Novel Druggable Sites of Insulin-Degrading Enzyme Identified through Applied Structural Bioinformatics Analysis

Suryani Lukman
Khalifa University, Abu Dhabi, United Arab Emirates
Suryani.lukman@kustar.ac.ae

Abstract
Insulin-degrading enzyme (IDE) plays critical roles in proteolysis of diverse substrates, like insulin and amyloid β. Pathologically, IDE is implicated in type 2 diabetes mellitus and Alzheimer’s diseases, but potent and selective regulators of IDE remain elusive. We have applied structural bioinformatics techniques to the largest ensemble of IDE structures determined hitherto, identified structural clusters associated with distinct conformational states and their respective clustroids. IDE utilizes its intrinsic large-scale structural motions to adopt multiple conformational states and perform molecular functions. The conformational space occupied by IDE structures can be shifted through mutations and inter-molecular interactions with other proteins, small molecules or substrate peptides. We observed that IDE-N is generally more dynamic than IDE-C and suggested that there are possibly other open conformational states of IDE whose structures remain unknown. We also identified novel druggable sites that are specific to particular conformational states of IDE, these sites can potentially be explored for designing investigative probes or therapeutic agents for specific spatiotemporal contexts.

1 Introduction
Insulin-degrading enzyme (IDE), also known as insulysin, is a zinc-ion-metalloprotease that belongs to the M16 metalloprotease family, comprising three subfamilies: M16A, M16B and M16C. Being a member of M16A, catabolic IDE plays important roles in protein clearance/turnovers, protein-protease networks, and proteostasis (protein homeostasis) to ensure structural and functional integrity of proteins. Depending on specific cell types, the subcellular localization of IDE varies; IDE has been mainly detected in cytoplasm [1], but is also found to be associated with cell surface [2], endosome [3], and extracellular milieu [1]. Pathologically, IDE has been identified as a susceptibility gene for diabetes and Alzheimer’s disease (Supplementary Info S1).

In terms of structures, IDE comprises N-terminal (IDE-N) and C-terminal (IDE-C) domains of ~50 kDa each. Both IDE-N and IDE-C are homologous and are linked by a short hinge loop of 26
residues. This linker is distinct among the three subfamilies of the M16 metalloprotease family. Both IDE-N and IDE-C make substantial contacts (~11,200-11,500 Å² of buried surface [4]) to form a catalytic site (also known as a catalytic chamber or a crypt, Figure S1) for entrapping, binding, and degrading peptide substrates of IDE [5]. Diverse substrates of IDE include insulin [5], glucagon [5], Aβ [5], macrophage inflammatory protein-1α (MIP-1α) [6], natriuretic peptides [7], IGF-II, TGF-α, and amylin [8]. The selective binding of these substrates to IDE is facilitated by potentially dynamic shape, volume (~13,000 Å³), and charge distribution [5] of IDE catalytic site. Nevertheless, the structural dynamics and molecular mechanisms that are critical for IDE functions, remain unclear.

Since multiple substrates bind to the catalytic site of IDE, we are particularly interested in potential chemical probes for studying IDE’s functions and therapeutic agents that selectively bind to novel druggable sites, distinct from the catalytic sites. To identify druggable sites on the IDE structures, we need to map the structures and structural dynamics of IDE. Knowledge of IDE druggable sites are prerequisites for identifying and designing IDE mutants, small molecules or stabilized peptides, that can be used to illuminate and regulate the functions of IDE in both physiological and pathological conditions, notably diabetes and Alzheimer’s diseases.

2 Materials and Methods

To investigate structural dynamics of IDE, we first searched UniProt for a high-resolution PDB entry with the highest number of residues resolved, followed by visual examination using PyMol. Examining UniProt entry P14735 with the aforementioned criteria, we decided to use the PDB entry 3E4A [9] as a starting structure to retrieve other IDE structures from PDB to form an ensemble of 190 IDE structures. Examining the alignment scores and their associated E-values, a sensible normalized score (-log(E-Value)) cutoff of 709 bits was used to prune the ensemble, resulting in 88 IDE structures. All IDE structures were resolved through X-ray crystallography; it will be interesting to compare the structural dynamics of crystal structures and those resolved in solutions [10], when NMR structures become available. To investigate the structural dynamics of IDE structures, we have performed iterative principal component analysis (PCA), root-mean square fluctuation (RMSF) analysis, normal mode analysis (NMA) and mapping of potential binding sites. We integrated the results to derive novel understandings of the IDE structures and the implications on their molecular functions.

3 Results and Discussion

3.1 IDE utilizes its multiple conformational states to perform molecular functions

To successfully probe into structure-dynamics-function relationships of studied proteins, comprehensive structural mapping of available structures have been performed using classic PCA on diverse proteins and domains, including titin immunoglobulin domain [11] and DNA binding domain of p53 [10]. In this study, we applied iterative PCA on the ensemble of IDE structures. After the first iteration, PDB entry 1Q2L from Escherichia coli was excluded from subsequent analyses because it is the single non-mammalian structure among all of the IDE structures in the ensemble and an outlier on the first principal component (PC1).

In the second iteration of PCA using the final ensemble of IDE structures comprising 87 structures derived from 42 PDB entries, we observed four prominent clusters (Figure S1) and the members of respective clusters are identical regardless of considering either the first two or three dominant principal components. Both PC1-PC2 capture 75.7% of structural variances within the IDE ensemble. PC1-PC2 separate the IDE ensemble into four distinct clusters (Supplementary Info S2).

It seems that a significant proportion of resolved IDE structures (cluster 1) are in closed conformational states. This observation correlates with the postulation that closed conformational state of IDE is the dominant state [4]. In addition, it is possible that closed conformational states are
preferentially stabilized in crystal structures, as compared to other conformational states with less extensive contacts between IDE-N and IDE-C. To hydrolyze substrates, IDE undergoes a switch between the closed and open conformational states [4]. Cluster 2 contains IDE structures that were resolved with a synthetic antibody fragment [12], we described them as having semi-closed conformational states. Cluster 3 contains IDE structures resolved with dual exosite IDE inhibitors [13]. Cluster 4 contains IDE structure 4IOF_B in an open conformational state, which is also termed as a swinging door state [12]. The door subdomain is formed by residues L170–L241 of IDE.

To identify a representative structure for each cluster, we calculated clustroid, i.e. the structure with the least divergence to the rest of the structures. Clustroid structures have been highlighted as beneficial for proposing potential therapeutic strategies [14]. The concept of clustroids has also been applied on catalytic residues, to find other residues having similar dynamical profiles to the catalytic residues, accounting for spatiotemporal subspaces [15]. For our subsequent structural analyses, we identified the following clustroids 2JBU_B, 4Q5Z_F, 4PF9_B, and 4IOF_B of clusters 1-4, respectively.

3.2 IDE has intrinsic large-scale structural motions to facilitate its functions

To perform structural dynamics analysis of amino acid residues for identifying flexible regions, we calculated the RMSF for each resolved residue of IDE across the final ensemble (Figure S2, Supplementary Info S3). Residual flexibility plays roles in protein conformational dynamics. IDE has been shown to adopt multiple conformational states: closed and open as studied through small angle X-ray scattering (SAXS) analysis [12]. The open conformational states of IDE enable substrates to freely diffuse in and out of the catalytic site. The closed conformational states of IDE were suggested to be stabilized by extensive contacts between IDE-N and IDE-C domains acting analogous to a “latch”, hence entrapping substrates inside the catalytic site and blocking other substrates from entering the catalytic site [5]. While the catalytic function of IDE requires its adoption of closed conformational state, IDE needs to undergo a conformational switch to open states to facilitate release of catalytic products and uptake of new substrates.

The conformational changes from the closed to open conformational states have been suggested to involve ATP [4] and its physiologically relevant concentrations [16]. ATP introduces negative charges that possibly reduce electrostatic attraction between IDE-N and IDE-C domains, hence promoting a conformational switch from closed to open states. Analogously, several ATP-binding proteins such as maltose transporters [17] are known to function as a switch between their closed, semi-open and open conformational states.

The open/swinging door conformational state of 4IOF_B comprises of only ~11 to 18Å opening [12], which is too small for larger substrates (e.g. insulin and Aβ) to enter the catalytic site of IDE. Thus, it can be hypothesized that other open conformational states exist for IDE. Experimental methods to test this hypothesis include small angle X-ray scattering (SAXS). A SAXS study showed that the radius of gyration (Rg) of IDE is noticeably larger than those in both closed and open/swinging door conformational states [12]. Computational methods to test this hypothesis include molecular dynamics simulations and normal mode analysis (NMA).

Since protein structural mobility predicted through NMA has been shown to be similar to that determined using time-consuming and computationally-intensive molecular dynamics simulations [18], we have performed NMA on an ensemble of IDE clustroid structures (2JBU_B, 4Q5Z_F, 4PF9_B, and 4IOF_B), in order to probe large-scale conformational changes and motions of IDE. The overall IDE motion is a superposition of its normal modes, and we are particularly interested in first few normal modes (with the lowest frequency) that correspond to large-scale motions. Based on NMA, we observed conformational changes involving more motion of IDE-N and IDE-C sub-regions, suggesting the intrinsic structural dynamics of IDE (Figure S2). This observation also aligns with the
view that conformational changes from closed to open conformational states, and vice versa, are likely to involve pivoting between IDE-N and IDE-C. In essence, low frequency normal modes of IDE describe the conformational changes between multiple conformational states of IDE clustroids. In the clustroid of closed and open/swinging door conformational states, the distance between residues E156 of IDE-N and K884 of IDE-C is 13.3 Å and 14.3 Å, respectively; we observed through NMA that this distance can increase up to 16.8 Å, implying other (more) open conformational states that can enable the entrance of larger substrates than those permissible into the open/swinging door conformational states. Our results suggest that IDE possesses global motions that are intrinsic to its structure, but the conformational landscape can be shifted by binding of substrates, activators, inhibitors, small molecules, or other proteins.

3.3 Identification of druggable sites

To map potential binding sites for small molecules or peptides or fragments of proteins that can regulate IDE, we utilized FTmap [19]. Residues surrounding IDE catalytic site (F108, E111 in wild-type IDE/Q111/N111, H112) were identified as residues forming binding sites in the clustroids of clusters 1-3, thus supporting the reliability of this method in detecting known and novel binding sites of IDE (Table S1). Interestingly, in cluster 4 (open/swinging door state), residues of IDE catalytic sites appear to less favor forming binding sites (Table S1, Figure S3). In addition, other binding sites, which are not available in the clustroids of clusters 1-2, are formed by residues E577, Y584, Y625, W695, and R765 of the clustroids of clusters 3-4. In contrast, residues S816, F820, R824, V833, and S835 are forming binding sites in the clustroids of clusters 1-2, but not in those of clusters 3-4. Uniquely present in cluster 1 but not in clusters 2-4, binding-site-forming residue E205 was reported to form a hydrophobic exosite where a quinoline molecule [13] and the IDE inhibitor 6bk [20] bind to. Interesting residues that we identified (Table S1) can be further explored through virtual screening, for designing probes/agents that (1) shift IDE conformational states from closed to open or vice versa, and (2) stabilize a particular conformational state of IDE for desirable timeframes and locations (Supplementary Info S4).

4 Conclusion

To examine structural dynamics in relation to molecular functions of IDE, we have applied an integrative sets of established bioinformatics methods in a rare combination, allowing discoveries of novel structural features of IDE. Structural analyses of IDE can enable us to learn how IDE alter its substrate specificity, distinguish the molecular mechanisms associated with the etiology of type 2 diabetes mellitus or Alzheimer's disease, and gather hints that may benefit the development of therapeutic agents, such as small molecules, stabilized peptides, or IDE mutants.

5 Supplementary Materials and Acknowledgments

Supplementary materials are available from http://tinyurl.com/idepapersup. We thank Dr. Kelvin Sim for comments and our colleagues at Khalifa University for critically reading our manuscript. We also acknowledge the 2015 Khalifa University Internal Research Fund Level 1 210070 for funding our research.

6 References


6. Ren M, Guo Q, Guo L, Lenz M, Qian F. Polymerization of MIP-1 chemokine (CCL3 and CCL4) and clearance of MIP-1 by insulin-degrading enzyme. EMBO J. emboj.embopress.org; 2010; Available: http://emboj.embopress.org/content/29/23/3952.abstract


