fig. S13) (36). This indicates that the revealed rotational motion is likely evolutionarily conserved and offers a potential mechanism by which IDE unfolds and repositions bound peptide substrates to degrade amyloid peptides.

Discussion

At its core, cryoEM is an ensemble method of structure determination, averaging tens of thousands of particles together to generate consensus structures. While methodologies have been well-established to separate out discrete conformational classes, proteins that exhibit constant gradients of structural heterogeneity remain problematic, although significant attention has been devoted to the issue in recent years, with varying levels of success (28, 29, 37). Currently, the best approaches use various dimensionality reduction techniques to analyze the principal components of particle heterogeneity. While a substantial innovation, it remains to be seen how the principal components of structural heterogeneity, derived from a series of static snapshots of a particle at the time of freezing, correlate with the real molecular motions of a protein in an aqueous environment. For many systems, especially enzymes of clinical significance, proper characterization of the conformational dynamics and transitional states of the protein are necessary for a full understanding of protein function and lays the foundation for future therapeutic development. MD simulations can provide this information, yet such studies are often met with skepticism by experimental researchers. By integrating MD simulations with experimental techniques, the weaknesses of one technique can be compensated for by the strengths of other techniques to provide comprehensive information about the system of interest. While this integrative approach has proven beneficial across a wide range of systems, including proteorhodopsin, the ribosome, and rotary ATPases, perhaps the best example of this approach is the CRISPR-Cas9 system, where an exhaustive number of studies have combined cryoEM, crystallography, and traditional biophysical techniques with traditional and accelerated MD methods to drive the field forward at an astounding pace (38–43). Here, we contribute to the growing body of literature fusing computational and experimental techniques. We used multibody analysis to reveal the dominant components of structural heterogeneity within our cryoEM particle population, which suggest the conformational dynamics of IDE include a significant rotational component that was not evident from analysis of the ensemble structures. All-atom MD simulations supported the conformational dynamics predicted from cryoEM and offered insight into how promiscuous interactions with a dynamic hydrogen bonding network at the IDE-N/C interface permit the domains to move against one another and suggested that R668 played a key role in mediating IDE-N/C interactions. MD simulations of an R668A mutant revealed significantly altered open-closed dynamics relative to WT. These results were further supported by experimental enzymatic kinetic assays and SEC-SAXS experiments.

IDE utilizes both unfoldase and protease activities to degrade clinically-relevant peptides, including three glucose-regulating hormones with contrary effects – insulin, amylin, and glucagon – and Aβ, the accumulation of which is associated with the progression of Alzheimer's disease. This has made IDE an attractive target for therapeutic intervention (9, 18). One major challenge is how to better control selectivity when the substrates IDE degrades are highly diverse in sequence and structure. It remains unknown how IDE unfolds substrates prior to cleavage. Our studies characterize the conformational dynamics of IDE and offer insight into the catalytic mechanisms, revealing that IDE-N rotates against IDE-C. We envision that this motion works in concert with the open-close transition to reposition and unfold substrates prior to cleavage. This rotation appears to be largely unconstrained in the open state and may play a role in promoting substrate capture. In the closed state, we found the rotational motion to be mediated by charge-

swapping events at the IDE-N/C interface. These events are supplemented by interactions between hydrophobic patches that provide a non-specific interaction interface and allow the domains to slide against one another. We found that IDE-N/C interactions were mediated by the key residue R668, and mutation of this residue altered the enzymatic activity and conformation of IDE, likely by impairing the open-close transition. Our data suggests that, while the R668A mutation globally impaired IDE activity, R668A may be more significantly impaired in its degradation of larger, folded substrates compared to shorter, disordered substrates due to its additional effect on Ki value for insulin. We envision that IDE conformational dynamics can regulate solvent access to the catalytic chamber and substrate binding, unfolding, and cleavage are all intimately linked.

Our results allow us to put forth a refined model describing how IDE recognizes amyloid peptides that are diverse in size and shape (Fig. 5F). The catalytic cycle of IDE starts with at least one subunit of dimeric IDE adopting the open state, exposing the interior of the catalytic chamber to potential substrates (21). The unconstrained motion of IDE-N likely facilitates substrate interaction in a variety of initial orientations. IDE-N and IDE-C have negatively and positively charged surfaces, respectively, and are able to attract peptide substrates with complementary charge profiles (21, 22). Such peptides have a high dipole moment and are often prone to aggregation. However, in the open state, the door subdomain, which contains the key catalytic residues, is highly flexible, rendering the open state catalytically incompetent (21, 24). Proper positioning and unfolding of substrate is necessary to stabilize the cleft prior to catalysis. Substrate binding has been suggested to enhance IDE closing through charge complementarity (44). It is likely that the rotation of IDE-N permits a range of closing pathways to accommodate non-optimal substrate binding geometry (Fig. 3). If the closing geometry permits R668 to interact with E381 or D309, the closing reaction continues to completion. If the closing geometry does not permit the R668-mediated interactions, this could be a signal of improper substrate binding, allowing IDE-N to open and either release improper substrate or reposition for another closing attempt. Once closed, IDE-N is capable of rotating against IDE-C mediated by charge-swapping event at the IDE-N/C interface (Fig. 5). The combined motions associated with the open to closed transition and rotation between IDE-N and IDE-C can lead to the distortion of the secondary structure of bound substrate and facilitate unfolding and proper positioning of substrate for cleavage. Our data indicates that IDE is able to accommodate any possible orientation of insulin tethering within the catalytic site (Fig. 1), suggesting that IDE is able to cut both chains of insulin in rapid succession without requiring significant repositioning (25). IDE then opens to release cleavage products and the cycle begins anew.

Our integrative structural analyses provide the structural basis for IDE conformational dynamics and their roles in IDE catalytic cycle and the framework for future studies. We speculate that multi-cycles of open-close transition and rotation driven by IDE conformational dynamics will be required for its unfoldase activity. However, the kinetic timescale of this reaction is fast, estimated to be ~500 ms, rendering investigation challenging (25). Furthermore, the mechanism by which IDE opens remains unknown. Future

integrative structural analysis for the better understanding and control of IDE catalytic cycle will provide a means to modulate IDE activity in the substrate specific manner to realize its therapeutic potential.

Materials and Methods

Expression and purification of human IDE:

Cysteine-free IDE (referred to as IDE in this study) was expressed in *E. coli* BL21 (DE3) cells (at 25 °C for 20 h, 0.5mM IPTG induction using T7 medium). The R668A mutation was introduced into this construct by site-directed mutagenesis with the following primer pair: CCGGAAATTGTTAAGAGATGCCATATATGCTTCTTTG (forward) CAAAGAAGCATATATGGCATCTCTTAACAACCGGGC (reverse) and verified by sequencing. Recombinant IDE proteins were purified by Ni-NTA, source-Q, and Superdex 200 columns as previously described (21). Protein was aliquoted and stored at -80^{o} C.

CryoEM data collection and analysis:

Thawed IDE aliquots were further purified by Superdex 200 chromatography using buffer containing 20 mM HEPES, pH 7.2, 150 mM NaCl, 10mM EDTA and then mixed with $Fab_{H11\text{-}E}$ at an equal molar ratio. $Fab_{H11\text{-}E}$-IDE complex was purified by Superdex 200 chromatography in the presence of insulin at molar ratio IDE:insulin=2:1. Insulin was purchased from SIGMA (91077C). All grids, 300 mesh carbon or gold holey nanowire grids were plasma cleaned with O_2 and H_2 for 10 secs using a Solarus plasma cleaner (Gatan) and plunged at 133 milli-second using Chameleon, Spotiton-based technology. All images were acquired using a Titan Krios microscope (FEI) operated at 300KeV with a Gatan K3 direct electron detector (Gatan) in counting mode. Images were automatically acquired using Leginon (45). Images were processed using software integrated into RELION (46). Frames were aligned using MotionCor2 (47) software with dose weighting, CTF was estimated using Gctf (48), particles were picked and extracted automatically using RELION. Particle stacks were processed through several rounds of 2D and 3D classification. Selected classes were then processed for high-resolution 3D refinement. The flowchart and detailed data processing is summarized in Figure S1. Finally, the overall map was improved by particle polishing in RELION and sharpening. The final resolution was estimated using Fourier Shell Correlation (FSC=0.143, gold-standard). The density fitting and structure refinement was done using UCSF CHIMERA (49), COOT (50), REFMAC5 (51) and PHENIX (52).

Multibody analysis:

We defined three rigid bodies as shown in figure 2A for the analysis. The size of the user-defined bodies has been theorized to play a significant role in the success of multibody refinement, most likely due to the fact that larger bodies will have a stronger signal-to-noise ratio (28). With a Fab bound to each of IDE-N within IDE dimer, we assessed the impact of body size on multibody refinement by examining how density corresponding to 2 F_v regions, 1 F_v region, or no F_v regions would affect the multibody refinement. Fab has a variable region (F_v), and a constant region (F_c) and we chose to subtract and mask out the

density corresponding to the Fab F_c, as we speculated that any motion between the Fab F_v and F_c domains would bias the results of the multibody refinement. The quality of the multibody results was assessed based on the subsequent improvement in the Coulomb potential map quality and calculated resolution. In the absence of any F_v density, the maps resulting from multibody refinement had the best calculated resolution, yet the map quality was quite poor overall; much of the density appears globular and featureless, particularly in the IDE-N regions (Fig. S14). Conversely, when F_v density was present on both IDE-N bodies, the resolution of the multibody output maps (4.5 Å) was worse than the resolution of the input map (4.3 Å). Thus, we found that the greatest improvement occurred when the F_v density was present on only the exterior body of the pO subunit. In this case, multibody refinement improved both the calculated resolution and density quality.

Following multibody refinement, the structural heterogeneity for each body was analyzed along six principal component vectors, resulting in 18 vectors describing the primary components of structural variation within the IDE dimer. We analyzed the variation along these vectors as a proxy for protein motion. The results of the multibody analysis indicate the degree of structural variation that is explained by motion along each of the primary component vectors. For all IDE states, we observed no single component vector could explain the majority of structural variance (Table S1-6). Given the underlying assumptions of multibody refinement, there is a question of whether any observed motions are biologically relevant or simply the result of improper particle alignment. We reasoned that the greater the degree of variation explained by motion along a specified component vector, the greater the likelihood was that the motion would be biologically relevant. Therefore, we focus on the top 9 component vectors, which cover ~75% of total variance for each state, in greater detail.

Molecular dynamics simulations:

All-atom MD simulations were prepared from the O/pO structure of IDE (PDB: 7RZK) and necessary CHARMM force field PARAM36 files using QwikMD (53). Gaps in the structure were modeled in using the AlphaFold model of IDE. A three-step minimization-annealing-equilibration process was used to generate an equilibrated system. Simulations were performed under NPT (constant numbers of particles N, pressure P, and temperature T) conditions at 310 K and 1 atm with periodic boundary conditions in NAMD3.0 (54). Explicit solvent was described with the TIP3P model (55). Custom TCL scripts in VMD were used to calculate the D1-D4 centers-of-mass distance and D1-D2-D3-D4 centers-of-mass dihedral angle. Residues 964-988 in D4 that are mostly absence in all static structures of IDE were built using Alpha-fold, which is highly flexible in our MD simulations. They were omitted from the analysis of MD simulations, as the inclusion of this flexible loop significantly altered the center-of-mass independent of global conformational change.

Enzymatic assays:

IDE constructs were exchanged into activity buffer comprised of 25 mM Tris-HCl pH 7.5, 150 mM NaCl, and 10 mM $ZnCl_2$. MCA-RPPGFSAFK(Dnp) was purchased from Enzo Life Sciences

(BMLP2270001), resuspended in dimethyl sulfoxide at a concentration of 5 mM and diluted to 5 μM with activity buffer. For activity measurements, 5 μM MCA-RPPGFSAFK(Dnp) was mixed with the desired IDE construct at concentrations ranging from 1-100 nM in the presence or absence of 1 mM ATP in a 200 μl reaction. Fluorescence was monitored every 30 sec for 30 minutes with excitation/emission wavelengths of 320/405 nm at 37° C. Initial velocity was calculated during the linear range. Insulin was purchased from MP Biomedicals (#0219390010), resuspended in 0.01N HCl at 1 mM and diluted to desired concentration with activity buffer. For competition assays, 5 nM IDE construct and 5 μM MCA-RPPGFSAFK(Dnp) was mixed with insulin ranging from 0-100 μM. Fluorescence was monitored every 30 sec for 30 minutes with excitation/emission wavelengths of 320/405 nm at 37° C. Initial velocity was calculated during the linear range for each concentration of insulin and normalized relative to the respective construct's velocity in the absence of insulin to maintain consistency across plates, yielding values for relative inhibition which were plotted vs [insulin] and fit to the Michaelis-Menten equation in xmgrace to generate apparent K_i values. All experiments were performed in triplicate.

Size-exclusion chromatography coupled small-angle X-ray scattering

SAXS/WAXS data collection was employed the Life Sciences X-ray Scattering (LIX) Beamline at the National Synchrotron Light Source II (NSLS II) at Brookhaven National Laboratory in Upton, NY. Briefly, 60ul of samples in solution were pipetted into PCR tubes, placed into a Bio-Inert Agilent 1260 Infinity II HPLC multisampler and measured using the isocratic SEC-SAXS format at the beamline (56). 50ul of sample was injected into the Phenomenex Biozen dSEC2 3 μm bead size with 200Å pore size column was utilized at a flow rate of 0.35mL/min for 25 minutes. SAXS and WAXS data are collected simultaneously on a Pilatus 1M (SAXS) and Pilatus 900K (WAXS) detectors with a 2 second exposure (57). Data from both detectors is then scaled and merged. Intensity is normalized using the water peak height at 2.0 $Å^{-1}$. Data processing and analysis was performed using py4xs and lixtools in jupyter lab. Buffer frames 100-125 were averaged and used for subtraction of averaged frames under the peak of interest.

Data Availability

CryoEM maps and refined models have been deposited to the EMDB and PDB, respectively with the following accession numbers:

O/O state: EMDB-24760, PDB 7RZH

O/pO state: EMDB-24759, PDB 7RZG

pO/pC state: EMDB-24757, PDB 7RZE

O/pC state: EMDB-24758, PDB 7RZF

pC/pC state: EMDB-24761, PDB 7RZI

References

1. M. S. Hipp, P. Kasturi, F. U. Hartl, The proteostasis network and its decline in ageing. *Nat Rev Mol Cell Biol* **20**, 421–435 (2019).

2. C. López-Otín, M. A. Blasco, L. Partridge, M. Serrano, G. Kroemer, The Hallmarks of Aging. *Cell* **153**, 1194–1217 (2013).

3. F. Chiti, C. M. Dobson, Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. *Annual Review of Biochemistry* **86**, 27–68 (2017).

4. F. Chiti, C. M. Dobson, Protein Misfolding, Functional Amyloid, and Human Disease. *Annual Review of Biochemistry* **75**, 333–366 (2006).

5. J. Greenwald, R. Riek, Biology of Amyloid: Structure, Function, and Regulation. *Structure* **18**, 1244–1260 (2010).

6. 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).

7. T. Saido, M. A. Leissring, Proteolytic Degradation of Amyloid β-Protein. *Cold Spring Harb Perspect Med* **2**, a006379 (2012).

8. W. G. Liang, *et al.*, Structural basis for the mechanisms of human presequence protease conformational switch and substrate recognition. *Nat Commun* **13**, 1833 (2022).

9. W.-J. Tang, Targeting Insulin-Degrading Enzyme to Treat Type 2 Diabetes Mellitus. *Trends in Endocrinology & Metabolism* **27**, 24–34 (2016).

10. L. Bertram, Evidence for Genetic Linkage of Alzheimer's Disease to Chromosome 10q. *Science* **290**, 2302–2303 (2000).

11. L. Bertram, M. B. McQueen, K. Mullin, D. Blacker, R. E. Tanzi, Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat Genet* **39**, 17–23 (2007).

12. W. Farris, *et al.*, Partial Loss-of-Function Mutations in Insulin-Degrading Enzyme that Induce Diabetes also Impair Degradation of Amyloid β-Protein. *The American Journal of Pathology* **164**, 1425–1434 (2004).

13. W. Farris, *et al.*, Insulin-degrading enzyme regulates the levels of insulin, amyloid β-protein, and the β-amyloid precursor protein intracellular domain in vivo. *PNAS* **100**, 4162–4167 (2003).

14. M. Kim, *et al.*, Decreased Catalytic Activity of the Insulin-degrading Enzyme in Chromosome 10-Linked Alzheimer Disease Families*. *Journal of Biological Chemistry* **282**, 7825–7832 (2007).

15. R. Sladek, *et al.*, A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* **445**, 881–885 (2007).

16. R. E. Tanzi, L. Bertram, Twenty Years of the Alzheimer's Disease Amyloid Hypothesis: A Genetic Perspective. *Cell* **120**, 545–555 (2005).

17. E. Zeggini, *et al.*, Replication of Genome-Wide Association Signals in UK Samples Reveals Risk Loci for Type 2 Diabetes. *Science* **316**, 1336–1341 (2007).

18. J. P. Maianti, *et al.*, Substrate-selective inhibitors that reprogram the activity of insulin-degrading enzyme. *Nature Chemical Biology* **15**, 565–574 (2019).

19. J. P. Maianti, *et al.*, Anti-diabetic activity of insulin-degrading enzyme inhibitors mediated by multiple hormones. *Nature* **511**, 94–98 (2014).

20. M. A. Leissring, *et al.*, Enhanced Proteolysis of β-Amyloid in APP Transgenic Mice Prevents Plaque Formation, Secondary Pathology, and Premature Death. *Neuron* **40**, 1087–1093 (2003).

21. Z. Zhang, *et al.*, Ensemble cryoEM elucidates the mechanism of insulin capture and degradation by human insulin degrading enzyme. *eLife* **7**, e33572 (2018).

22. Y. Shen, A. Joachimiak, M. Rich Rosner, W.-J. Tang, Structures of human insulin-degrading enzyme reveal a new substrate recognition mechanism. *Nature* **443**, 870–874 (2006).

23. P. Li, W.-L. Kuo, M. Yousef, M. R. Rosner, W.-J. Tang, The C-terminal domain of human insulin degrading enzyme is required for dimerization and substrate recognition. *Biochemical and Biophysical Research Communications* **343**, 1032–1037 (2006).

24. L. A. McCord, *et al.*, Conformational states and recognition of amyloidogenic peptides of human insulin-degrading enzyme. *PNAS* **110**, 13827–13832 (2013).

25. M. Manolopoulou, Q. Guo, E. Malito, A. B. Schilling, W.-J. Tang, Molecular Basis of Catalytic Chamber-assisted Unfolding and Cleavage of Human Insulin by Human Insulin-degrading Enzyme*,. *Journal of Biological Chemistry* **284**, 14177–14188 (2009).

26. E. N. Baker, *et al.*, The structure of 2Zn pig insulin crystals at 1.5 Å resolution. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* **319**, 369–456 (1997).

27. L. Busto-Moner, C.-J. Feng, A. Antoszewski, A. Tokmakoff, A. R. Dinner, Structural Ensemble of the Insulin Monomer. *Biochemistry* **60**, 3125–3136 (2021).

28. T. Nakane, D. Kimanius, E. Lindahl, S. H. Scheres, Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. *eLife* **7**, e36861 (2018).

29. A. Punjani, D. J. Fleet, 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *Journal of Structural Biology* **213**, 107702 (2021).

30. I. Bahar, C. Chennubhotla, D. Tobi, Intrinsic Enzyme Dynamics in the Unbound State and Relation to Allosteric Regulation. *Curr Opin Struct Biol* **17**, 633–640 (2007).

31. I. Bahar, A. Rader, Coarse-grained normal mode analysis in structural biology. *Current Opinion in Structural Biology* **15**, 586–592 (2005).

32. J. V. King, *et al.*, Molecular Basis of Substrate Recognition and Degradation by Human Presequence Protease. *Structure* **22**, 996–1007 (2014).

33. H. Im, *et al.*, Structure of Substrate-free Human Insulin-degrading Enzyme (IDE) and Biophysical Analysis of ATP-induced Conformational Switch of IDE *. *Journal of Biological Chemistry* **282**, 25453–25463 (2007).

34. R. M. Vernon, *et al.*, Pi-Pi contacts are an overlooked protein feature relevant to phase separation. *eLife* **7**, e31486 (2018).

35. C. T. Armstrong, P. E. Mason, J. L. R. Anderson, C. E. Dempsey, Arginine side chain interactions and the role of arginine as a gating charge carrier in voltage sensitive ion channels. *Sci Rep* **6**, 21759 (2016).

36. A. Ben Chorin, *et al.*, ConSurf-DB: An accessible repository for the evolutionary conservation patterns of the majority of PDB proteins. *Protein Science* **29**, 258–267 (2020).

37. E. D. Zhong, T. Bepler, B. Berger, J. H. Davis, CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. *Nat Methods* **18**, 176–185 (2021).

38. A. Saha, P. R. Arantes, G. Palermo, Dynamics and mechanisms of CRISPR-Cas9 through the lens of computational methods. *Current Opinion in Structural Biology* **75**, 102400 (2022).

39. G. Palermo, Y. Miao, R. C. Walker, M. Jinek, J. A. McCammon, Striking Plasticity of CRISPR-Cas9 and Key Role of Non-target DNA, as Revealed by Molecular Simulations. *ACS Cent. Sci.* **2**, 756–763 (2016).

40. F. Jiang, J. A. Doudna, CRISPR–Cas9 Structures and Mechanisms. *Annual Review of Biophysics* **46**, 505–529 (2017).

41. S. Hirschi, D. Kalbermatter, Z. Ucurum, T. Lemmin, D. Fotiadis, Cryo-EM structure and dynamics of the green-light absorbing proteorhodopsin. *Nat Commun* **12**, 4107 (2021).

42. S. Arenz, *et al.*, A combined cryo-EM and molecular dynamics approach reveals the mechanism of ErmBL-mediated translation arrest. *Nat Commun* **7**, 12026 (2016).

43. S.-H. Roh, *et al.*, Cryo-EM and MD infer water-mediated proton transport and autoinhibition mechanisms of Vo complex. *Science Advances* **6**, eabb9605 (2020).

44. R. E. Hulse, L. A. Ralat, W.-J. Tang, Structure, Function, and Regulation of Insulin-Degrading Enzyme. *Vitam Horm* **80**, 635–648 (2009).

45. C. Suloway, *et al.*, Automated molecular microscopy: The new Leginon system. *Journal of Structural Biology* **151**, 41–60 (2005).

46. J. Zivanov, *et al.*, New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).

47. S. Q. Zheng, *et al.*, MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331–332 (2017).

48. K. Zhang, Gctf: Real-time CTF determination and correction. *Journal of Structural Biology* **193**, 1–12 (2016).

49. E. F. Pettersen, *et al.*, UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612 (2004).

50. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta Cryst D* **60**, 2126–2132 (2004).

51. G. N. Murshudov, *et al.*, REFMAC5 for the refinement of macromolecular crystal structures. *Acta Cryst D* **67**, 355–367 (2011).

52. P. D. Adams, *et al.*, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst D* **66**, 213–221 (2010).

53. J. V. Ribeiro, *et al.*, QwikMD — Integrative Molecular Dynamics Toolkit for Novices and Experts. *Sci Rep* **6**, 26536 (2016).

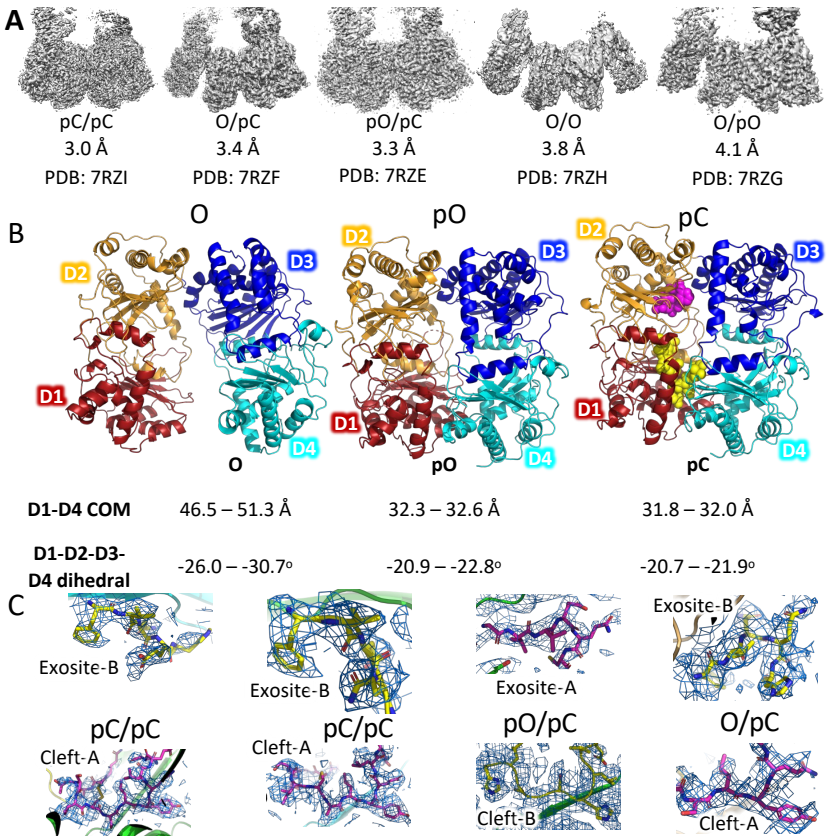
54. J. C. Phillips, *et al.*, Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J Chem Phys* **153**, 044130 (2020).

55. 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).

56. L. Yang, *et al.*, Tools for supporting solution scattering during the COVID-19 pandemic. *J Synchrotron Rad* **28**, 1237–1244 (2021).

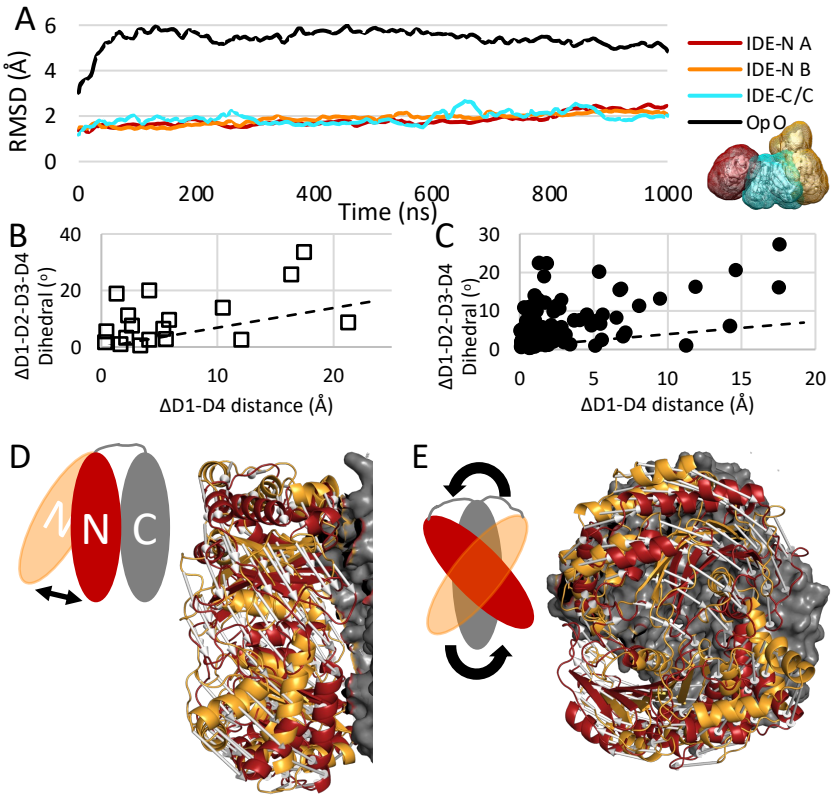
57. L. Yang, *et al.*, Solution scattering at the Life Science X-ray Scattering (LiX) beamline. *J Synchrotron Rad* **27**, 804–812 (2020).

575 **Figures**



576

577 **Figure 1: CryoEM structures. (A)** Overview of the cryoEM structures. See figure S1 for processing details.
578 **(B)** Comparison of the open (O), partial open (pO), and partial closed (pC) subunit states present in our
579 cryoEM structures with domain organization. The distance between the D1 and D4 domain centers-of-mass
580 (D1-D4 COM) along with the dihedral angle formed by the D1-D2-D3-D4 domain centers-of-mass (D1-D2-
581 D3-D4 dihedral) described in Zhang et al. (21) and depicted in Fig. S4 were used as biologically important
582 criteria to quantify observed conformations. **(C)** Insulin density and corresponding model in our cryoEM
583 structures. Both the A chain (magenta) and B chain (yellow) can fit the density in the exosite and catalytic
584 cleft.



585

586 **Figure 2: Conformational dynamics of IDE implied by structural heterogeneity. (A)** All-atom MD
587 simulations analysis. The primary source of structural variance (RMSD) results from the IDE-N moving
588 against IDE-C as a rigid body. Rigid bodies were defined as colored for multibody refinement in RELION.
589 **(B-C)** Multibody analysis. The range of conformational variance described by the top principal component
590 vectors displays an unexpectedly high degree of rotational motion, as measured by the change in D1-D2-
591 D3-D4 dihedral angle across each vector's gradient of structural heterogeneity, in both the absence **(B)** and
592 presence **(C)** of insulin compared to the expected open-close transition pathway predicted from a linear
593 interpolation of the experimentally determined structures of IDE (dashed line, Fig. S4). Two dominant
594 components of structural variance are revealed from multibody analysis: **(D)** where IDE-N swings relative
595 to IDE-N about the inter-domain linker, and **(E)** where IDE-N rotates against IDE-C. Starting (orange) and
596 ending (red) states of IDE-N shown with pathway depicted by arrows. IDE-C shown as gray surface.

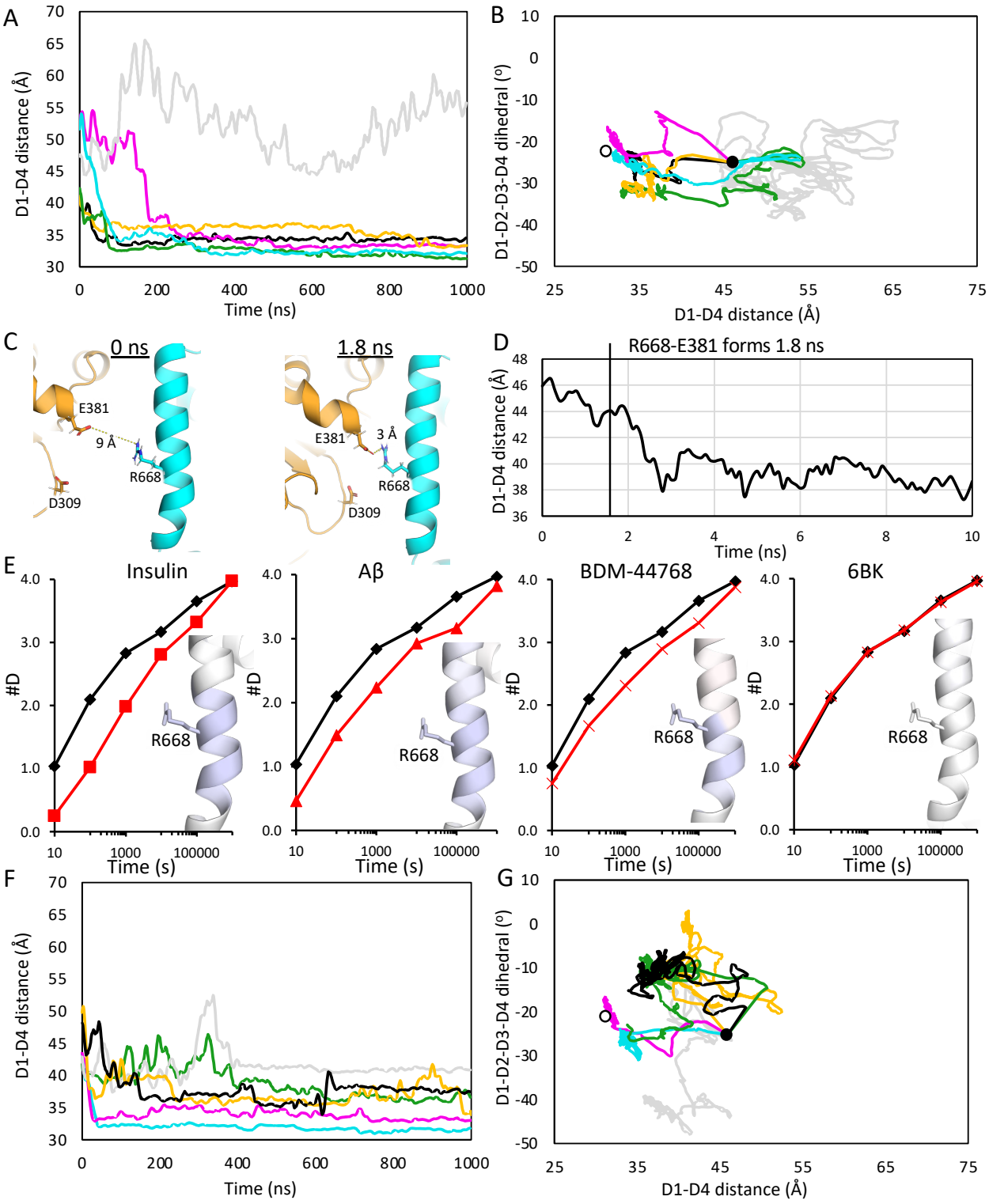
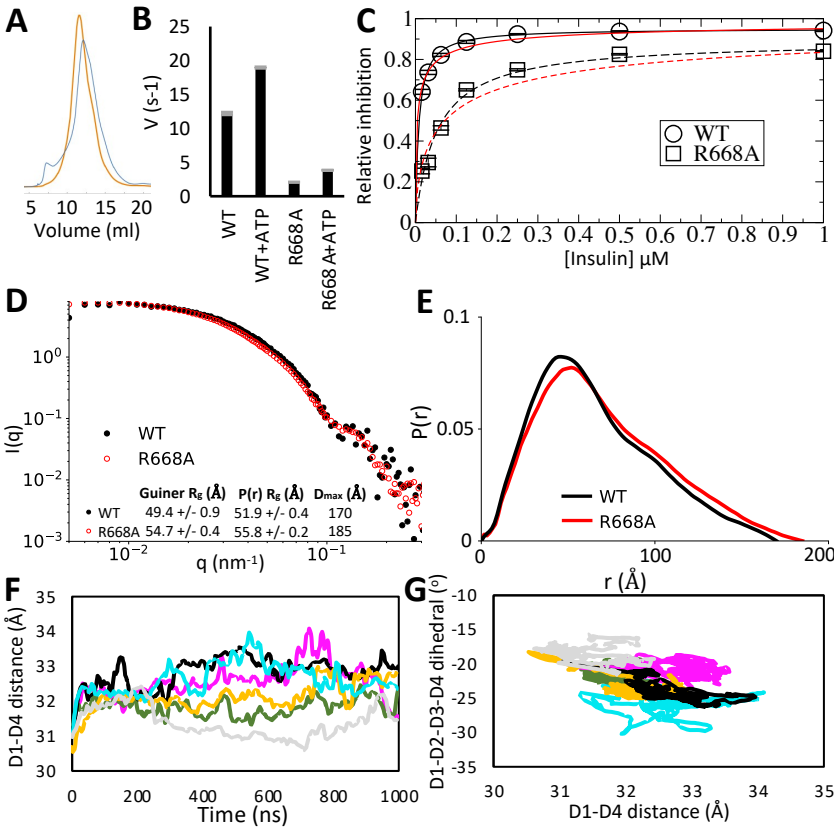


Figure 3: All-atom MD reveals a molecular basis for IDE conformational dynamics. **(A)** Measurements of the O subunit D1-D4 distance over the course of six separate microsecond long all-atom MD simulations of WT IDE. Of which, the open subunit closed in 5 of the 6 simulations. **(B)** Plot of the O subunit D1-D4 distance vs the D1-D2-D3-D4 COM dihedral angle over the course of the simulation of WT IDE. The open

602 subunits displayed a variety of closing pathways and did not close to a consensus structure. Starting
603 structure shown as black dot, pO structure shown as white dot. **(C)** R668 acts as a guidepost residue,
604 rapidly interacting with D309 or E381. Formation of this interaction is associated with rapid closing, as
605 measured by a decrease in D1-D4 distance **(D). (E)** Hydrogen-deuterium exchange mass spectrometry
606 highlights the importance of R668 in mediating the open-close transition. In the presence of insulin (panel
607 1, red), Aβ (panel 2, red) , and BDM-44768 (panel 3, red), all of which promote IDE closing, the peptide
608 containing R668 shows reduced deuterium exchange relative to apo-IDE (black), yet in the presence of 6bk
609 (panel 4, red), which does not promote closing, there is no difference in the exchange rates for the R668
610 containing peptide relative to apo-IDE (black). Helix containing R668 colored by red (increase) – white (no
611 change) – blue (decrease) gradient depicting the degree of deuterium exchange relative to apo-IDE. **(F)**
612 Measurements of the O subunit D1-D4 distance over the course of six separate microsecond long all-atom
613 MD simulations of IDE R668A. **(G)** Plot of the O subunit D1-D4 distance vs the D1-D2-D3-D4 COM dihedral
614 angle over the course of the simulation of IDE R668A. The six separate microsecond long simulations
615 indicate that an R668A mutation significantly alters the closing dynamics of IDE **(F)** and increases the
616 rotational motion **(G)** relative to WT (panels **A** and **B** respectively). Starting structure shown as black dot,
617 pO structure shown as white dot.



618

619 **Figure 4 R668A alters IDE activity *in vitro* (A)** Elution profile of WT IDE (blue) compared to the R668A
620 mutant (orange) from a S200 SEC column. (**B**) Degradation of the fluorescent substrate MCA-
621 RPPGFSAFK(Dnp) by WT IDE and the R668A construct in the presence and absence of ATP. Data

represents the average initial velocities of three replicates performed at a protein concentration of 3.125 nM. Error bars (gray) represent the standard error. **(C)** Inhibition of MCA-RPPGFSAFK(Dnp) degradation by WT IDE (circles, solid fit lines) and IDE R668A (squares, dashed fit lines) in the presence of varying amounts of insulin. Data was fit to the Michaelis-Menten (black) and Hill equations (red). Relevant parameters, Michaelis-Menten: WT: χ^2=0.001, V_{max}=0.951, K_i=8.3 nM; R668A: χ^2=0.005, V_{max}=0.892, K_i=52 nM; Hill: WT: χ^2=0.009, n=0.55, K_i=51 nM; R668A: χ^2=0.055, n=0.61, K_i=198 nM. Error bars represent standard error, data points represent the average of three replicates. **(D)** SEC-SAXS profile of WT (black) and R668A (red) constructs with Rg values calculated by both the Guinier and Porod methods along with Dmax derived from the P(r) function **(E)**. **(F)** Measurements of the pO subunit D1-D4 distance over the course of six separate microsecond long all-atom MD simulations of WT IDE. **(G)** Plot of the pO D1-D4 distance vs the D1-D2-D3-D4 COM dihedral angle over the course of the simulation of WT IDE.

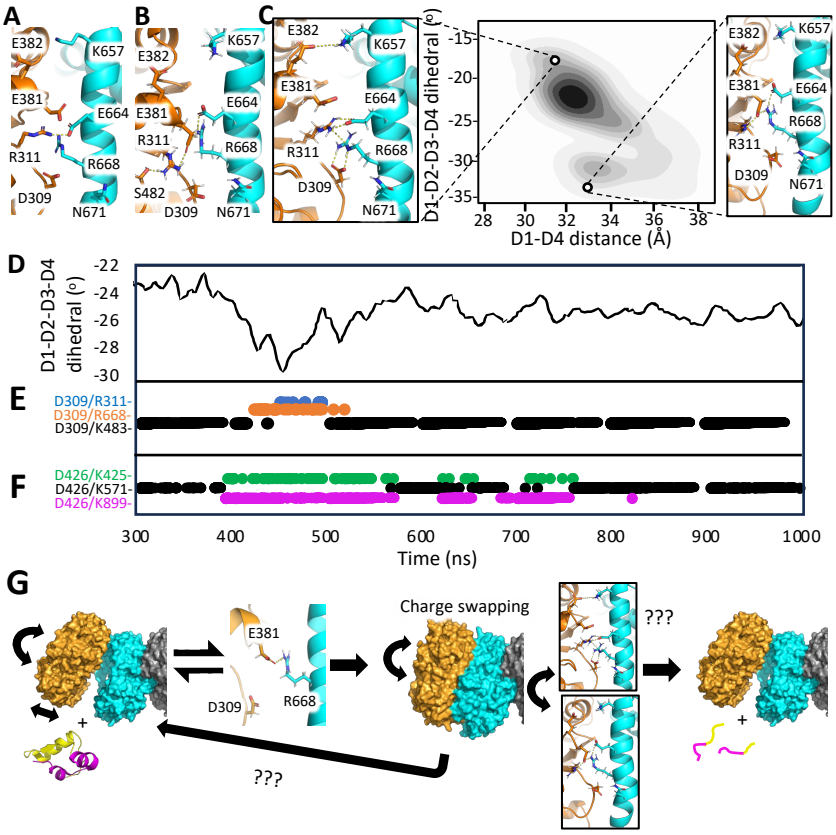


Figure 5 Structural basis of closed state conformational dynamics. (A) IDE-N/C interface previously solved crystal structures (PDB:2G47 shown) shows side chains are ill-positioned for interaction. **(B)** IDE-N/C interface formed upon open subunit closing in our MD simulations reveals a complex hydrogen bonding network. **(C)** Heat map showing conformational geometries that were preferentially sampled in our MD simulations by the open subunits upon closing. Insets highlight how the IDE-N/C interface changes to permit interdomain motion. **(D)** Plot of the O subunit D1-D2-D3-D4 dihedral angle during a subset of a single WT IDE MD simulation after the open-close transition has been completed. Charge-swapping between residues at the IDE-N/C interface is associated with changes in the D1-D2-D3-D4 dihedral. **(E)** For most of the

simulation, D309 interacts with K483 (black), however, this interaction is broken for ~100 ns, during which D309 instead interacts with R311 (blue) and R668 (orange). **(F)** For most of the simulation, D426 interacts with K571 (black), yet this interaction is periodically broken, and D426 instead interacts with K425 (green) and K899 (magenta). When these events of charge-swapping coincide with D309 charge-swapping **(E)**, they are associated with a large change in the D1-D2-D3-D4 dihedral angle **(D)**. When they occur alone, the effect on D1-D2-D3-D4 dihedral is smaller. **(G)** Model for the catalytic cycle of IDE. IDE-N colored orange, IDE-C colored cyan. Insulin colored by chain (A: magenta, B: yellow). Single subunit of the dimer shown for simplicity, second subunit colored gray.

Table 1: CryoEM data collection, refinement, and validation statistics

Data collection and processing					
Microscope	Titan Krios				
Camera	Gatan K3				
Automation software	Leginon				
Magnification	81,000				
Voltage (kV)	300				
Frames collected per micrograph	50				
Dose per frame (e-/Å2)	1.36				
Total electron dose (e-/Å2)	67.9				
Defocus range (μM)	0.7 to 1.5				
Total micrographs	7,611				
Initial particle images (no.)	7,206,464				
	O/O state	**O/pO state**	**pO/pC state**	**O/pC state**	**pC/pC state**
Pixel size (Å)	1.0842	1.0842	1.0842	1.0842	1.0842
Final particle images (no.)	77,973	328,870	76,379	304,011	1,341,061
Symmetry imposed	C1	C1	C1	C1	C1
Map resolution (Å)	3.8	4.1	3.3	3.4	3.0
FSC threshold	0.143	0.143	0.143	0.143	0.143
EMDB	EMD-24760	EMD-24759	EMD-24757	EMD-24758	EMD-24761
Refinement					
Model resolution					
FSC 0.5	7.4 (7.7)[a]	4.5 (7.4)[a]	3.5 (3.8)[a]	3.6 (3.9)[a]	3.1 (3.2)[a]
FSC 0.143	3.6 (3.8)[a]	3.4 (3.8)[a]	3.1 (3.2)[a]	3.3 (3.3)[a]	3.0 (3.0)[a]
Sharpening B factor	-73.0	-70.0	-87.1	-47.7	-109.8
Refinement package	PHENIX & COOT	PHENIX & COOT	PHENIX & COOT	PHENIX & COOT	PHENIX & COOT
Model composition					

Protein residues	1867	1888	1898	1902	1926
Total atoms	15300	15465	15545	15584	15775
B factors					
Protein	64.38	129.58	70.19	80.87	26.93
RMS deviations					
Bond length	0.006	0.006	0.005	0.005	0.005
Bond angle	1.148	1.126	1.031	1.062	0.985
Ramachandran (%)					
Favored	96.99	94.36	96.34	95.92	97.07
Allowed	3.01	5.64	3.66	4.08	2.93
Outliers	0	0	0	0	0
Validation					
MolProbity score	1.51	1.67	1.34	1.49	1.3
Poor rotamers (%)	0.24	0.06	0.18	0.12	0.23
Clash score	6.07	5.26	3.03	4.25	3.44
Correlation coefficient	0.68	0.68	0.77	0.74	0.79
Cbeta outliers	0.11	0	0	0.06	0.05
CaBLAM outliers	1.89	2.73	1.76	1.86	1.9
EMRinger score	0.51	1.27	2.52	2.28	2.92
PDB ID	7RZH	7RZG	7RZE	7RZF	7RZI
[a]Unmasked resolution is given in parentheses					

650

651