Received: 5 October 2020 | Revised: 13 April 2021 | Accepted: 16 April 2021

DOI: 10.1002/prot.26087

RESEARCH ARTICLE

PROTEINS WILEY

Marker residue types at the structural regions of transmembrane alpha-helical and beta-barrel interfaces

Sercan Beytur

Faculty of Engineering and Natural Sciences, Department of Bioinformatics and Genetics, Kadir Has University, Istanbul, Turkey

Correspondence
Sercan Beytur, Faculty of Engineering and Natural Sciences, Department of Bioinformatics and Genetics, Kadir Has University, Istanbul, Turkey.
Email: sercan.beytur@live.fr

Abstract

Membrane proteins play a variety of biological functions to the survival of organisms and functionalities of these proteins are often due to their homo- or hetero-complexation. Encoded by ~30% of the genome in most organisms, they represent the target of over half of nowadays drugs. Spanning the entirety of the cell membrane, transmembrane proteins are the most common type of membrane proteins and can be classified by secondary structures: alpha-helical and beta-barrel structures. Protein-protein interaction (PPI) have been widely studied for globular proteins and many computational tools are available for predicting PPI sites and construct models of complexes. Here, the structural regions of a non-redundant set of 232 alpha-helical and 37 beta-barrel transmembrane complexes and their interfaces are analyzed. Using the residue composition, frequency and propensity, this study brings the light on the marker residue types located at the structural regions of alpha-helical and beta-barrel transmembrane homomeric protein complexes and of their interfaces. This study also shows the necessity to relate the frequency to the composition into a ratio for immediately figuring out residue types presenting high frequencies at the interface and/or at one of its structural regions despite being a minor contributor compared to other residue types to that location's residue composition.

KEYWORDS
composition, frequency, membrane proteins, propensity, protein-protein interface

1 | INTRODUCTION

Membrane proteins represent around 30% of the proteome in most organisms and are targeted by over 50% of nowadays drugs due to their diversity of biological functions necessary to the survival of organisms, such as signal transduction, electron transport and ion conductance.[1-3] Functionalities of membrane proteins, as globular proteins, are often due to their homo- or hetero-complexation and, therefore, searchers have considered protein-protein interactions (PPI) as increasingly important therapeutic targets.[4] Many diseases are related to the disfunction of membrane protein complexes and contrary to globular protein complexes, for which plenty of available three-dimensional structures and studies deepening the understanding of the structural and interacting components of these protein complexes and prediction tools exist, it needs to be deepened, notably the knowledge of the residue types occurring at their interfaces.[5-6]

Membrane proteins can be classified by secondary structure type: alpha-helical and beta-barrel structures.[7] Alpha-helical transmembrane proteins represent 27% of all proteins in humans and are mostly present in the inner membranes of most cellular membranes.[8] Beta-barrel transmembrane proteins are found only in outer membranes of gram-negative bacteria, mitochondria, and chloroplasts and have a simplest up-and-down topology, reflecting a common evolutionary origin and folding mechanism.[9-10]

An interface is formed upon the complexation of identical monomers (homo-complexation) or different monomers (hetero-complexation) and corresponds to the contacts between the involved monomers via the residues building the interface. A residue was

defined as being an interface residue if the change in its accessible surface area (ASA) upon complexation (going from a monomeric state to a dimeric state) is larger than 1 Å[2,11-12] Then, it has been shown that residues contributing the most to the binding energy are protected from the solvent.[13] Therefore, an interface is not a uniform entity and a first model was proposed where a central desolvated region is surrounded by residues in contact with water and, by analogy to the interior-surface dichotomy for protein structure folding, a core-rim dichotomy was proposed for protein-protein interfaces: where rim is formed by residues containing solvent-accessible atoms only and where the remaining residues of the interface are forming the core.[14] Different studies supported this model and have shown that amino acids forming the interface core tend to be more hydrophobic than over the rim.[15-17] It is also known that they are more frequently hotspots and, therefore, usually more conserved.[18] Deepening these studies, a formal structural definition of regions in protein complexes was proposed and a new structural region, the support, was introduced.[19] In this last study, structural regions of protein complexes and of their interfaces were defined by a combination of two measurements: the relative ASA (rASA) and the difference in rASA (ΔrASA) upon complexation of each amino acid within a protein-protein complex.

Plenty of studies exist for globular proteins and one of them especially has shown the hydrophobic aspect of their interfaces, crucial for the stabilization of protein-protein complexes, especially aromatic residues which can form strong hydrophobic interactions between the bulky hydrophobic side chains.[20-21] In transmembrane proteins, it has been shown a tendency of large hydrophobic residues to be located at the protein surfaces facing the lipids in beta-barrel proteins, when in the alpha-helical ones these residues are equally distributed between the interior and the surface of the protein.[22]

Recently, it has been shown that the core of protein-protein interfaces has similar amino acid compositions to that of the membrane-embedded regions of transmembrane proteins but some differences are seen in composition of alpha-transmembrane proteins like a higher frequency of Ala and Gly in their core, which is consistent with the GLY-XXX-GLY motif found in transmembrane helix-helix association.[23] This motif is involved in transmembrane helix-helix-interactions modulation.[24] Based on Levy's model, it was also revealed that support residues are significantly more conserved than the rest of the protein, whereas rim residues are significantly less conserved.[25] Core residues display intermediate profiles. This information permitted to update the evolutionary information regard to Levy's definition of structural regions in proteins.

However, all these studies do not provide information about the difference in marker residues between alpha-helical or beta-barrel transmembrane protein interfaces. For contributing to the knowledge deepening of these categories, the structural regions of a nonredundant set of 232 alpha-helical and 37 beta-barrel transmembrane homomeric complexes and their interfaces, obtained from OPM database, have been analyzed.[26-27] Using the residue composition, frequency and propensity, this study brings the light on the crucial residue types located at the structural regions of these complexes, especially the structural regions of their interfaces. This study points also the necessity to relate the residue frequency to the residue composition into a ratio for immediately figuring out residue types having high frequencies at the interface and/or at one of its structural regions, despite being a minor contributor compared to other residue types to that location's residue composition.

2 | MATERIALS AND METHODS

2.1 | Amino acid hydrophobicity scale

The amino acid classification follows the hydrophobicity scale proposed by Moret and Zebende in 2007, based on the variation of the amino acid's accessible surface area (ASA).[28] Amino acids are classified into three categories: hydrophobic, intermediate and polar.

2.2 | Transmembrane homomers dataset

A total of 232 alpha-helical and 37 beta-barrel nonredundant and biologically relevant membrane homo-dimeric and trimeric structures downloaded from OPM database are constituting the database (Table 1). OPM database is providing information about the orientation of proteins within membranes, useful for defining inner-membrane from outer-membrane surfaces. Alpha-helical and beta-barrel datasets are containing a total of 208 637 and 41 053 residues, respectively, including 33 875 and 8120 interface residues (Table 2).

2.3 | Definition of the structural regions: Levy's model

Five structural regions can be defined in Levy's paper[19] based on the relative accessible surface area (rASA) of each monomer of a protein complex, in its free state (rASAm), after complexation (rASAc) and their difference upon complexation (ΔrASA = rASAm − rASAc). rASA is the normalized form of the ASA. For a residue *i*, its measured ASA within a protein is divided by its maximum theoretical ASA value, which corresponds to its ASA in a free state. These maximum theoretical ASA values are taken from Tien et al.[29] and the calculation is made as follows:

$$\text{for a residue } i,\ \text{rASA}_i = \left(\frac{\text{measured ASA}_i}{\text{maximum theoretical ASA}_i}\right)$$

Interior region is characterized by a rASA <0.25 and the surface regions by a rASAc >0.25, in addition of a ΔrASA = 0 for both non-interface regions. Interface region is characterized by a ΔrASA >0 and its three structural sub-regions are defined by the addition of: a rASAm >0.25 for the core, rASAm >0.25 and rASAc <0.25 for the rim, rASAm <0.25 for the support region.

TABLE 1 List of non-redundant transmembrane alpha-helical and beta-barrel homomeric complexes

Alpha-helical (n = 232)							Beta-barrel (n = 37)
1ap9	3oe0	4kjs	5cfy	5wuf	6jbj	6vp0	1a0s
1e12	3org	4mnd	5ctg	5xls	6jpf	6wc9	1af6
1kpl	3qnq	4mrs	5d92	5xmj	6kkt	6wm5	1pho
1l7v	3t56	4mt1	5dqq	5xu1	6kuw	6xdc	2fgq
1mhs	3tij	4myc	5egi	5y79	6l3h	6xwr	2mpr
1ots	3tui	4ntf	5eik	5z1f	6l47	6y5r	2o4v
1q16	3ug9	4o6m	5gko	5zih	6m96	6y9b	2por
2a65	3ukm	4o6y	5h35	5zlg	6m97	7bp3	2xe1
2b2f	3um7	4oh3	5h36	5zov	6mgv	7bve	3a2s
2bs2	3vvk	4or2	5i6c	6a2w	6n51	7bx8	3nsg
2ei4	3w9i	4pl0	5i9k	6ak3	6nf4		3prn
2hyd	3wme	4q2e	5iji	6b87	6nf6		3upg
2m3g	3x3b	4qi1	5iws	6bhp	6npl		3wi5
2mpn	3ze5	4qnc	5jnq	6bqo	6nq0		4aip
2nq2	4a01	4qnd	5jsi	6c96	6nt6		4aui
2ns1	4ain	4qtn	5khn	6caa	6nt7		4d65
2onk	4av3	4r0c	5l22	6cb2	6nwd		4rjw
2q7r	4ayt	4r1i	5l24	6coy	6nwf		5dqx
2qfi	4bpm	4ri2	5l25	6csm	6o84		5fq6
2uuh	4bw5	4rng	5lil	6d0j	6oce		5fvn
2vpz	4bwz	4rp8	5mju	6d79	6oht		5ldv
2yvx	4cz8	4ry2	5nj3	6dz7	6ows		5mdq
2z73	4czb	4ryi	5o5e	6e1h	6pis		5nuq
3b5x	4djh	4tl3	5o9h	6eid	6pzt		5nxn
3b60	4dkl	4twk	5oc9	6eu6	6qd5		5o67
3b9y	4dx5	4uc1	5och	6eyu	6qp6		5o78
3cap	4ezc	4wis	5oge	6fv7	6qq5		5o79
3d31	4fbz	4x5m	5oyb	6gyh	6qti		5t4y
3dh4	4g1u	4xu4	5sv9	6h59	6r72		5xdo
3fi1	4gpo	4ymu	5sy1	6hcy	6rtc		6ehb
3hd6	4h1d	4yzf	5t0o	6i1z	6rv2		6ehf
3hfx	4j72	5a1s	5tqq	6iql	6rvx		6ene
3j08	4j7c	5a43	5uen	6is6	6s3k		6eus
3k3f	4j9u	5aex	5uld	6iu3	6su3		6hcp
3l1l	4jkv	5ah3	5v6p	6iyx	6tqe		6sln
3m73	4jr8	5b57	5vrf	6iz4	6v1q		6ucu
3nd0	4k0e	5c78	5wue	6jbh	6vja		6v78

2.4 | Residue composition

For a particular residue type *i* in a structural region *j* of a single homomeric structure, the residue composition ($RC_{i,j}$) is defined as follows:

$$RC_{i,j} = 100 \times \frac{N_{i,j}}{\sum_{k=1}^{n_j} N_{k,j}} \quad (1)$$

where, $N_{i,j}$ defines the number of residue type *i* appearing in region *j*, n_j describes the number of the residue types (*k*) appearing in region *j* and $N_{k,j}$ denotes the total count of all residue types (*k*) found in region *j*. Therefore, RC calculates the percentage of occurrence of a particular amino acid type in a specific structural location with respect to the other residue types.

TABLE 2 Total number of the different amino acid types located at the interior, the surface, the interface and its three structural regions (in gray) in alpha-helical and beta-barrel membrane proteins

	Alpha-helical						Beta-barrel					
Residue type	Interior	Surface	Interface	Core	Rim	Support	Interior	Surface	Interface	Core	Rim	Support
ALA	12 220	5078	2544	1178	440	926	1954	846	644	368	82	194
ARG	2435	4056	1713	532	842	339	767	396	405	79	51	275
ASN	2706	2710	1230	485	465	280	1351	1045	560	292	103	165
ASP	2421	3179	1040	339	532	169	1089	1586	577	192	234	151
CYS	1597	497	302	105	68	129	16	12	2	2	0	0
GLN	2187	2576	1198	473	449	276	934	422	302	69	88	145
GLU	2497	4082	1347	419	677	251	791	582	296	92	153	51
GLY	9846	4514	1855	741	466	648	2811	1039	895	216	122	557
HIS	1102	1457	556	197	273	86	190	113	102	73	17	12
ILE	8412	5145	2694	1210	604	880	658	476	349	266	31	52
LEU	13 621	8606	4759	2259	1193	1307	1155	986	603	405	88	110
LYS	1527	4786	1152	216	795	141	612	1158	462	99	280	83
MET	3578	1379	1062	483	225	354	323	155	141	91	14	36
PHE	6428	4043	2488	1218	615	655	667	802	588	353	105	130
PRO	3304	3627	1273	494	473	306	392	304	116	57	37	22
SER	7108	4057	1963	826	525	612	1706	611	515	139	164	212
THR	6317	3526	1733	666	450	617	1563	667	507	204	148	155
TRP	1639	1569	718	300	276	142	234	251	173	91	37	45
TYR	3665	2027	1473	679	373	421	1151	1013	455	233	57	165
VAL	10 216	5022	2775	1249	562	964	1139	966	428	203	83	142
Total	102 826	71 936	33 875	14 069	10 303	9503	19 503	13 430	8120	3524	1894	2702

2.5 | Residue frequency

For a particular residue type i in a structural region j of a single homomeric structure, the residue frequency ($RF_{i,j}$) is defined as follows:

$$RF_{i,j} = 100 \times \frac{N_{i,j}}{N_i} \quad (2)$$

where, $N_{i,j}$ defines the number of a particular amino acid i appearing in region j and N_i denotes the total count of this particular amino acid within the complete structure. Thus, RF calculates the percentage of occurrence of a particular residue type in a specific region with respect to its total count in all regions.

2.6 | Frequency/composition ratio

For a particular residue type i in a structural region j of a single homomeric structure, the ratio ($FCR_{i,j}$) of its frequency (Equation 2) and composition (Equation 1) is obtained as follows:

$$FCR_{i,j} = \frac{RF_{i,j}}{RC_{i,j}} \quad (3)$$

If $FCR_{i,j} < 1$, the residue composition is too important at this location for considering a real tendency of that residue to be there. If $FCR_{i,j} > 1$, the residue frequency at this location is not due to the residue composition and presents a real tendency to be located there.

2.7 | Residue normalized propensity

For a particular residue type i in a structural region j of a single homomeric structure, the normalized propensity is calculated using two ratios.[12] The first ratio (Equation 4) gives the residue composition at a given structural region ($p_{i,j}$) and is similar to the Equation 2:

$$p_{i,j} = \frac{N_{i,j}}{\sum_{k=1}^{n_j} N_{k,j}} \quad (4)$$

The second ratio (Equation 5) gives the residue composition in the entire structure (p_i):

$$p_i = \frac{N_i}{\sum_{k=1}^{n} N_k} \quad (5)$$

where, $N_{i,j}$ defines the number of a particular amino acid i appearing in region j and n denotes all amino acid types appearing across all regions.

Then, using these two ratios, the residue normalized propensity ($NP_{i,j}$) (Equation 6) is obtained as follows:

$$NP_{i,j} = \frac{p_{i,j}}{p_i} \quad (6)$$

If for a residue, its $NP_{i,j} > 1$, the tendency of this residue to be at this location is high. If $NP_{i,j} < 1$, then this means the opposite, as the residue avoid being in this location. If, on the other hand, $NP_{i,j} = 1$, then the presence or absence tendency of a residue in this particular position is equal.

2.8 | Neighboring residues

Residues are considered to neighbor a central residue if the distance between any of their heavy atoms is below a defined value and are located on the same subunit. Here, it has been computed for two different distances: 6 or 9 Å.

3 | RESULTS

3.1 | Secondary structures effect on amino acid distribution and frequency

3.1.1 | Mean residue composition: few differences

The number of residues per complex is higher in beta-barrel proteins and at their interfaces (1110 and 219 residues, respectively) than the alpha ones (899 residues and 146 at the interface; Tables 2 and 3). This is also the case at each structural location. However, for both categories, profiles of their composition are close at the whole-protein complex level (Figure 1A) and when decomposing it into interior, surface and interface structural regions (Figure 1B). The only and main difference concerns the percentage of hydrophobic residues which is higher in all regions of alpha-proteins, when that is more polar in beta-structures (Table 4).

Deepening the decomposition of the interface region, differences can be observed within each category and between them (Figure 1C). At the core and support regions of alpha-helical structures, hydrophobic residues are mostly observed and besides being the most important component, Leu only represents, respectively, 17.4% and 14.4% of both locations' composition. The four most polar residues especially (Glu, Asp, Arg and Lys) have higher values at the rim region than the rest of the interface, and are constituting around 25% of the rim. Looking at the core and rim regions of beta-structures, as for alpha-structures, the first region is mostly hydrophobic while the second is mostly polar.

At the core of beta-structures, a lower percentage of hydrophobic residues than the alpha-ones can be observed (48.5% and 57.2%, respectively; Table 4) and there is no residue type standing out the others. At the rim also a lower percentage of hydrophobics, almost by half, is observable (21.9% and 41.7%, respectively) and besides having their higher values at this location, Asp and Lys are also representing over 27% of its composition when, alone, above 20% of the support region of beta-structures is composed of Gly.

3.1.2 | Mean residue frequency: for each secondary structure, specific residue types

The residue composition is not enough and needs to be supplemented by the frequency for progressing into the analysis. At the surface of beta-structures, hydrophobic residues have the highest frequency compared to polar residues (38% and 34.8%) when the surface of alpha-structures tends to be polar (32.1% and 42.6%; Table 4). These tendencies are confirmed when looking at the frequencies of each residue type (Figure 2A). At the interior, it is the opposite and in both structure types, intermediate residues have the highest frequency at the interior (besides the hydroxylic residues and Gln in beta-structures) and the lowest at the surface and interface. A high frequency of Cys at the interior of alpha-structures is also observed (65.6%) besides the hydroxylic residues (50.2% for Thr and 52.4% for Ser). Except a lower frequency for the intermediate residue category at the interface of alpha-structures (14.4%), the percentage frequencies are close at the interface whatever is the secondary structure or the hydrophobicity group: respectively 19.9% for hydrophobic and 18.5% for polar residues in alpha-structures and 22.4% and 19.9% in beta-structures. Frequencies are close between residue types and only His is reaching 40% at the beta-interfaces. An absence of Cys can also be noticed there despite its rarity in beta-structures, contrary to the alpha-helical interfaces. Decomposing the interface, in regard to the whole alpha or beta-protein level, shows higher frequencies of hydrophobic residues at core and polar at rim. At the core of beta-structures, hydrophobic residues are twice more frequent than polar (13.6% and 6.5%) when that is the opposite at the rim (3.3% and 6.7%).

Looking at the frequency of each amino acid type (Figure 2B), we can notice that at the core of alpha-structures, most of the hydrophobic residues are occurring there and polar residues at the rim. Besides, while His and some polar residues have the higher frequency at the rim, there is no residue type standing out the others at the support region of alpha-proteins. In beta-structures, we also observe higher frequencies of hydrophobic residues at core but the highest frequency concerns His, one out of three is located at the interface rims (34%). At the beta-rim, Lys appears to be the only residue having a frequency over 10%. Contrasting with alpha-helical supports, where polar residues and especially Arg have higher frequencies (20%), followed by Gly (11.9%).

3.1.3 | Frequency/composition ratio: a supplement pointing specific residue types

Percentage composition is showing the amount of a residue type at a location regard to all residue types, frequency is showing the amount of a residue type at a location regard to the total amount of that residue type in the entirety of the complex. Combining them into a ratio

TABLE 3 Average number of the different amino acid type located at the interior, the surface, the interface and its structural regions (in gray) in alpha-helical and beta-barrel membrane proteins (and their associated 95% confidence interval values)

	Alpha-helical						Beta-barrel					
Residue type	Interior	Surface	Interface	Core	Rim	Support	Interior	Surface	Interface	Core	Rim	Support
ALA	53 ± 5	22 ± 2	11 ± 1	5 ± 1	2 ± 0	4 ± 1	53 ± 6	23 ± 4	17 ± 3	10 ± 2	2 ± 1	5 ± 2
ARG	10 ± 1	17 ± 2	7 ± 1	2 ± 0	4 ± 1	1 ± 0	21 ± 4	11 ± 3	11 ± 2	2 ± 1	1 ± 1	7 ± 2
ASN	12 ± 1	12 ± 1	5 ± 1	2 ± 0	2 ± 0	1 ± 0	37 ± 7	28 ± 4	15 ± 3	8 ± 2	3 ± 1	4 ± 1
ASP	10 ± 1	14 ± 2	4 ± 1	1 ± 0	2 ± 0	1 ± 0	29 ± 4	43 ± 5	16 ± 3	5 ± 1	6 ± 1	4 ± 1
CYS	7 ± 1	2 ± 0	1 ± 0	0 ± 0	0 ± 0	1 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
GLN	9 ± 1	11 ± 1	5 ± 1	2 ± 1	2 ± 0	1 ± 0	25 ± 4	11 ± 2	8 ± 2	2 ± 1	2 ± 1	4 ± 1
GLU	11 ± 1	18 ± 2	6 ± 1	2 ± 0	3 ± 1	1 ± 0	21 ± 3	16 ± 3	8 ± 2	2 ± 1	4 ± 1	1 ± 1
GLY	42 ± 4	19 ± 2	8 ± 1	3 ± 1	2 ± 0	3 ± 0	76 ± 9	28 ± 4	24 ± 4	6 ± 1	3 ± 1	15 ± 3
HIS	5 ± 1	6 ± 1	2 ± 0	1 ± 0	1 ± 0	0 ± 0	5 ± 2	3 ± 1	3 ± 1	2 ± 1	0 ± 0	0 ± 0
ILE	36 ± 4	22 ± 2	12 ± 1	5 ± 1	3 ± 0	4 ± 0	18 ± 5	13 ± 2	9 ± 2	7 ± 1	1 ± 0	1 ± 1
LEU	59 ± 6	37 ± 2	21 ± 2	10 ± 1	5 ± 1	6 ± 1	31 ± 6	27 ± 4	16 ± 2	11 ± 2	2 ± 1	3 ± 1
LYS	7 ± 1	21 ± 2	5 ± 1	1 ± 0	3 ± 0	1 ± 0	17 ± 3	31 ± 6	12 ± 2	3 ± 1	8 ± 2	2 ± 1
MET	15 ± 2	6 ± 1	5 ± 1	2 ± 0	1 ± 0	2 ± 0	9 ± 4	4 ± 1	4 ± 1	2 ± 1	0 ± 0	1 ± 0
PHE	28 ± 3	17 ± 1	11 ± 1	5 ± 1	3 ± 0	3 ± 0	18 ± 4	22 ± 3	16 ± 2	10 ± 1	3 ± 1	4 ± 1
PRO	14 ± 2	16 ± 2	5 ± 1	2 ± 0	2 ± 0	1 ± 0	11 ± 4	8 ± 3	3 ± 1	2 ± 1	1 ± 1	1 ± 0
SER	31 ± 3	17 ± 2	8 ± 1	4 ± 1	2 ± 0	3 ± 0	46 ± 8	17 ± 4	14 ± 2	4 ± 1	4 ± 1	6 ± 1
THR	27 ± 3	15 ± 1	7 ± 1	3 ± 0	2 ± 0	3 ± 0	42 ± 7	18 ± 4	14 ± 2	6 ± 1	4 ± 1	4 ± 1
TRP	7 ± 1	7 ± 1	3 ± 0	1 ± 0	1 ± 0	1 ± 0	6 ± 2	7 ± 2	5 ± 1	2 ± 1	1 ± 1	1 ± 0
TYR	16 ± 2	9 ± 1	6 ± 1	3 ± 0	2 ± 0	2 ± 0	31 ± 5	27 ± 2	12 ± 2	6 ± 1	2 ± 1	4 ± 1
VAL	44 ± 5	22 ± 2	12 ± 1	5 ± 1	2 ± 0	4 ± 1	31 ± 6	26 ± 3	12 ± 2	5 ± 1	2 ± 1	4 ± 1
Total	443 ± 42	310 ± 23	146 ± 13	61 ± 6	44 ± 4	41 ± 4	528 ± 79	363 ± 36	219 ± 24	95 ± 11	51 ± 5	73 ± 10

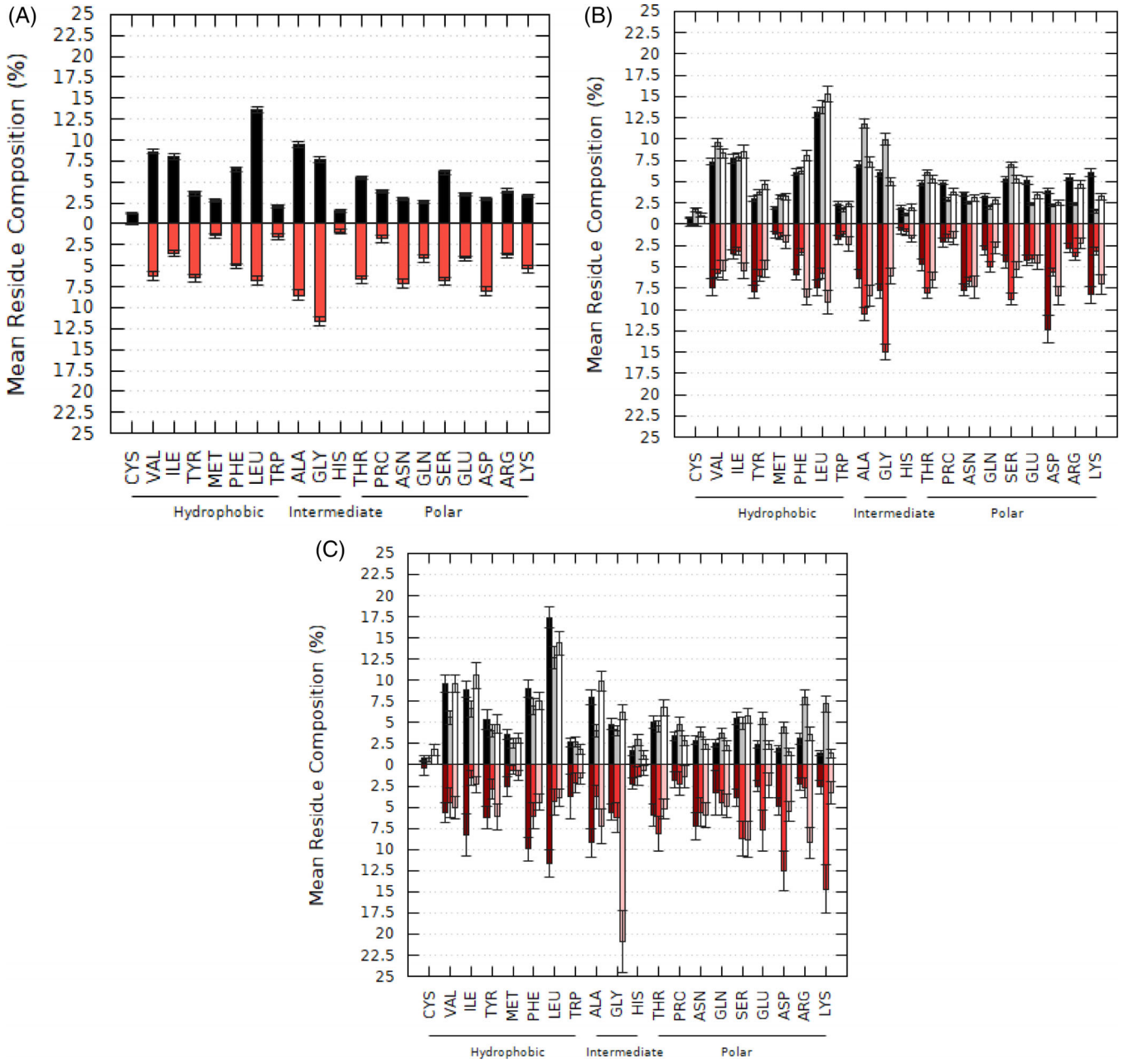

FIGURE 1 (A) Mean residue composition of alpha-helical (in black) and beta-barrel complexes (in red); (B) at the surface, interior and interface (respectively, in black at the surface, gray at the interior and white at the interface of alpha-helical structures and dark-red at the surface, red at the interior and pink at the interface of beta-barrel structures) and (C) at the structural regions of their interfaces (respectively, in black at the core, gray at the rim and white at the support of alpha-helical structures and dark-red at the core, red at the rim and pink at the support of beta-barrel structures). 95% confidence interval bars are in black [Color figure can be viewed at wileyonlinelibrary.com]

leads to the immediate identification of residue types presenting a low contribution at a location's composition but having also a high frequency there (FCR > 1), meaning that such residue types are susceptible to be crucial there. The frequency of a residue type at a location is divided by the composition in that residue type at the same location.

At the interface of alpha-structures, we can notice the important ratio of Trp (7.5 ± 1.5), His (6.2 ± 1.3), Met (4.6 ± 1.1), Tyr (4.3 ± 0.9) and polar residues - except the hydroxylics (Figure 3A). Focusing at core, Trp especially (2.4 ± 0.7) presents an important frequency despite its amount there (Figure 3B). At rim the ratio of His (1.5 ± 0.3) is notable (Figure 3C) when at the support region, we do not see a particular residue type standing out from the others (Figure 3D).

The interface of beta-structures presents the same three amino acids as alpha-proteins: ratio of His is striking (27.6 ± 10.3), followed by Met and Trp. Decomposing the interface, His appears specific to the core region (12.2 ± 4.7), and is followed by Met (Figure 3E). Besides, no residue type is standing out the others at the rim

TABLE 4 Mean percentage of residue composition (C), frequency (F), and mean residue normalized propensity (P) at each structural region of alpha-helical and beta-barrel structures and of their interfaces (in gray), according to the hydrophobicity scale category (and their associated 95% confidence interval values)

		Interior			Surface			Interface			Core			Rim			Support		
		C	*F*	*P*	*C*	*F*	*P*	*C*	*F*	*P*	*C*	*F*	*P*	*C*	*F*	*P*	*C*	*F*	*P*
Alpha-helical	Hydrophobic	48.0 ± 0.9	48.0 ± 1.6	1	42.4 ± 1.2	32.1 ± 1.0	0.9	51.4 ± 1.4	19.9 ± 1.5	1.1	57.2 ± 2.0	9.1 ± 0.7	1.2	41.7 ± 2.0	5.1 ± 0.5	0.9	53.5 ± 2.2	5.7 ± 0.5	1.2
	Intermediate	22.9 ± 0.9	56.3 ± 1.8	1.2	15.1 ± 0.7	28.9 ± 1.2	0.8	14.4 ± 0.8	14.9 ± 1.6	0.8	14.6 ± 1.2	6.3 ± 0.8	0.8	11.3 ± 1.1	3.4 ± 0.4	0.6	17.5 ± 1.5	5.2 ± 0.8	0.9
	Polar	29.4 ± 0.7	38.8 ± 1.4	0.8	42.8 ± 1.1	42.6 ± 0.8	1.2	34.5 ± 1.3	18.5 ± 1.6	1	28.4 ± 1.7	6.5 ± 0.7	0.8	47.3 ± 1.9	7.6 ± 0.7	1.4	29.3 ± 1.8	4.5 ± 0.5	0.8
Beta-barrel	Hydrophobic	26.6 ± 1.1	39.6 ± 2.7	0.9	35.4 ± 1.3	38.0 ± 2.7	1.1	34.7 ± 1.7	22.4 ± 2.6	1.3	48.5 ± 3.2	13.6 ± 1.7	1.6	21.9 ± 3.7	3.3 ± 0.7	0.7	24.7 ± 3.0	5.5 ± 0.9	0.8
	Intermediate	26.4 ± 2.2	57.3 ± 2.0	1.3	15.0 ± 1.5	23.5 ± 2.1	0.7	19.4 ± 1.9	19.2 ± 2.5	0.7	17.0 ± 2.2	7.7 ± 1.4	0.8	11.4 ± 2.7	2.5 ± 0.6	0.5	28.7 ± 3.5	9.0 ± 1.3	1.4
	Polar	47.0 ± 2.5	45.3 ± 1.8	1	49.7 ± 1.5	34.8 ± 2.3	1	45.9 ± 1.8	19.9 ± 2.6	1	34.5 ± 2.5	6.5 ± 1.0	0.7	66.7 ± 4.6	6.7 ± 0.8	1.4	46.6 ± 3.7	6.7 ± 1.0	1

(Figure 3F) and Arg appears specific to the support (1.8 ± 0.3), followed by Met (2.8 ± 1.3), Trp (2.8 ± 1.3) (Figure 3G).

3.2 | Mean normalized propensity: confirmation and supplementary specificities

Corresponding to the percentage composition of a residue type at a location, divided by the percentage of that residue in whole-structure, normalized propensity gives us the tendency of an amino acid type to occur at a location.

No important differences are observed between alpha-helical and beta-barrel interior, surface and interface regions, when comparing propensities of each amino acid category (Table 4). Looking at the residue type level, at the surface of alpha-structures (Figure 4A), polar residues and Trp have higher propensity values, while Cys and hydroxylic residues tends to occur at the interior for contributing to the helical interactions and structural stability. At the interface, few hydrophobic residues (Tyr, Phe and Trp), the most polar residues (Arg and Lys) and His as the only amino acid for the intermediate group present the higher propensity values. At the surface of beta-structures Asp and Lys have the highest propensities to be located there. At the interior, Ala, Gly, Gln and Ser present a higher values and tendencies to be there. Looking at the interface the important propensity of His is immediately observable, followed by Lys and Phe.

Decomposing the interface, specific residue types can be related to each of its structural regions (Figure 4B). Alpha-helical and beta-barrel cores have mostly hydrophobic residues having high propensity values but when Tyr and Phe have higher values in alpha-structures, His is high alone at beta-cores with a strikingly high propensity for that location. At beta-rims, when the most polar residues higher propensity in both categories, His appears specific to alpha-helical and Lys in beta. At the support region of alpha-helical interfaces, Ala and Ile have the highest propensity, followed by Cys and Thr contributing to the intra-chain structural stability. When at support region of beta-structures, Arg and Gly present higher propensity values.

4 | DISCUSSION

Looking at the overall structure and when decomposing it into surface, interior and interface regions, including its sub-structural locations, beta-barrel structures appear to be more polar than alpha-helical proteins (Table 3 and Figure 1). Moreover, the difference in hydrophobicity between each alpha-helical and beta-barrel structural regions separately, or considering the entirety of the structures, shows the hydrophilic aspect of beta-barrel structures compared to the alpha-helical proteins (Table 5) and due to its permanent contact with the solvent, rims present a more hydrophilic profile than interface cores for both categories, which is suitable to the location of polar residues and certainly to their potential involvement in forming interactions. It has been shown that at close distances the interactions

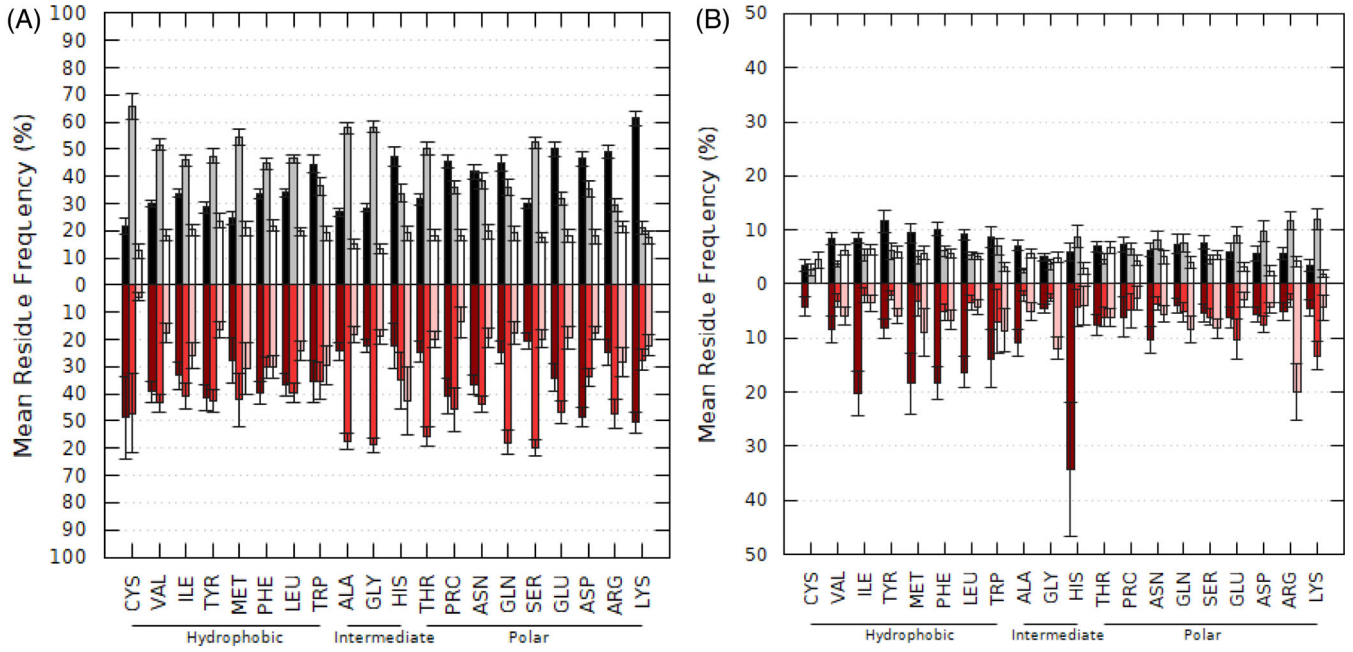

FIGURE 2 (A) Mean residue frequency at the surface, interior and interface of alpha-helical and beta-barrel complexes (respectively, in black at the surface, gray at the interior and white at the interface of alpha-helical structures and dark-red at the surface, red at the interior and pink at the interface of beta-barrel structures) and (B) at the structural regions of their interfaces (respectively, in black at the core, gray at the rim and white at the support of alpha-helical structures and dark-red at the core, red at the rim and pink at the support of beta-barrel structures). 95% confidence interval bars are in black [Color figure can be viewed at wileyonlinelibrary.com]

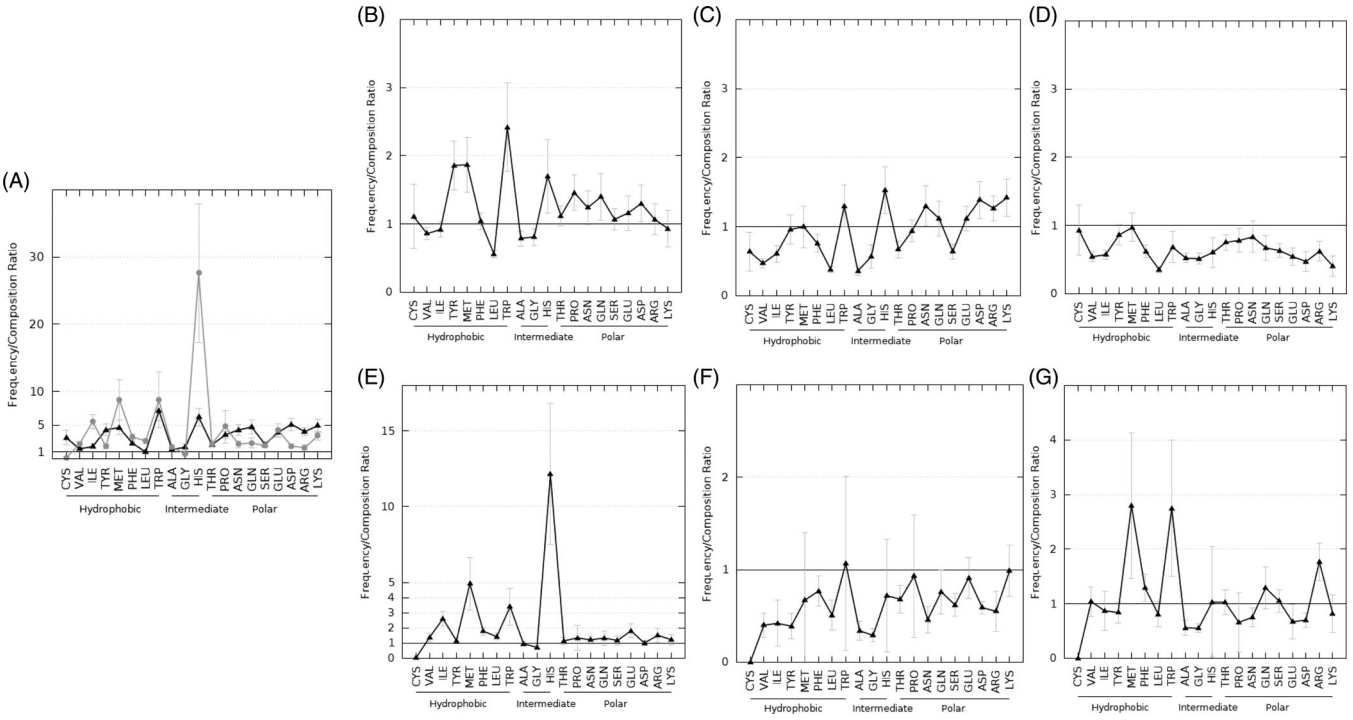

FIGURE 3 (A) Ratio of the residue frequency to the residue composition at the interface of alpha-helical (in black) and beta-barrel complexes (in gray) and at the core, rim and support of these complexes (B-D for alpha-helical core, rim and support regions. E-G for beta-barrel complexes). 95% confidence interval bars are in light gray

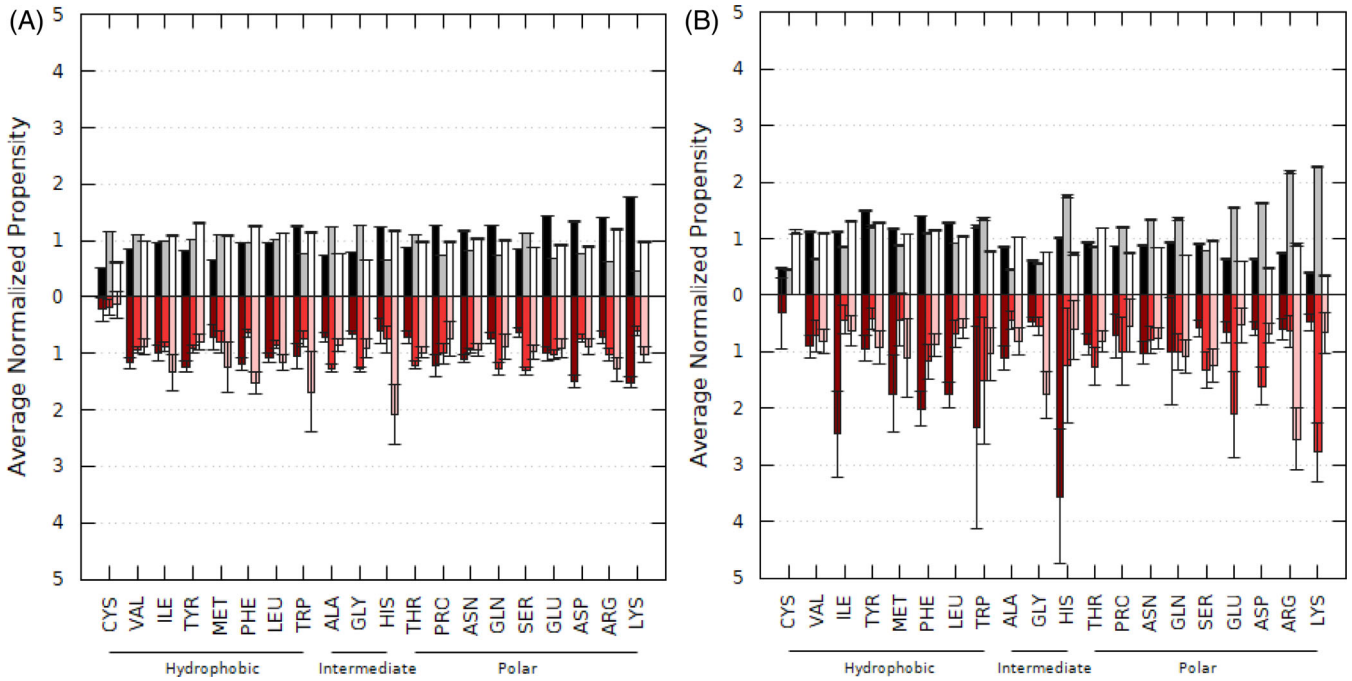

FIGURE 4 (A) Mean normalized propensity at the surface, interior and interface of alpha-helical and beta-barrel complexes (respectively, in black at the surface, gray at the interior and white at the interface of alpha-helical structures and dark-red at the surface, red at the interior and pink at the interface of beta-barrel structures) and (B) at the structural regions of their interfaces (respectively, in black at the core, gray at the rim and white at the support of alpha-helical structures and dark-red at the core, red at the rim and pink at the support of beta-barrel structures). 95% confidence interval bars are in black [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 5 Mean hydrophobicity scales of alpha-helical and beta-barrel homomers and at their different structural locations

	Hydrophobicity scale	
	Alpha-helical	Beta-barrel
Overall structure	0.57	−0.51
Interior	1.00	−0.45
Surface	−0.02	−0.65
Interface	0.32	−0.22
Outer membrane part of the surface	−1.29	−1.70
Transmembrane part of the surface	2.10	1.04
Core	0.91	0.39
Rim	−0.47	−1.35
Support	0.98	−0.80

between pairs of polar residues are predominant when hydrophobic interactions are more important at longer distances.[30]

This difference in hydrophobicity can be explained by the folding process itself, when alpha-helices are tightly packed, creating an enclosed part and, therefore, intra-protein interactions, beta-sheets are forming the layer of an almost-cylindric barrel maintained by H-bonds. Studies enlightened that the folding process of membrane alpha-proteins is not due to nonspecific effects but must be driven by specific interactions such as close packing, salt-bridge, and hydrogen bond formation.[31] In beta-barrel proteins, hydrophobic residues are oriented into the interior of the barrel, forming a hydrophobic core. Polar residues are oriented toward the outside of the barrel, on the solvent-exposed surface—complete reviews are available about the folding and insertion of beta-barrel membrane proteins in lipid bilayers.[32-33] In both secondary structure types, the average number of neighboring residues located at the core region is almost the double than at the rim region (Table 6), suggesting a more important packing, and it would be interesting to deepen this through the identification of potential specific hydrogen bonding pairs or salt bridges in interface rims and cores, taking into account potential pKa shifts.

The necessity to find new features or scores to calculate is often expressed for improving the protein-protein interaction prediction tools[34] but for observing specific and deepened differences, it is also necessary to consider the interface, not as a uniform entity, but as a complex entity needing to be decomposed. Therefore, it has to be taken in account when studying protein-protein interaction. In this study, important differences between the secondary structure types can be noticed after decomposing the interface. It can be seen that the core of alpha-protein interfaces is characterized by Tyr, Met and Phe - when looking at the propensity values - or Trp - when looking at the frequency/composition ratio. At their rim, besides the four most polar residues, His especially seems to be a specific marker when at their support region, Cys and aliphatic residues appear as characterizing that region without presenting any striking tendencies. At the beta-core, His especially is a striking marker, followed by Met and Ile

TABLE 6 Mean number of residues neighboring a central residue in alpha-helical and beta-barrel interfaces and its structural locations (in gray) using a distance of 6 and 9 Å

	Distance ($Å^2$)	Mean number of neighboring residues			
		Interface	Core	Rim	Support
Alpha-helical	6	0.95	0.8	0.5	1.7
	9	4.1	3.7	2.8	5.9
Beta-barrel	6	1	0.9	0.6	1.5
	9	4.2	3.7	2.9	5.9

when at the rim and support, Lys and Arg appear, respectively, as the main markers.

His is an aromatic residue carrying an imidazole ring and appears crucial at the rim of alpha-helical and core of beta-barrel interfaces. It can interact by its rings with nonpolar and aromatic groups, as a heteroaromatic moiety, but also participate in hydrogen bonds by its heteroatoms. Besides, it can also be involved in salt-bridges with acidic groups, depending on the protonation state.[35] His side chain features a non-negligible protonation probability at physiological pH and, due to a permanent contact of the interface rims with the solvent, and can justify its appearance. At the core of beta-barrel interfaces, which is hydrophobic compared to the rim and completely buried after complexation, His appears also crucial but its imidazole side chain may be affected by the hydrophobic environment entailing a pKa shift leading to deprotonation.[36-38]

The remaining intermediate residues present higher percentage of composition than His, especially at the interior and support regions (Figure 1). When almost 10% of the alpha-helical support regions is composed of Ala, in beta-structures, Gly alone represents around 20% of the residues and has one of the most important propensity values at the support region. A past study noted the overrepresentation of Ala and Gly in TM interfaces compared to the soluble ones, especially Gly, due to the favorable hydrogen bonding configuration of these residues in alpha-helices notably and hypothesized that the flat interfaces formed by the packing of beta-sheets also constrain the amino acids at the interface to be small as well as hydrophobic.[39-40]

In proteins, another study points the π-π stacking interaction between neutral histidine and aromatic amino acids (Phe, Tyr, and Trp) are larger than the van der Waals energies (from −3.0 to −4.0 kcal/mol).[41] Over 50% of Phe, Tyr and Trp residues in a protein-protein interface are involved π-π interactions, showing their importance as hot-spots especially in the aim of stabilizing the structure.[42-43] Determining the precise position of the transmembrane region in lipid membranes, aromatics contribute to lipid-protein or protein-protein interactions and the side chain orientation affects the stabilizing effect of π-π stacking.[44-45] Propensity values of most of the aromatics at the interface of membrane proteins are within the highest values, especially at the core and rim (Figure 4). It is known that aromatics play a role of anchors to stabilize the TM regions through interactions with the lipid head groups or other TM segments and they are known to be over-represented near the ends of transmembrane helices.[46-47] Also, in alpha-helical complexes, Phe and Tyr are the only aromatics having a higher propensity at the interface than the surface while Trp seems to present similar propensity value for both locations. From this observation, Trp can be observed at the interface due only to a homogeneous distribution at the surface of these proteins or to a geometrical effect like the presence of an aromatic belt.[22]

An absence of Cys can be noticed at the interface of beta-barrel complexes and Cys is almost absent in these structures, contrary to the alpha-helical complexes where disulfide-bonds have a crucial role for the structural stability and form important pairs.[48] The importance of hydroxylic residues at the interior of alpha-structures supports the importance of disulfide-bonds.

When looking at the frequency/composition ratio (Figure 3), Met appears as one of residue standing out the others at different locations of the interface. Its importance for stabilizing a protein structure with the formation of hydrophobic interactions and at long distances has been shown, when forming a motif with aromatic residues.[49] In his review, Aledo defines it as the Cinderella of the proteinogenic amino acids due to its variety of property, from its side chain, that makes of it an optimal amino-acid for protein-protein interaction and central to molecular recognition.[50] Besides Met, Phe and Trp appear to be markers of some interface regions in alpha and beta-proteins and is consistent with an observation made by Ma and Nussinov, stating these amino acids as potential targets in drug design.[51]

The present study is pointing potential residues characterizing the interface of alpha-helical and beta-barrel homomeric complexes through the different use of the residue occurrence at the different structural locations. In future work, using the same set of membrane homomers, it can be supplemented with different direction of analyses in the aim to predict protein-protein interaction of membrane proteins. The paradigm of strictly binding interfaces, for example, has been questioned in a study demonstrating that the majority of a globular protein can geometrically participate in an interaction surface.[52] This is certainly applicable to membrane proteins, even if the embedding of proteins within a lipidic membrane appears to be a major constraint, but using a set of known structures the current study presents residues appearing as specific to the interface and to its different structural locations. Moreover, it has been demonstrated, in 2009, that weakly stable strands of beta-barrel oligomers are involved in the oligomerization processes, forming the interface.[53] Using sequence information only, this study successfully predicted with 82% the protein-protein interaction interface. Thus, a combination of both approaches can be effective for building more accurate prediction tools. However, these global approaches should be supplemented by the identification of hot-spots or interface residues as done in the present study for aiming to be potentially more accurate.

5 | CONCLUSION

The analysis of protein interfaces is an area in progress and, for deepening it, usable features are actually limited. In this study, through the only use of the residue occurrence as a basis and the use of Levy's model, I aimed to bring to light residues characterizing the interface of

alpha-helical or beta-barrel membrane proteins, after decomposing it. The residue composition and frequency are useful but give us incomplete information when used separately. Combined into a simple ratio, this information highlights residues having high frequencies at a location despite their low occurrences, which can be an interesting and simple indicator, notably for spotting key residues at the interface. In parallel to this ratio, propensity calculations are pointing similar residues and permits me to propose different residues characterizing each interface location.

ACKNOWLEDGMENTS

This work was supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK): 1002 - Short-Term Research and Development Funding Program n°119M188. I also thank my colleague Nuray Sogunmez for her technical assistance.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Sercan Beytur https://orcid.org/0000-0002-0202-9128

REFERENCES

1. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol*. 2001;305:567-580.
2. Fagerberg L, Jonasson K, von Heijne G, Uhlen M, Berglund L. Prediction of the human membrane proteome. *Proteomics*. 2010;10:1141-1149.
3. Overington JP, Al-Lazikani B, Hopkins AL. How many drug targets are there? *Nat Rev Drug Discov*. 2006;5:993-996.
4. Yin H, Flynn AD. Drugging membrane protein interactions. *Annu Rev Biomed Eng*. 2016;18:51-76.
5. Sanders CR, Nagy JK. Misfolding of membrane proteins in health and disease: the lady or the tiger? *Curr Opin Struct Biol*. 2000;10(4):438-442.
6. Ng DP, Poulsen BE, Deber CM. Membrane protein misassembly in disease. *Biochim Biophys Acta*. 1818;2012:1115-1122.
7. Kennedy SJ. Structures of membrane proteins. *J Membr Biol*. 1978;42:265.
8. Almèn MS, Nordström KJV, Fredriksson R, Schiöth HB. Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC Biol*. 2009;7:50.
9. Freeman T, Landry S, Wimley W. The prediction and characterization of YshA, an unknown outer-membrane protein from *Salmonella typhimurium*. *Biochim Biophys Acta*. 1808;2001:287-297.
10. Wimley WC. The versatile beta-barrel membrane protein. *Curr Opin Struct Biol*. 2003;13:404-411.
11. Lee B, Richards M. The interpretation of protein structures: estimation of static accessibility. *J Mol Biol*. 1971;55:379-400.
12. Jones S, Thornton JM. Principles of protein-protein interactions. *Proc Natl Acad Sci U S A*. 1996;93(1):13-20.
13. Bogan AA, Thorn KS. Anatomy of hot spots in protein interfaces. *J Mol Biol*. 1998;280:1-9.
14. Lo Conte L, Chothia C, Janin J. The atomic structure of protein-protein recognition sites. *J Mol Biol*. 1999;285(5):2177-2198.
15. Bahadur RP, Chakrabarti P, Rodier F, Janin J. Dissecting subunit interfaces in homodimeric proteins. *Proteins*. 2003;53:708-719.
16. Chakrabarti P, Janin J. Dissecting protein-protein recognition sites. *Proteins*. 2002;47:334-343.
17. Janin J, Rodier F, Chakrabarti P, Bahadur RP. Macromolecular recognition in the protein data bank. *Acta Crystallogr Sect D*. 2007;63:1-8.
18. Guharoy M, Chakrabarti P. Conservation and relative importance of residues across protein-protein interfaces. *Proc Natl Acad Sci U S A*. 2005;102:15447-15452.
19. Levy ED. A simple definition of structural regions in proteins and its use in analyzing Interface evolution. *J Mol Biol*. 2010;403:660-670.
20. Yan C, Wu F, Jernigan RL, Dobbs D, Honavar V. Characterization of protein-protein interfaces. *Protein J*. 2008;27(1):59-70.
21. Moreira IS, Martins JM, Ramos RM, Fernandes PA, Ramos MJ. Understanding the importance of the aromatic amino-acid residues as hotsports. *Biochim Biophys Acta*. 1834;2013:404-414.
22. Ulmschneider MB, Sansom MSP. Amino acid distributions in integral membrane protein structures. *Biochim Biophys Acta*. 2001;1512(1):1-14.
23. Gaines JC, Acebes S, Virrueta S, Butler M, Regan L, O'Hern CS. Comparing side chain packing in soluble proteins, protein-protein interfaces, and transmembrane proteins. *Proteins*. 2018;86(5):581-591.
24. Cymer F, Veerappan A, Schneider D. Transmembrane helix-helix interactions are modulated by the sequence context and by lipid bilayer properties. *Biochim Biophys Acta*. 2012;1818(4):963-973.
25. Laine E, Carbone A. Local geometry and evolutionary conservation of protein surfaces reveal the multiple recognition patches in protein-protein interactions. *PLoS Comput Biol*. 2015;11(12):e1004580.
26. Lomize MA, Pogozheva ID, Joo H, Mosberg HI, Lomize AL. OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res*. 2012;40:D370-D376.
27. Lomize AL, Pogozheva ID, Lomize MA, Mosberg HI. Positioning of proteins in membranes: a computational approach. *Protein Sci*. 2006;15(6):1318-1333.
28. Moret MA, Zebende GF. Amino acid hydrophobicity and accessible surface area. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2007;75(1):011920.
29. Tien MZ, Meyer AG, Sydykova DK, Spielman SJ, Wilke CO. Maximum allowed solvent accessibilites of residues in proteins. *PLoS One*. 2013;8(11):e80635.
30. Bahar I, Jernigan RL. Inter-residue potentials in globular proteins and the dominance of highly specific hydrophilic interactions at close separation. *J Mol Biol*. 1997;266:195-214.
31. Ben-Tal N, Honig B. Helix-helix interactions in lipid bilayers. *Biophys J*. 1996;71:3046-3050.
32. Kleinschmidt JH. Folding of β-barrel membrane proteins in lipid bilayers: unassisted and assisted folding and insertion. *Biochim Biophys Acta*. 1848;2015:1927-1943.
33. Chatuverdi D, Mahalakshmi R. Transmembrane β-barrels: evolution, folding and energetics. *Biochim Biophys Acta*. 1859;2017:2467-2482.
34. Xue LC, Dobbs D, Bonvin AM, Honavar V. Computational prediction of protein interfaces: a review of data driven methods. *FEBS Lett*. 2015;589(23):3516-3526.
35. Bhattacharyya R, Saha RP, Samanta U, Chakrabarti P. Geometry of interaction of the histidine ring with other planar and basic residues. *J Proteome Res*. 2003;2:255-263.
36. Rabenstein B, Knapp EW. Calculated pH-dependant population and protonation of carbon-monoxy-myoglobin conformers. *Biophys J*. 2001;80:1141-1150.
37. Panahi A, Brooks CL. Membrane environment modulates the p*Ka* values of transmembrane helices. *J Phys Chem B*. 2015;119:4601-4607.
38. Martfeld AN, Greathouse DV, Koeppe RE. Ionization properties of histidine residues in the lipid bilayer membrane environment. *J Biol Chem*. 2016;291(36):19146-19156.
39. Senes A, Ubarretxena-Belandia I, Engelman DM. The Cα–H···O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. *Proc Natl Acad Sci U S A*. 2001;98(16):9056-9061.

40. Duarte JM, Biyani N, Baskaran K, Capitani G. An analysis of oligomerization interfaces in transmembrane proteins. *BMC Struct Biol*. 2013;13:21.
41. Liao SM, Du QS, Meng JZ, Pang ZW, Huang RB. The multiplel roles of histidine in protein interactions. *Chem Cent J*. 2013;7(1):44.
42. Tewari AK, Dubey R. Emerging trends in molecular recognition: utility of weak aromatic interactions. *Bioorg Med Chem*. 2008;16(1):126-143.
43. Burley SK, Petsko GA. Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science*. 1985;229(4708):23-28.
44. Schlebach JP, Sanders CR. The safety dance: biophysics of membrane protein folding and misfolding in a cellular context. *Q Rev Biophys*. 2015;48:1-34.
45. Makwana KM, Mahalakshmi R. Implications of aromatic-aromatic interactions: from protein structures to peptide models. *Protein Sci*. 2015;24(12):1920-1933.
46. White SH, Wimley WC. Membrane protein folding and stability: physical principles. *Annu Rev Biophys Biomol Struct*. 1991;28:319-365.
47. Cymer F, von Heijne G, White SH. Mechanisms of integral membrane protein insertion and folding. *J Mol Biol*. 2015;427:999-1022.
48. Glaser F, Steinberg DM, Vakser IA, Ben-Tal N. Residue frequencies and pairing preferences at protein-protein interfaces. *Proteins*. 2001;43(2):89-102.
49. Valley CC, Cembran A, Perlmutter JD, et al. The methionine-aromatic motif plays a unique role in stabilizing protein structure. *J Biol Chem*. 2012;287(42):34979-34991.
50. Aledo JC. Methionine in proteins: the Cinderella of the proteinogenic amino acids. *Protein Sci*. 2019;28(10):1785-1796.
51. Ma B, Nussinov R. Trp/Met/Phe hot spots in protein-protein interactions: potential targets in drug design. *Curr Top Med Chem*. 2007;7:999-1005.
52. Tonddast-Navaei S, Skolnick J. Are protein-protein interfaces special regions on a protein's surface? *J Chem Phys*. 2015;143(24):243149.
53. Naveed H, Jackups R Jr, Liang J. Predicting weakly stable regions, oligomerization state, and protein-protein interfaces in transmembrane domains of outer membrane proteins. *Proc Natl Acad Sci USA*. 2009;106(31):12735-12740.

How to cite this article: Beytur S. Marker residue types at the structural regions of transmembrane alpha-helical and beta-barrel interfaces. *Proteins*. 2021;1–13. https://doi.org/10.1002/prot.26087