Comparative analysis of permanent and transient domain-domain interactions in multi-domain proteins

Ramanathan Sowdhamini[1], Swayam Prakash Das Sidhanta[1], and Narayanaswamy Srinivasan[1]

[1]Indian Institute of Science Molecular Biophysics Unit

June 1, 2023

Abstract

Protein domains are structural, functional, and evolutionary units. These domains bring out the diversity of functionality by means of interactions with other co-existing domains and provide stability. Hence, it is important to study intra-protein inter-domain interactions from the perspective of types of interactions. Domains within a chain could interact over short timeframes or permanently, rather like protein-protein interactions (PPIs). However, no systematic study has been carried out between two classes, namely permanent and transient domain-domain interactions (DDIs). In this work, we studied 264 two-domain proteins, belonging to either of these classes and their interfaces on the basis of several factors, such as interface area and details of interactions (number, strengths, and types of interactions). We also characterized them based on residue conservation at the interface, correlation of residue motions across domains, its involvement in repeat formation, and their involvement in particular molecular processes. Finally, we could analyse the interactions arising from domains in two-domain monomeric proteins, and we observed significant differences between these two classes of domain interactions and a few similarities. This study will help to obtain a better understanding of structure-function and folding principles of multi-domain proteins.

1. Introduction

The existence and functioning of any organism can be seen to be solely due to proteins in its cellular environment. Most of the functionalities arise due to several interactions of proteins with various macromolecular entities like nucleic acids, lipids, carbohydrates, etc., or with other proteins. Among these interactions, protein-protein interactions (PPIs) are of primary importance as these interacting complexes play crucial roles in several cellular processes like replication, transcription, translation, regulation, signaling, etc.. These protein-protein complexes (PPCs) can be categorized into different groups based on the proportion of interacting protomers or stability of protomers or the lifetime of interactions into homo/hetero complexes or obligate/non-obligate complexes or permanent/transient complexes, respectively. The complexes where the protomers become unstable when they are separated are obligates, while the complexes where the interactors remain stable even though they are separated are non-obligates. On the other hand, the complexes where the protomers interact throughout their functional lifetime are permanent, while the complexes where the protomers associate and dissociate temporarily are transient complexes. In general, obligate complexes are permanent both structurally and functionally, while non-obligate complexes are mostly transient associations with a few permanent associations like antibody-antigen complexes. One of the great examples of such interaction could be the heterotrimeric G protein. G protein consists of three subunits: α, β, and γ, where β and γ subunits interact throughout their lifetime, making it a permanent interaction. Instead, the α subunit interacts transiently to $\beta\gamma$ complex when inactive and dissociates when active, making it a transient interaction. Among these, transient interactions are of utmost importance as these complexes are crucial for various biological processes as they act as hubs in protein-protein interaction networks (PPINs), are multi-specific, are great drug targets and are involved in various cellular processes.

There are several studies that distinguish structural characteristics of such interaction types amongst proteins and PPIs. Few such physicochemical properties which discriminate permanent and transient interactions are contact area, interface shape and size, number of contacts, polarity, hydrophobicity, complementarity of the interface, involvement of secondary structures at the interface, evolution of the interface, etc. Based on these properties, several groups focused on distinguishing these two types of PPIs. Few groups focused solely on the physicochemical properties or interfacial properties to predict permanent and transient PPIs. Some groups represented these physicochemical properties into vectors for better prediction using machine learning approaches, while some researchers used desolvation energy explicitly to predict permanent and transient. Apart from these, some used sequence features to predict permanent or transient and few developed algorithms which do not require information about binding partners for prediction.

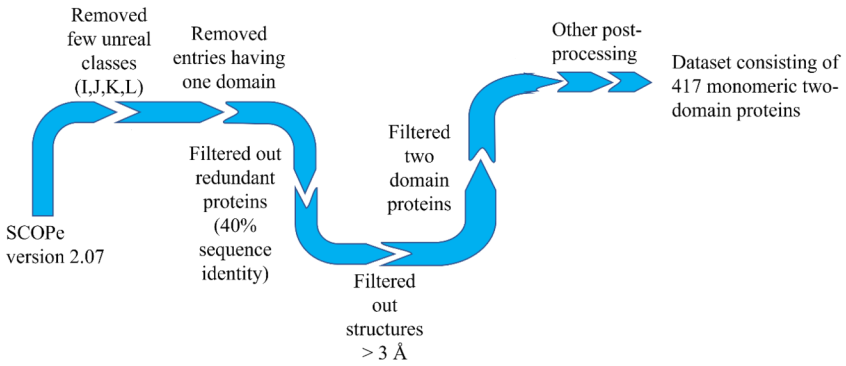
After the first enzyme was solved, it was found that there are some distinct lobes present within the protein. However, the term 'domain' was coined by Wetlaufer by defining these entities as structurally independent regions within proteins. These domains are also referred to as units of protein evolution. Several structure-based identification and analysis of domains have been performed and organized as databases. The vast functional diversity of a protein arises by combining such domains into a single polypeptide chain, calling it a multi-domain protein, and most proteins, even in a simple proteome, are multi-domain proteins. The interactions amongst multi-domain proteins with other such proteins are mainly carried out by a portion of the protein structure, a protein domain, rather than the whole protein. The interactions between domains are called domain-domain interactions (DDIs), and they generally facilitate protein interactions. It is also observed that interacting domain pairs tend to co-evolve with each other in an interaction in order to maintain a better interaction. The domain pairs are also consistent with their parent protein interactions. There are few studies that take help of known structural DDIs to predict PPIs, whether these are input sequences or structures. Deng and coworkers used maximum likelihood approach to estimate the probabilities of domain pairs in protein interactions to predict PPIs. A J Gonzalez and Li Liao used fisher scores derived from the domain interaction profiles as features to predict DDI using SVM, which can be used to predict PPI. Instead of using generative methods of predicting PPI, Zhao *et al.* used information of both PPI and non-PPI to infer DDI, which in turn can be used again to predict PPI from the inferred DDIs. Similarly, Sprinzak and Margalit used correlated sequence signatures in proteins to predict DDI.

Often, functional characteristics of a multi-domain protein is dependent on the arrangement of the domains in it and interactions among them, which can be compared to arrangement of words to form meaningful sentences in natural languages. Interactions among the domains facilitate proper functioning of multi-domain proteins. These domains are also known to be responsible for functional and evolutionary relationship of proteins. The occurrence of multiple domains also confers additional stability to individual domains and hence the whole protein. Hence, there is a need to study inter-domain interactions, mostly in monomeric proteins, for their resident time of interactions or strength of interactions which could provide immense knowledge about the functional and structural aspects of multi-domain proteins. However, unlike studies differentiating PPIs into permanent and transient interactions, there is no systematic and organized approach to classify DDIs into permanent and transient interactions. Instead, there are a few studies which investigate DDIs in a single polypeptide chain and regard such interactions to be either permanent interactions or to have characteristics intermediate between PPIs.

In this work, we extended the concept of permanent and transient interactions to intra-protein inter-domain interactions and characterized the underlying interaction types. Using a dataset of monomeric two domain proteins whose domain definitions are taken from SCOPe, we could identify such domain interactions to be either permanent or transient. Permanent and transiently interacting domains are not much different in terms of evolution of the interface, and the type of functions they are involved in, when investigated human proteome only. However, we found that these two types of DDI differ in their physicochemical properties of their interface, dynamically correlated motion of their residues, and preference for choosing its interacting partner. This work would shed light on the principles of domain interactions, prediction of domain orientation, and protein functioning by these rules of domain interactions. Structurally, this study would also help to understand the folding of multi-domain proteins correctly in the near future.

2. Materials and methods

2.1 Dataset creation: monomeric two domain proteins



We used SCOPe (dir.cla.scope 2.07) database for structural domain definitions. Entries or proteins having either single domain or only single domain available in the database were removed. Further, some unsuitable SCOPe classes (such as low resolution protein structures, peptides, designed proteins, and artifacts belonging to classes I, J, K, L, respectively) were removed. For the analyses to be conducted on a non-redundant protein set, a 40% sequence identity was set for clustering proteins using CD-HIT. The resulting entries were filtered for monomeric proteins solved by X-ray crystallographic method in RCSB filter using parameters like asymmetric unit, biological unit, experimental method, and structures with 3Å or better resolution. Proteins having only two domains were next alone considered through SCOPe definitions (only continuous domains were taken). Finally, the structure having the best resolution was taken as the representative structure for the RCSB entries of proteins. The various filtering steps for dataset creation are summarized in Figure 1.

2.2 Identification of domain-domain interactions

Identification was done only for those protein structures whose domains interact with each other. To define interacting domains, 5-5 rule was used, which states that interacting domains have at least five residue contacts within 5Å. The distance criterion was adopted using our in-house PIC software. Further, at least five interactions arising from residues of domains were considered that are at least six residues apart to consider a short linker connecting two domains, which would include linkers of varying lengths. The classification of DDIs in monomeric multi-domain proteins to obligate and non-obligate (here permanent and transient) ones were done using NOXclass. This tool is a SVM classifier which is based on the physicochemical properties of the interface. For this study, we used the parameters which showed highest accuracy using multi-stage SVM. Necessary PDB manipulations were done using pdb-tools. To get a cutoff to define the interaction as obligate or non-obligate, this was tested on Block *et al.* dataset and tried to match the accuracy of prediction of NOXclass with different cutoffs, resulting in a cutoff of 70% to distinguish the interaction as obligate and non-obligate. We also used multiple structures of proteins to check large structural deviations (>2Å) using MUSTANG, and reclassified the interaction, wherever needed, based on literature.

2.3 Interfacial properties

We used our in-house server, PIC, to obtain interactions arising from two domains by taking care of multiple occupancies of atoms.

A Python script obtained from Pymol (The PyMOL Molecular Graphics System, Schrödinger, LLC) was used for the recognition of interfacial residues, which is based on the change in solvent accessibility upon complex formation.

Interaction energies of DDIs were calculated using PPCheck, which measures energies as sum of van der

Waals, hydrogen bond, and electrostatic interactions. The energy of the proteins was minimized using GROMACS for those proteins which showed unfavourable calculated energies.

2.4 Gene Ontology studies

We used PANTHER to carry out gene ontology studies on both interaction types. As these studies are difficult for a dataset containing multiple genomes, we only considered the highest occurring genome in the dataset, i.e., human genome, consisting of 24 and 25 two-domain proteins containing permanent and transient domains, respectively.

2.5 Conserved interfacial residues

ConSurf-DB was employed to identify conserved residues across domains. It is a database for evolutionary rates of residues of a protein of known structure. We used the ConSurf colors greater than 7 to define conserved positions. Common residues to both ConSurf and interface residues were considered as conserved interfacial residues.

2.6 Correlated residue movement

ProDy was used to perform Anisotropic Network Model (ANM) based Normal Mode Analysis (NMA) calculations. Twenty modes were calculated, and the same were used to calculate cross-correlated motions of C^{α} atoms, keeping a cutoff of 0.7 correlation value to define high correlation. Residue cross-correlation of domain1 with domain2 was only considered.

2.7 Repeats analysis

Uniprot was used to know the presence of sequence repeat containing proteins in the dataset. RepeatsDB was used to get structural repeats populating at least one domain in proteins in the dataset. We used SCOPe "sccs" id till superfamily level to define homodomain containing proteins and used fold information to get folds of domains.

3. Results and discussions

3.1 Intra-protein domain interactions can be further classified as permanent and transient

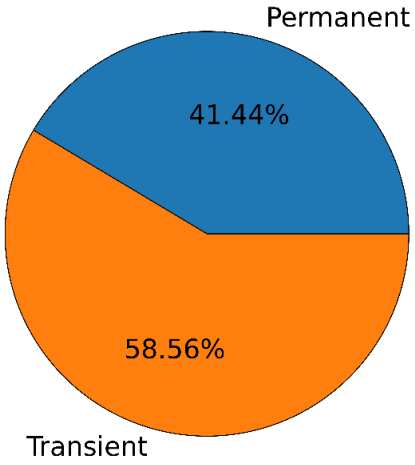
Interactions between domains in multi-domain proteins are now possible to study with the availability of larger such entries in structural databanks. Moreover, structures are more conserved than sequences, which makes protein studies more accessible. In order to avoid complications of higher order domain interactions, we have considered only interactions arising from two structural domains within a single polypeptide chain which would eliminate the interfering effect of another domain in other chain(s). For this study, we created a protein structural dataset of monomeric two-domain proteins (Figure 1).

Next, the dataset was classified into DDIs which permanently or transiently interact. As the domain interface and subunit interface are somewhat similar, we searched for a tool to predict PPIs as permanent and transient, which could be used for domain interactions within proteins. We chose NOXclass since this is one of the highly accurate classifiers and is easy to use.

From a dataset of 417 monomeric two-domain proteins, only 264 proteins could be classified due to stringent cutoff, and the rest of the proteins were of lower confidence. We observed that around 109 proteins retain permanent inter-domain interactions, while around 154 of them showed to have transient interactions (Figure 2A, Supplementary Table S1) at a stringent cutoff. Comparatively large number of proteins showed to have transient interactions, which could prove their inherent flexibility to accommodate any function of the protein.

There are few studies which compare inter-chain protein interactions to intra-chain interactions and comment on their resident time of interaction. In one such study, the authors analyzed protein interaction sites by taking 750 transient PPIs and 2000 domain interactions within a chain. The authors assumed such domain interactions within the same protein chain as obligate interactions. In another such study, the authors analyzed six different types of interfaces in protein structures, and domain-domain interface within

a single chain was one of the six interface types. They viewed such interaction as permanent interaction between independent folding units and compared these with hetero-obligomers. On the other hand, it is also seen that most of the domain interfacial properties within a chain are intermediate between inter-chain permanent and non-obligate complexes. The observations from our study clearly suggest that intra-chain DDIs can also be classified as permanent and transient interactions. Moreover, we could further diverge domain interfaces which are intermediate between permanent and transient PPI into permanent and transient domain interactions, which are discussed in detail in the following sections.



3.2 Permanent domains have enhanced interfacial properties

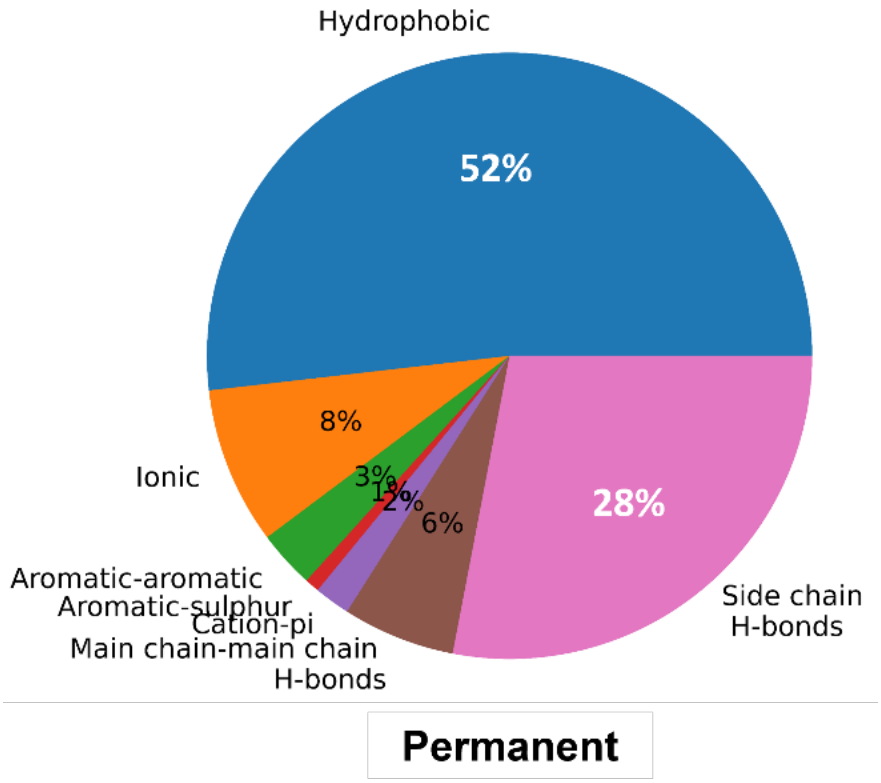
To understand the interfacial features of DDIs within a single chain which could be responsible for their permanent and transient behaviour, we first investigated the total number of interactions through which two domains are held. We observed that most of the proteins having transient domains have a comparatively smaller number of interactions (Figure 2B). Also, around 91% of the transient domain containing proteins have interactions less than 75 in number. On the other hand, comparatively more permanent domain proteins have a larger number of interactions. The total number of interactions between permanent domains follow a near similar uniform distribution throughout different number of interaction ranges. We then computed interfacial areas to find the interface size of different domain-domain interfaces. From Figure 2C, we infer that proteins having transient domain interactions have comparatively smaller distribution of interface areas than proteins having permanent domain interactions. The interfacial areas of transient domain interactions are concentrated around the median of the distribution which suggests that these domains have smaller interfaces consistently. Instead, permanent domains of proteins have widespread interfacial areas, of which most of the domains have larger interfaces which is evident from comparatively large difference between upper quartiles in the boxplot. The domain pairs which have larger interfaces than 4000Å^2 are listed in Supplementary Table S2, and most of the large transient domain interfaces are outliers (Figure 2C). This shows that permanent domains harbor larger elaborate interfaces than transient domains. The better the interaction energy of a complex, the stronger the binding and stability. For this case, we next checked the strength of domain interactions in these two kinds of interacting domains. From Figure 2D, it is observed that permanent domains indeed have better interaction energies than transient domains, as measured through PPCheck. The energies associated with permanent domains are more stabilizing than transient domains. It

is also observed that the energies of transient domains are concentrated to a comparatively lower stabilizing energy, while the energies of permanent domains have a wide range of interaction strengths.

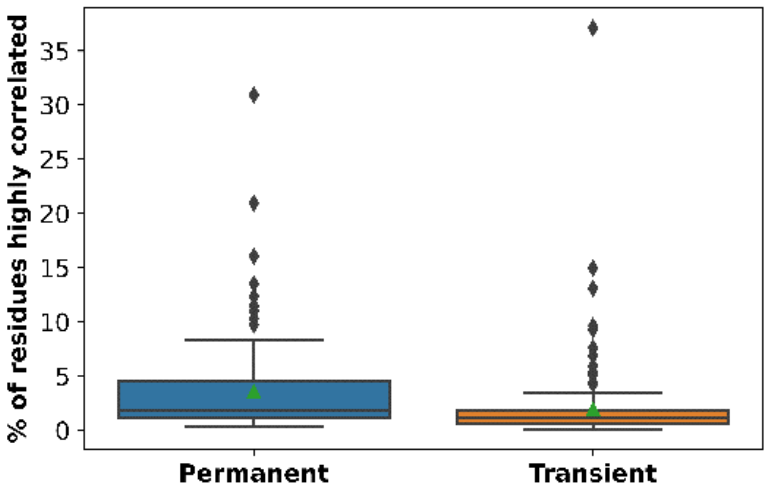
It is noteworthy that the number of permanent and transient domain containing proteins in our dataset are not equal, where we had comparatively a greater number of domain pairs in transient interactions than permanent ones. Therefore, we sampled random number of entries from transient domain pair dataset to match permanent domain pair dataset, and the observed trends are very similar to asymmetric dataset. Although, permanent domains have higher interfacial physical properties, it is also observed that both permanent and transient have similar average number of interactions per interfacial residue, 0.685 and 0.641 interactions per interfacial residue, respectively, which would mean that the residue interaction networks at the interface are not much different. The average number of interactions per interfacial residue is a proportional value and hence could be the reason for such similarity. The interface of such domain interaction types reveals many discriminatory facts that would help to distinguish permanent and transient domain interactions.

3.3 A tie between hydrophobic interactions and hydrogen bonds: permanent and transient domains

It is clear from the previous analysis that permanent domains have a larger number of interactions between interacting domains than transient domains. To obtain a clearer perspective on the interactions, we probed atomic interactions of interfacial residues. We observed around 52% of the total number of interactions in the case of proteins having permanent domains are hydrophobic (Figure 3A). On the contrary, as shown in Figure 3B, transient domains have only 37% of hydrophobic interactions. These hydrophobic interactions are known to drive different PPIs and are known to comprise major interactions in the biomolecules which stabilize interacting complexes. Moreover, we found that transient domains have large proportions of side chain associated hydrogen bonds in comparison to permanent domains (Figure 3A, 3B and Supplementary Figure S1), and such polar interactions are known to bring out specificity. Apart from hydrophobic interactions and side chain associated hydrogen bonds, all other interaction types were similarly populated in the interface, which would be required for sustained domain interactions and the functioning of multi-domain proteins.



3.4 Residues of permanent domains have higher correlated motion

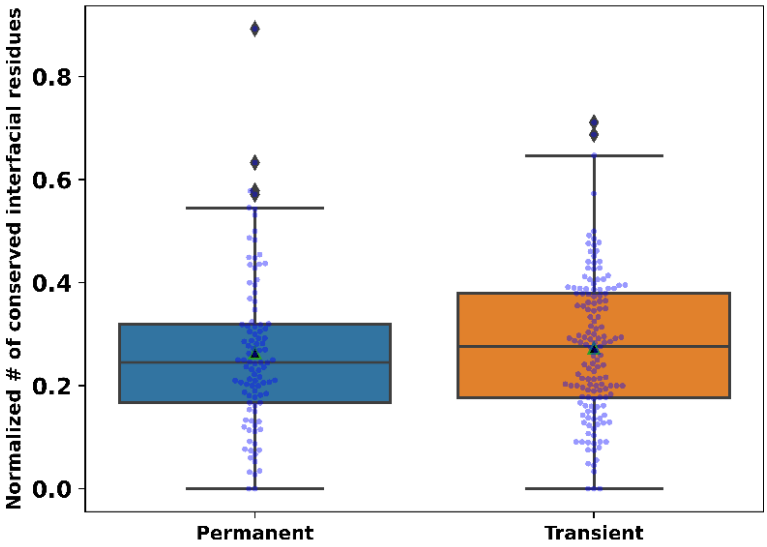


A variety of functions of proteins are achieved by cooperative motions of their constituent atoms. This cooperativity is further achieved by crosstalk among domains of the proteins either by physical contacts or by correlated motions of its atoms. To explore any discrimination of dynamic behavior of residues between

permanent and transient domains, we studied residue cross-correlation of motions, which are represented by C^{α} atoms using Anisotropic Network Model (ANM) of Normal Mode Analysis (NMA). Cross-correlation values range from -1 to 1, and we considered those residue pairs to be highly correlated if their value is greater than 0.7 to keep a balance between the number of correlated residues and their high correlation. Figure 4 shows the differences in inter-residue correlated motions between domains interacting permanently and transiently. Residues from permanent domain pairs showed a wider range of highly correlated motions than residues from transient domains. When the data were plotted in a histogram to better understand the difference, we found maximum number of transient domain pairs to have extremely low percentages of highly correlated residues, while consistently, a greater number of permanent domain pairs showed a higher percentage of highly correlated residues (Supplementary Figure S2). The same analysis was tested with different cutoffs to define high residue correlations ranging from 0.5 to 0.9, and the patterns obtained were similar. This observation implies that the domains which interact transiently carry out short range correlated motions, whereas in case of permanent domains, extensive interactions are going on across domains and can engage in long range correlated motion. The range here conveys the strength of interaction or force of movement using a large number of residues (long) or a small number of residues (short).

Such behavior of permanent domains could be thought of due to their lifetime of interactions. These residues are needed to be synchronous to maintain the integrity of the domain interface, and this correlated motion would help the domains to maintain resonance for the stability of the monomeric protein. On the other hand, transient domains would need to associate and dissociate frequently. Comparatively lower percentage of highly correlated residues between domains would be enough to maintain the interface and hence the transient nature. This clearly conveys how dynamics is associated with the long-term interactions within protein interiors.

3.5 Number of conserved interfacial residues is similar in permanent and transient domains



Residues present in the interface of interacting partners are solely responsible for communication between the partners. There are several studies on PPI, which state higher conservation of interfacial residues than remaining protein surface. It is also known that interfaces of permanent protein complexes have a lower evolution rate than transient interactions, which allows better co-evolution with its interacting protein

partner. Similarly, in this case, we ought to look into the conservation at the domain interfaces of different DDI types in a protein chain. Due to the requirement of ConSurf-DB to have at least a certain number of homologues to follow the evolutionary rate, some of the proteins in our dataset could not be retrieved from the database. Hence, we analyzed only 101 and 147 permanent and transient domain pairs in proteins, respectively. We computed the absolute number of interfacial residues which are conserved in domain pairs and observed that permanent domain pairs have a little wider distribution of the number of conserved interfacial residues than transient domain pairs, which could be due to a large number of interfacial residues arising from larger interfaces. But the number of such residues in both domain types is not significantly different to account for any dissimilarity (with a Mann-Whitney p-value: 0.1341 and two sample KS test: 0.1882). Next, we also computed the normalized number of conserved interfacial residues, that is, the number of conserved interface residues per total number of interface residues in a domain pair of a protein, and we observed near similar distribution of such residues with respect to their interface residues with near identical mean and median (Figure 5). These observations suggest that both permanent and transient domain pairs have similar proportions of conserved interfacial residues. Unlike PPIs, this similarity in maintaining the conservation at the interface could be due to the fact that the interacting partners, here domains, are referred as semi-independent and evolutionary units which are thought to be conserved. Like permanent PPI, residues in the permanent domains might be under evolutionary pressure to co-evolve with partner domains. On the contrary, two transient domains might be harboring functional sites at their interface and hence the obligation to preserve the interfacial residues. Each domain interaction type has its compulsion to maintain the interface geometry, resulting in similar preferences to conserve the domain interfaces in the two-domain protein.

3.6 Permanent domains structurally prefer similar folds

Different types of domain interactions in a protein chain might have some influence on the anatomy of the protein structure landscape. Hence, we next explored a few of the structural aspects which could be discriminated by permanent and transient domain interactions within a protein. Firstly, we aimed to look at their preferences to have repeats. Repeats can be of two types, viz, sequence repeats and structural repeats. Structural repeats can be further classified into different classes. Using different databases (see methods) to map the proportion of proteins in our dataset to have repeats, we found a few proteins in both repeat types where proteins having permanent domain interactions showed a little more preference for sequence and structural repeats. However, this observation cannot be relied upon due to the sparse number of proteins. Among the structural repeats, the repeating units (domains) of bead-on-string repeats (class-IV) are thought to either interact loosely or not interact, which could have been interesting examples of transient domains in multi-domain proteins. However, from the proteins having at least one structural repeat containing domain, there was no protein which belonged to this class. This could be due to the limited amount of information in the database or due to the limited number of domains in our study to represent multi-domain proteins. Secondly, to overcome this limitation, we defined homodomains, where both domains have same class, fold, and superfamily according to SCOPe. Thus, these domains will have similar architecture and are evolutionarily related to each other, which are supposed to be originated by duplication. Using such a definition, we observed a comparatively higher proportion of permanent domain containing proteins to have homodomains, 37.3% in comparison to 28.6% of homodomains in the dataset. Although these homodomains may not be true tandem repeats, such domains can provide functional and structural advantages to the proteins having permanent domains due to evolutionary pressure and topological constraints, respectively. Thirdly, to investigate their structural constraints, we explored their fold distribution in homodomains. We found that proteins having permanent homodomains have a comparatively lower number of unique folds than transient homodomains, which could signify the capability to re-use folds. This suggests that if domains interact permanently in a protein, there is a greater chance of finding another interacting domain of common ancestry and similar structural topology. This observation is similar to the observations of PPI, where obligate PPI tends to have more homo-DDIs. When we considered the whole dataset to look into the number of unique folds, both permanent and transient domain pairs showed a similar count of unique folds quantitatively. However, qualitatively, we observed a few biases of folds toward permanent and transient

domain interactions (Table 1). Superfolds such as TIM beta/alpha-barrel, OB fold, and beta-grasp showed an inclination towards transient domains. On the other hand, 7-bladed beta-propeller, Ribonuclease H-like motif fold, and a few others showed inclinations towards permanent domains. Apart from that, superfolds like Immunoglobulin-like beta-sandwich, DNA/RNA-binding 3-helical bundle showed preferences for both permanent and transient domains. Other sparsely occurring folds (frequency: less than 5) showed little or no bias (Supplementary Table S3 and S4). These observations show the structural preferences of different domain interaction types and also justify how a limited number of folds are re-used to sample various protein structural landscapes in DDI following a power-law. This will enlighten the basic principles of domain interaction type prediction, given that we know the interacting domains in a protein, their topology, and evolutionary information.

3.7 Functional classification of domain interaction types using Gene Ontology

Next, we explored if there are functional preferences between permanent and transient domains. As our dataset consists of proteins from various genomes, we only considered the genome which populates the maximum number of proteins in our dataset. Using Gene Ontology functional classification analysis, we observed that both permanent and transient domain containing proteins in humans are involved in similar kinds of biological processes, molecular functions and belong to similar protein classes, and there is no bias (Supplementary Figure S3). The variance in functional roles may be seen if the whole dataset is compared, and this needs more sophisticated algorithms, which are out of scope at present.

4. Conclusions

Interactions between domains are responsible for the functionality of a protein. Apart from functional advantages, domains in multi-domain proteins provide additional stability to proteins and to the neighboring domains. Studying types of domain interactions in multi-domain proteins, focusing on their resident time becomes crucial to understand the intra-protein interactions. In this current work, we recognized DDI arising from two domains in a monomeric multi-domain protein as permanent and transient using an algorithm used to classify PPIs. We demonstrate that permanently interacting domains have larger interfaces that facilitate larger number of interactions between the domains, which in turn support stronger interactions. Their interfaces are populated by a larger proportion of hydrophobic interactions, while transient domain interfaces have comparatively lower hydrophobic interactions, which are compensated by large number of side chain associated hydrogen bonding. A comparatively increased number of residues in permanent domains have highly correlated motions. Domains interacting permanently have a higher chance of interacting with a structurally similar domain, and there are a few topological biases for each interaction type. Furthermore, both permanent and transient domains have equal number of conserved interfacial residues, and the domains in the human genome do not discriminate upon the functions or processes they are associated with. We note that few of these observations are consistent with the way permanent and transient PPIs differ from each other.

This work will be very useful to understand the molecular basis of function and how the functional sites are disposed in 3D structure. This analysis provides objective realization that two-domain monomeric proteins which are permanently interacting are more likely to adorn their interface by hydrophobic residues. This observation is certainly of predictive value to obtain clues on biochemical function and to recognize reasonable poses while performing domain-domain docking and modeling.

Acknowledgments

The authors thank lab members for helpful discussions. SPDS thanks Dr. Ashraya Ravikumar for discussions on the work. SPDS thanks Hongbo Zhu for providing NOXclass files. SPDS thanks CSIR for financial support. RS acknowledges funding and support provided by JC Bose Fellowship (JBR/2021/000006) from Science and Engineering Research Board, India and Bioinformatics Centre Grant funded by Department of Biotechnology, India (BT/PR40187/BTIS/137/9/2021). RS would also like to thank Institute of Bioinformatics and Applied Biotechnology for the funding through her Mazumdar-Shaw Chair in Computational Biology (IBAB/MSCB/182/2022).

Author contributions

NS conceived the idea and concept for the work. RS designed the work. SPDS carried out all the analysis and wrote the manuscript. RS shaped and finalized the manuscript.

Conflict of interest

The authors declare no conflict of interest.

Data availability

The dataset is available in the supplementary materials and the PDB codes can be accessed using RCSB site.

ORCID

Swayam Prakash Das Sidhanta https://orcid.org/0009-0003-3707-5452

Narayanaswamy Srinivasan https://orcid.org/0000-0002-6566-6239

Ramanathan Sowdhamini https://orcid.org/0000-0002-6642-2367

References

Supporting information

Additional information can be found in the supplementary materials.