



**Kompleks Proteinlerin Farklı Bölgelerindeki Amino Asit
Dağılımı ve Etkileşimini Etkileyen Çevresel ve Yapısal
Faktörlerin İncelenmesi:
Hücre Zarı Ortamı, Sekonder Yapı ve Kompleks Yapıdaki
Monomer Sayısı**

Program Kodu: 1002

Proje No: 119M188

Proje Yürütücüsü:
Prof. Dr. E. Demet AKDOĞAN

Bursiyer:
Sercan BEYTUR

EKİM 2020



İSTANBUL

ÖNSÖZ

Bu projede protein kompleks yapılarıdaki monomerik birim sayısı, sekonder yapı gibi yapısal farklılıklarla, hücre içi ve dışı çevresel farklılıkların kompleks yapının arayüzey özelliklerini nasıl etkilediği istatistiksel analizlerle ortaya çıkartılmıştır. Bu proje, bir sonraki geliştirilmesi planlanan arayüzey tahmin algoritmasında kullanılacak kritik öneme sahip bulgular sağlamıştır. Kompleksi oluşturan monomerik birim sayısı, sekonder yapının türü ve hücre içi/dışı üç temel faktörün arayüzeydeki etkileşim türlerini ve derecelerini ciddi biçimde etkilemesinin yanı sıra, amino asitlerin bileşim yüzdeleri, SASA değerleri, arayüzey dışında kalan bölgelerdeki bulunma sıklıklarını da etkilediği gösterilmiştir. Dolayısıyla, bu projenin çıktıları, arayüzey tahmin algoritmaları geliştirirken kompleks sistemleri kendi yapısal/çevresel farklılığını temsil eden kategorilerde bağımsız ele almanın ve farklı parametrelerle temsil etmenin önemini ve gerekliliğini açıkça ortaya koymuştur.

TÜBİTAK 1002 destekli 119M188 no'lu bu proje, Kadir Has Üniversite bünyesi altında, 1 Eylül 2019 – 1 Eylül 2020 tarihleri arasında yürütülmüştür.



İÇİNDEKİLER

Önsöz	2
Şekil Listesi.....	4
Özet.....	7
Abstract.....	8
1. Giriş.....	9
2. Literatür Özeti	9
3. Gereç ve Yöntem	11
3.1. Membran protein kompleks yapılar ve hazırlanması.....	11
3.2. Yüzey alana (SASA) göre protein kompleksinin farklı bölgelere ayrılması.....	13
3.3. Farklı bölgeler için amino asit eğilimlerinin hesaplanması.....	15
3.4. Arayüzeyde etkileşen rezidülerin kontakt derecesi ve görülme sıklıkları Hesaplanması	17
4. Bulgular ve Tartışma.....	18
4.1. Alfa-sarmal membran protein komplekslerde ortalama amino asit miktarı.....	18
4.2. Alfa-sarmal membran protein komplekslerde ortalama SASA/rASA değeri.....	22
4.3. Alfa-sarmal membran protein komplekslerdeki rezidülerin ortalama eğilim (propansite) değerleri.....	24
4.4. Alfa-sarmal membran protein komplekslerdeki arayüzey bölgelerde kontakt rezidü çiftlerinin belirlenmesi.....	28
4.5. Homotrimerik ve Homodimerik Alfa-sarmal membran protein komplekslerin karşılaştırmalı analizi.....	31
4.6. Homotetramerik, Homotrimerik ve Homodimerik Alfa-sarmal membran protein komplekslerin karşılaştırmalı analizi.....	37
4.7. Beta-silindir ve Alfa-sarmal membran protein komplekslerin karşılaştırmalı analizi....	42
5. Sonuç.....	48
6. Referanslar.....	50

ŞEKİL LİSTESİ

- Şekil 1. OPM veri bankasından elde edilen alfa-sarmal ve bata-tabaka sekonder yapıdaki membran protein komplekslerinin yapısal uygunluklarının test edilmesi ve tekrar eden (redundant data) yapıların elenme işlemi için izlenen iş akışı.....12
- Şekil 2. OPM veribankasında tetramerik kategorisinde olup ta heterodimerik bir arayüzey oluşumu gösteren kompleks yapılar (a) PDB id: 1qlc ve (b) PDB id: 6rko.....13
- Şekil 3. Hücre zarı alfa-sarmal homodimerik bir yapı olan çinko-taşıyıcı YiiP (PDB id:2qfi) için (a) cartoon gösterimi ve (b) yüzey alan üzerinde iç (sarı), dış yüzey (koyu mavi), destek arayüzey (turuncu), çekirdek arayüzey (yeşil) ve kenar arayüzey (kırmızı) bölgeler.....14
- Şekil 4. 168 alfa-sarmal homodimerik protein kompleksinde rezidülerin (a) toplam sayısına ve (b) farklı bölgelerindeki dağılım grafikleri. (c) hidrofobisite sabitine (Eisenberg vd. (1984)) göre ortalama rezidü bileşimi.....20
- Şekil 5. 168 alfa-sarmal homodimerik protein kompleksinde hidrofobisite sabitine (Eisenberg vd. (1984)) göre farklı bölgelerdeki ortalama rezidü bileşimi. (a) İç bölge, (b) yüzey bölge, (c) çekirdek arayüzey, (d) destek arayüzey ve (e) kenar arayüzey.....21
- Şekil 6. 168 alfa-sarmal homodimerik protein kompleksinde (a) hidrofobisite sabitine (Eisenberg vça., 1984) göre ortalama SASA değeri. (b) rASA dağılım grafiği.....22
- Şekil 7. 168 alfa-sarmal homodimerik protein kompleksinde rezidülerin farklı bölgelerdeki ortalama SASA değerleri. (a) yüzey, (b) kenar arayüzey, (c) çekirdek arayüzey ve (d) destek arayüzey.....24
- Şekil 8. 168 alfa-sarmal homodimerik protein kompleksinde rezidülerin farklı bölgeler için propansite (eğilim) değerleri. (a) yüzey, (b) kenar arayüzey, (c) çekirdek arayüzey ve (d) destek arayüzey.....25
- Şekil 9. 168 alfa-sarmal homodimerik protein kompleksinde hidrofobisite sabitine (Eisenberg vd. (1984)) ve tüm yüzeye göre (a) arayüzey ve (b) kenal arayüzey propansite değerleri.....26
- Şekil 10. Rezidü türlerinin lipid ve periferik ortamdaki arayüzeylerdeki kontakt derecesi..... 27

- Şekil 11. (a),(c) lipid ve (b),(d) periferik ortamlardaki arayüzeyde gözlenen (*observed*) kontakt frekansı KP_{ij} ve skor değerleri.....30
- Şekil 12. Dimerik ve trimerik alfa sarmal yapılarda rezidülerin ortalama (a) bileşim yüzdeleri ve (b) SASA değerleri.....32
- Şekil 13. 60 alfa-sarmal homotrimerik protein kompleksinde rezidülerin farklı bölgelerdeki ortalama SASA değerleri. (a) yüzey, (b) kenar arayüzey, (c) çekirdek arayüzey ve (d) destek arayüzey.....33
- Şekil 14. 60 alfa-sarmal homotrimerik protein kompleksinde rezidülerin farklı bölgeler için propansite (eğilim) değerleri. (a) arayüzey / yüzey, (b) kenar arayüzey / yüzey.....34
- Şekil 15. Dimerik ve trimerik yapılarda, rezidülerin (a) lipid ve (b) periferik ortamda bulunan arayüzey bölgelerdeki kontakt dereceleri.....35
- Şekil 16. Alfa-sarmal trimerik yapılarda (a),(c) lipid ve (b),(d) periferik ortamlardaki arayüzeyde gözlenen (*observed*) kontakt frekansı KP_{ij} ve skor değerleri.....36
- Şekil 17. Dimerik, trimerik ve tetramerik alfa sarmal yapılarda rezidülerin ortalama (a) bileşim yüzdeleri ve (b) SASA değerleri.....38
- Şekil 18. 99 alfa-sarmal homotetramerik protein kompleksinde rezidülerin farklı bölgelerdeki ortalama SASA değerleri. (a) yüzey, (b) kenar arayüzey, (c) çekirdek arayüzey ve (d) destek arayüzey.....39
- Şekil 19. 99 alfa-sarmal homotetramerik protein kompleksinde rezidülerin farklı bölgeler için propansite (eğilim) değerlerinin dimerik yapılardaki propansite değerleriyle karşılaştırılması. (a) arayüzey / yüzey, (b) kenar arayüzey / yüzey.....40
- Şekil 20. Dimerik, trimerik ve tetramerik yapılarda, rezidülerin (a) lipid ve (b) periferik ortamda bulunan arayüzey bölgelerdeki kontakt dereceleri.41
- Şekil 21. Tetramerik yapılarda (a),(c) lipid ve (b),(d) periferik ortamlardaki arayüzeyde gözlenen (*observed*) kontakt frekansı KP_{ij} ve skor değerleri.....42

Şekil 22. Alfa-sarmal ve beta-silindir yapılarda rezidülerin ortalama (a) bileşim yüzdeleri ve (b) SASA değerleri.....43

Şekil 23. 39 beta-silindir kompleks yapıdaki rezidülerin farklı bölgelerdeki ortalama SASA değerlerinin alfa-sarmal yapıdakilerle karşılaştırılması. (a) yüzey, (b) kenar arayüzey, (c) çekirdek arayüzey ve (d) destek arayüzey.....45

Şekil 24. 39 beta-silindir kompleks yapıdaki rezidülerin farklı bölgeler için propansite (eğilim) değerlerinin alfa-sarmal yapılardaki propansite değerleriyle karşılaştırılması. (a) arayüzey / yüzey, (b) kenar arayüzey / yüzey.....45

Şekil 25. Alfa-sarmal ve beta-silindir yapılarda rezidülerin (a) lipid ve (b) periferik ortamda bulunan arayüzey bölgelerdeki kontakt dereceleri.....46

Şekil 26. Beta-silindir yapılarda (a),(c) lipid ve (b),(d) periferik ortamlardaki arayüzeyde gözlenen (*observed*) kontakt frekansı KP_{ij} ve skor değerleri.....47



ÖZET

Bu projede kompleks yapılardaki amino asit türleri, dağılımları, görülme sıklıkları, çözücü erişilir yüzey alanı (SASA), arayüzeyde etkileşen amino asit çiftlerinin görülme sıklıkları ve kontakt derecelerini etkileyen yapısal ve çevresel faktörler ele alınarak, bu değişimler ayrıntılı bir şekilde ortaya konmuştur. Bu faktörlerden biri olan kompleks yapının içinde bulunduğu ortam ele alındığında, hücre zarı proteinin hücre dışına taşan periferik ortamdaki arayüzey bölgesinin, hücre içi lipid ortamdaki arayüzey bölgesinden her yönden farklı olduğu gösterilmiştir. Bir diğer faktör de kompleks yapıyı oluşturan monomerik birim sayısıdır. Alfa-sarmal kategorisinde yer alan dimerik, trimerik ve tetramerik yapılar amino asitlerin tür ve dağılımları, SASA, propansite ve arayüzey kontakt dereceleri, frekansları ve skor değerlerine göre karşılaştırıldığında her yönden dikkat çekici farklılıklar ortaya çıkmıştır. Son olarak sekonder yapıdaki değişikliğin ne derece etkili bir faktör olacağını göstermek amacıyla, beta-silindir yapılar kendi içlerinde incelenmiş ve en büyük farklılık ta bu kategoride ortaya çıkmıştır. Bu bulgular doğrultusunda bir sonraki daha kapsamlı proje önerimizde geliştirilmesi planlanan arayüzey tahmin algoritması için kritik öneme sahip farklılıklar ortaya çıktığından, her kategorinin kendi içinde ele alınması gerektiği ve faktörlerden en fazla hücre içi/dışı ortam ile sekonder yapının dikkate alınması gerektiği vurgulanmıştır.

Anahtar kelimeler: protein-protein arayüzey, sekonder yapı, hücre zarı ortamı, lipid, periferik, monomerik birimler



ABSTRACT

In this project, various structural and environmental factors affecting the amino acid type, composition, propensity, solvent-accessible surface area (SASA), contact degree and observed frequency at the interface region in complex proteins have been investigated and highlighted. One of the factors having a significant impact on these structural features is the environment in which the complex protein is found. The interface regions which are found on either peripheric or lipid sides display completely different profiles. Another factor affecting the interaction profile in complexes is the number of monomeric units. Accordingly, all alpha-helical membrane protein complexes were further categorized as dimer, trimer and tetramer and each category was analyzed independently. As a result, significant differences were observed in almost every structural feature of the complex, such as amino acid composition, mean SASA, residue propensity for a specific region, contact degree at the interface, observed frequency of a specific pair of residue at the interface. Finally, a third factor which was the secondary structure was investigated using beta-barrel membrane proteins. This category showed the most striking differences in all features studied so far. Based on these findings, a prediction tool for interface region will be developed in the near future and the algorithm will eventually take into account all these differences, treating each protein category separately.

Keywords: protein-protein interface, secondary structure, membrane environment, lipid, peripheric, monomeric units



Giriş

Biyokimyasal düzeyde, proteinler işlevlerini nadiren yalnız başlarına yerine getirirler. Polipeptid zincirleri arası etkileşimler biyolojik olayların temelini oluşturmaktadır. Bu etkileşimlerin altında yatan moleküler mekanizmanın anlaşılması için çok sayıda monomer ve kompleks yapının üç boyutlu yapısından faydalanılmıştır. Bu sayede, monomerler arası etkileşimin çok farklı düzeylerde konformasyonel değişimleri de beraberinde getirdiği ortaya çıkartılmıştır. Bu değişimler, yan zincirlerdeki lokal konformasyonel değişimler olabilirken, daha global ölçekte bölgesel değişimler de olabilmektedir (Gerstein vd. (1994)). Bu projedeki amaç, arayüzey olarak ta adlandırılan monomerlerin etkileşim alanına özgü özellikleri ortaya çıkartmak, bu sayede daha kapsamlı bir projede geliştirilmesi planlanan kompleks yapı tahmin programı için kullanılabilir son derece yararlı bir veri kaynağı oluşturmaktır.

Hücre zarı proteinlerinin içinde buldukları lipid çevre, membran dışında işlevini sürdüren sitoplazmik proteinlerin buldukları sulu ortamdan oldukça farklıdır. Hücre zarı proteinleri çoğunlukla lipid tabaka ile etkileşim içinde olup, su moleküllerinden çok fazla etkilenmemektedirler. Lipid tabakayla temas halinde olan yüzeydeki amino asitlerin iç bölgeden daha hidrofobik olduğu bilinmektedir. Bunun yanı sıra, kompleks membran proteinlerin arayüzey bölgelerine bakıldığında tek tip sekonder yapı görülmektedir, ya tamamı heliks ya da beta-silindir olmaktadır. Tüm bu çevresel ve yapısal farklılıklar, amino asit görülme sıklığı ve dağılımının aynı olmayacağını işaret etmektedir. Buna en çarpıcı örnek Cys-Cys rezidü çiftidir. Sitoplazmik proteinlerde en fazla sayıda görülen rezidü çifti olmalarına rağmen, hücre zarı proteinlerde çok nadir rastlanmaktadır (Xu vd. (2007)). Literatürde, kompleks yapıların arayüzey bölgesinin, lipid çevre ya da sekonder yapıdan nasıl etkilendiğine dair herhangi bir sistematik çalışma bulunmamaktadır. Bu projede gerçekleştirilen çalışma bu açığı kapatmayı hedeflemektedir. Aynı zamanda elde edilecek veriler, bir sonraki projede geliştirilecek algoritma için temel oluşturacaktır.

1. Literatür Özeti

Hücre içi kimyasal süreçlerin çoğu protein kompleksleri tarafından yürütülmektedir. Bunlar arasında, hücre içi sinyalleşme, çoğalma, farklılaşma, DNA onarımı ve bağışıklık sayılabilir. Protein kompleks yapıları oluşturan arayüzeyler, son yıllarda ilaç geliştirme çalışmalarında en fazla hedef alınan bölgeler arasında yer alarak ciddiyetini korumaktadır. Bu arayüzey bölgelerin kendine özgü karakteristik özellikleri, protein-protein etkileşiminin (kısaca PPI) derecesini ve özgüllüğünü etkilemektedir.

Transmembran proteinler proteomun %30'unu temsil etmektedir ve sinyal geçişi, hücre içi ve dışına elektron taşınması, iyon/molekül iletimi ve hücrelerarası iletişim gibi çok sayıda farklı işleve sahiptirler (Jayasinghe vd. (2001)). Birçok hastalığın sebebi bu işlevlerin aksaması sonucu oluşmaktadır. Dolayısıyla, transmembran proteinlerin kompleks yapılarını ve birbirleri arasındaki etkileşimi ortaya koymak ilaç tasarım çalışmalarına önemli katkılar sağlayacaktır. Bunun yanı sıra, transmembran protein komplekslerinin, hücre içi ya da dışında serbest dolaşan protein komplekslerine göre ne tür farklı yapısal özellikler ortaya koyduğunu belirlemek, lipid çevrenin kompleks (kuaterner) yapıların oluşumundaki etkisinin anlaşılmasını mümkün kılacaktır. Literatürde, protein-protein etkileşim arayüzey bölgelerin karakteristik özelliklerini ortaya koymak için birçok deneysel ve hesaplamalı çalışma mevcut olsa da, hiç biri lipid çevrenin arayüzey bölgeye olan etkisini incelememiştir. Dolayısıyla bu projenin en özgün tarafı, farklı iki ortamdaki protein gruplarını (transmembran ve membran-dışı) inceleyerek lipid çevrenin yapısal özelliklerin oluşumundaki rolünü ortaya çıkartmak olacaktır.

Hücre zarı proteinlerinin içinde buldukları lipid çevre, membran dışında işlevini sürdüren globüler proteinlerin buldukları sulu ortamdan oldukça farklıdır. Hücre zarı proteinleri çoğunlukla lipid tabaka ile etkileşim içinde olup, su moleküllerinden çok fazla etkilenmemektedirler. Bu proteinlerin hücre zarına gömülü olan bölgesindeki amino asitlerin büyük çoğunluğunun hidrofobik olduğu bilinmektedir (Ulmschneider vd. (2005)). Bunun yanı sıra, lipid tabakayla temas halinde olan yüzeydeki amino asitlerin iç bölgeden çok daha fazla hidrofobik olduğu gösterilmiştir. Protein arayüzey etkileşimlerinde hidrofobik kuvvetlerin kritik rol oynadığı birtakım çalışmalarda gösterilmiş olsa da (Young vça., 1994; Berchanski vd. (2004); Yan vd. (2008); Glaser vd. (2001)), hidrofilik özelliğin daha baskın olduğu Ben-Naima'nın 2006 yılındaki çalışmasında öne sürülmüş, "sıcak-nokta" ("hot-spot") rezidülerinin kompleks oluşumundaki baskın rolü ortaya çıkartılmıştır (Ben-Naima (2006); Keskin vd. (2005)). 2011 yılında Kysilka ve Vondrasek'a ait bir başka çalışmada, arayüzey bölgedeki yan zincirlerin intramoleküler bir yapılanma gösterdiği ortaya konmuştur. Ayrıca, amino asit dağılımı ve monomerler arası kontakt haldeki amino asit çiftleri incelendiğinde arayüzey bölgenin diğer bölgelere göre oldukça farklı bir profil sergilediği gösterilmiştir (Kysilka ve Vondrasek (2012)).

Transmembran proteinlerde amino asit dağılımını ortaya koyan bir çalışmada, hidrofobik rezidülerin (ör. *Phe*, *Ile*, *Leu*, *Val*, *Ala*) beta-silindir sekonder yapısına sahip proteinlerde proteinin dış yüzeyinde kalarak lipid tabaka ile etkileşim içinde bulduklarını, ancak heliks yapıdaki proteinlerde herhangi bir bölgeyi tercih etmedikleri gözlemlenmiştir (Ulmschneider ve Sansom (2001)). Bu sonuç, sekonder yapının proteinlerdeki amino asit dağılımına olan

etkisini açıkça göstermektedir. Dolayısıyla, arayüzey bölgeler incelenirken, sekonder yapıları farklı olan protein kompleksleri farklı kategorilerde ele alınmalıdır.

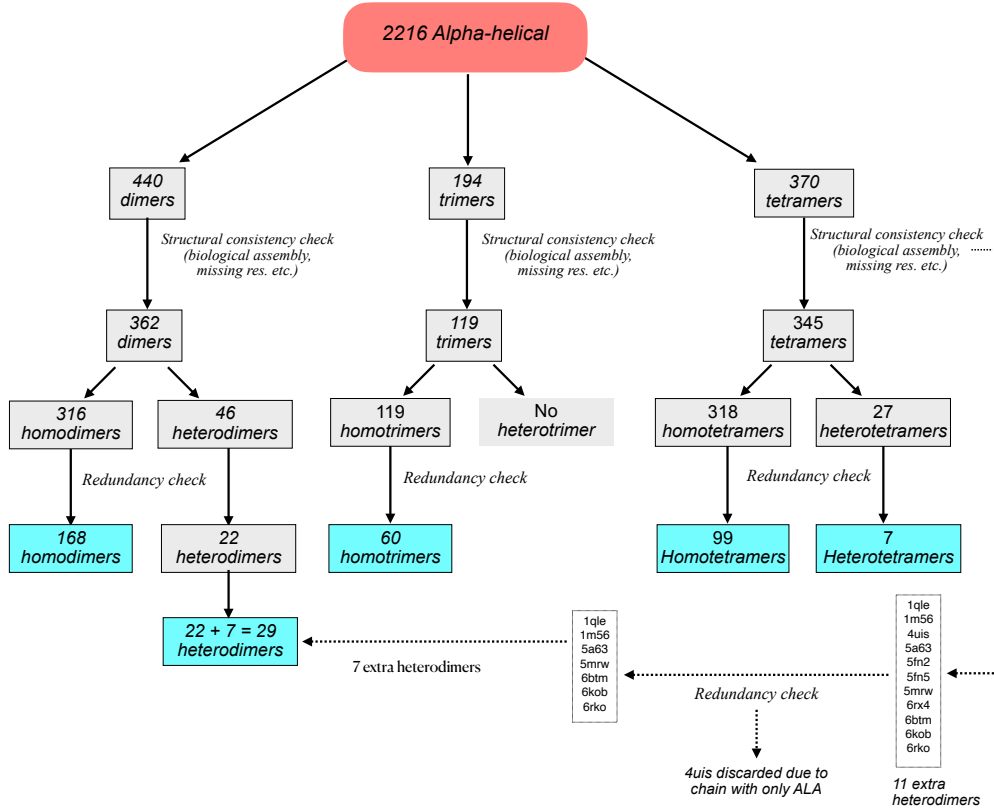
Bu projeyi, literatürde var olan diğer çalışmalardan ayıran en belirgin özellik, lipid çevrenin, sekonder yapının ve kompleks yapıdaki monomer sayısının, protein komplekslerinin arayüzey bölgelerindeki amino asit (i) dağılımı, (ii) türü ve (iii) etkileşimlerinde ne gibi farklılıklar yarattığını ortaya koyacak olmasıdır. Bu farklar, daha sonra geliştirilecek algoritma için önemli belirteçler olarak kullanılacaktır. Örneğin, transmembran bir protein kompleks yapısı tahmin edilecekse, kullanılacak parametreler, membran-dışı protein kompleksleri için belirlenen parametrelerden çok daha farklı olması beklenmelidir. Bu doğrultuda, projenin çıktıları, daha sonra geliştirilecek algoritma için bir alt yapı niteliğinde olacaktır.

2. GEREÇ ve YÖNTEM

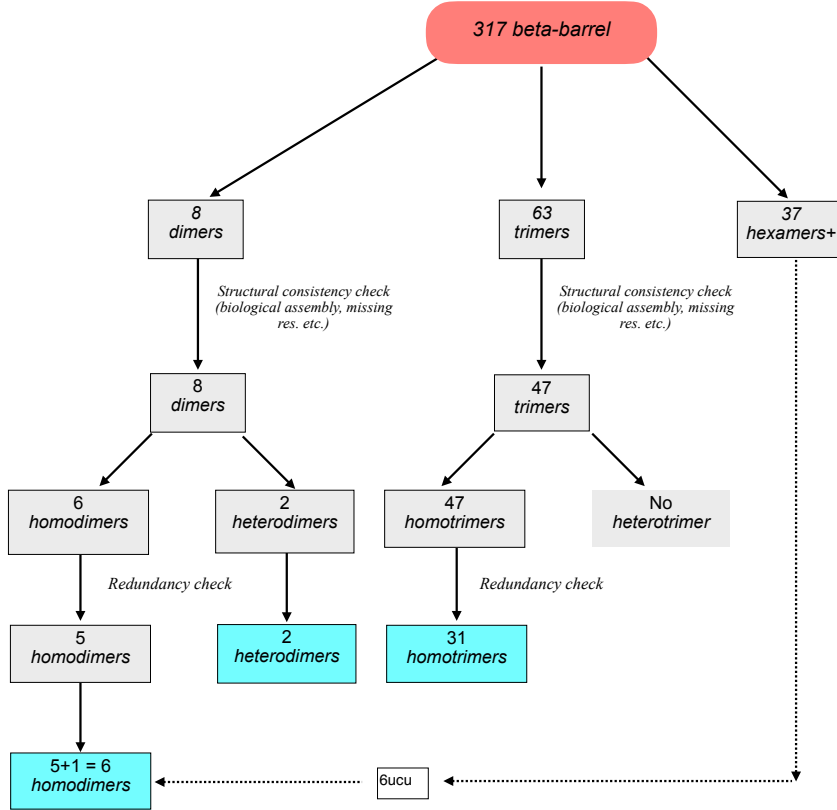
3.1. Membran protein kompleks yapılar ve hazırlanması:

Membran protein komplekslerinin üç boyutlu yapıları OPM (2019) veri bankasından alınmıştır. OPM membran proteinlerin hücre zarındaki açılma eğimini verdiği için simülasyon çalışmaları açısından son derece yararlıdır. Protein Veri Bankası'nda (2020) mevcut tüm membran proteinleri OPM'de de bulunmaktadır ve PDB'e yeni proteinler yüklendikçe OPM de kendini otomatik olarak güncellemektedir. Bu çalışmada OPM'in kullanılmış olmasının sebebi, tüm membran proteinlerin bir arada bulunuyor olmasındandır. Toplamda 2216 adet alfa-sarmal ve 317 adet beta-silindir yapıya sahip hücre zarı proteini veri bankasından alınmıştır. Oldukça kapsamlı ve titiz bir elemeye sonra, Şekil 1'de gösterildiği üzere, alfa-sarmal protein kompleksinin %85'i beta-silindir yapıların ise %88'i elelenmiştir. Beta-silindir sekonder yapıya sahip membran proteinlerin toplam sayılarının 39 gibi düşük bir seviyede olması yeteri kadar örnekleme oluşturmadığından güvenilir istatistiksel sonuçlar elde edilemeyecektir. Bu yüzden raporda verilen analiz sonuçları daha çok alfa-sarmal membran yapılar üzerinde yoğunlaştırılmıştır.

Kompleks yapıların istatistiksel analiz için hazırlık aşaması projenin en fazla zaman gerektiren aşaması olmuştur. Zira, OPM veri bankasındaki bu yapılar çoğu kez kendini tekrar etmekte, yapısal olarak gerçek anlamda bir arayüzey oluşumu göstermemekte ya da doğal olmayan deneysel kristalografik kafes yapıları (ör. PDB id: 2i36, rhodopsin) temsil etmektedir. Ancak kendini tekrar eden yapılar, Şekil 1'de de gösterildiği gibi veri bankasındaki kompleks yapıların sayısını en fazla düşüren başlıca etken olmuştur.



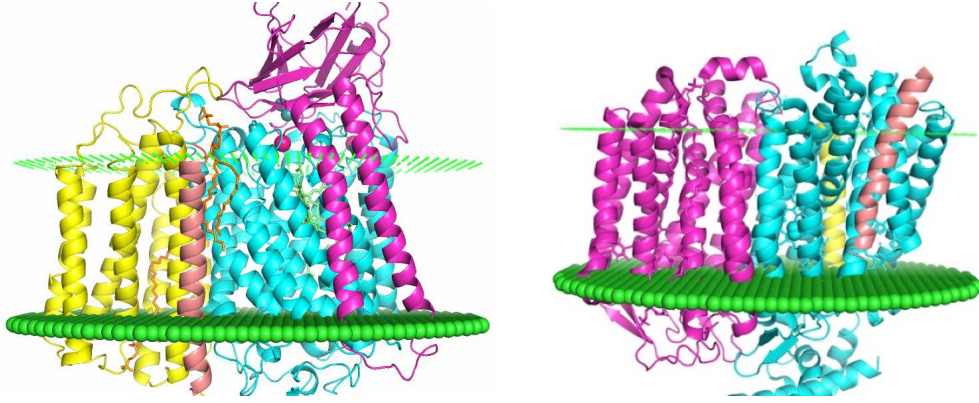
(a)



(b)

Şekil 1. OPM veri bankasından elde edilen alfa-sarmal ve bata-tabaka sekonder yapıdaki membran protein komplekslerinin yapısal uygunluklarının test edilmesi ve tekrar eden (redundant data) yapıların elenme işlemi için izlenen iş akışı.

Alfa-sarmal yapılardan tetramerik olanlar yakından incelendiğinde beklenmedik bir şekilde heterodimer arayüzeyler görülmektedir (Bknz Şekil 2). Dış kısımda etkileşim içine girmiş bir ya da iki alfa-sarmal yapının bulunması bu yapıları tetramerik kategorisine sokmuştur. Ancak, bir ya da iki adet sarmal yapıdan oluşan bir monomerin kompleks yapı oluşumuna etkisi göz ardı edilebileceğinden, bu uzantılar yapıdan çıkartılmış ve geriye kalan iki monomerik birimden oluşan heterodimerik kompleks yapı dikkate alınmıştır. Diğer bir deyişle, bu türden kompleks yapılar, heterodimerik bir arayüzey oluşturduğundan, projede heterodimer kategorisinde incelenmiştir.



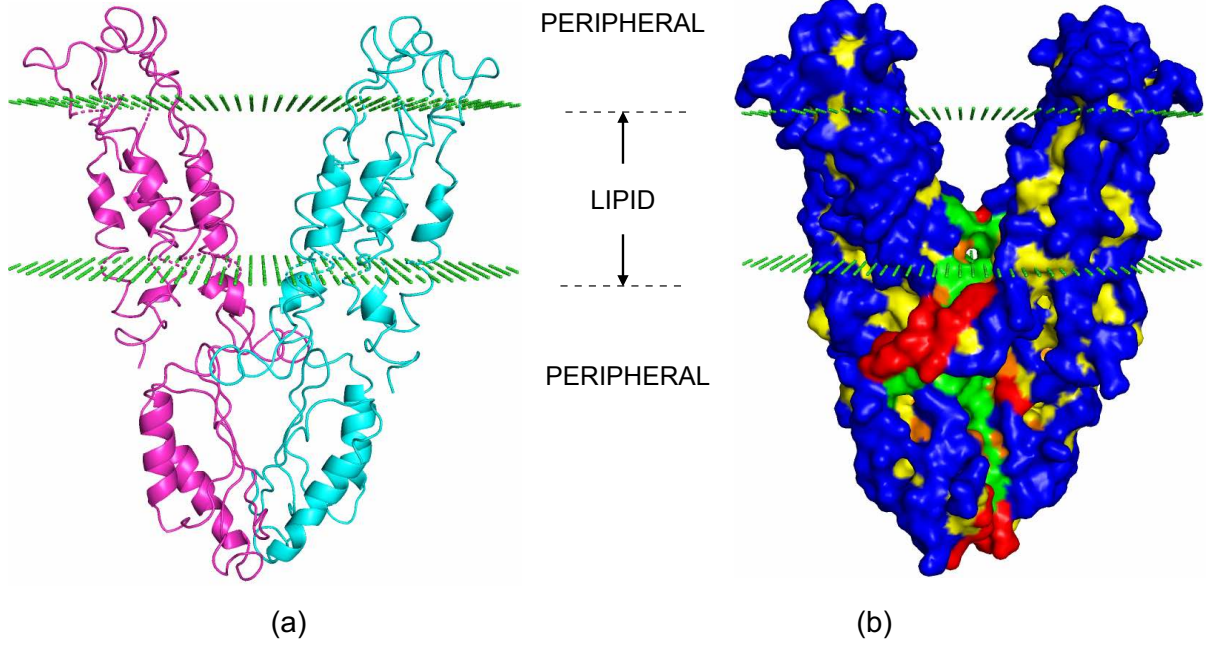
Şekil 2. OPM veribankasında tetramerik kategorisinde olup ta heterodimerik bir arayüzey oluşumu gösteren kompleks yapılar (a) PDB id: 1qle ve (b) PDB id: 6rko.

Beta-silindir kompleks yapıların alfa-sarmallara göre hücre zarı ortamında çok nadir rastlandığı taramalar sonucunda ortaya çıkmıştır. Buna göre, toplamda 317 adet yapı içeren OPM veri bankasında toplam sayı, eleme işlemlerinden sonra 76'a düşmüştür. Bunlar arasında 6 homodimer, 2 heterodimer, 31 homotrimer ve 37 hexamer ve daha yukarı seviye kompleks oluşumlar dikkat çekmektedir. Az sayıda örnekleme istatistiksel bir analizin güvenilirliğini sorgulatacağından sonuçlar değerlendirilirken alfa-sarmal yapılar üzerinde yoğunlaşmıştır.

3.2. Yüzey alana (SASA) göre protein kompleksinin farklı bölgelere ayrılması:

Protein kompleksleri \AA^2 cinsinden ifade edilen *çözücü-erişir yüzey alan* (solvent-accessible surface area) kısaca SASA olarak bilinen özelliğine bağlı olarak 5 farklı bölgeye ayrılmıştır: iç bölge, dış yüzey, destek arayüzey, çekirdek arayüzey ve kenar arayüzey bölgeler (Bknz Şekil 3). Daha sonra, her bir bölge kendi içinde "lipid" ve "periferik" olmak üzere iki farklı bölgeye daha ayrılmıştır. Şekil 3'te örnek olarak verilmiş homodimerik çinko taşıyıcı protein yapısına bakıldığında hücre zarının dışına taşmış periferik bölgenin proteinin yarısından

fazlasını oluşturduğu görülmektedir. Arayüzey bölgeler incelenirken *lipid* ve *periferik* ayrımının yapılması son derece önemlidir. Zira, periferik bölgedeki arayüzey oluşumu ve özelliklerinin lipid içinde kalan arayüzeyden farklı olacağı muhtemeldir. Veri bankasından elde edilen birçok alfa-sarmal homodimerik yapının benzer şekilde hücre dışına taşan ve oldukça fazla sayıda rezidüden oluşan periferik yapılar içerdiği görülmüştür.



Şekil 3. Hücre zarı alfa-sarmal homodimerik bir yapı olan çinko-taşıyıcı YiiP (PDB id:2qfi) için (a) cartoon gösterimi ve (b) yüzey alan üzerinde iç (sarı), dış yüzey (koyu mavi), destek arayüzey (turuncu), çekirdek arayüzey (yeşil) ve kenar arayüzey (kırmızı) bölgeler.

Rezidülerin kompleks yapıda ait oldukları bölgeleri belirlemek için SASA değeri kullanılmıştır. Kompleks yapıyı oluşturan her bir monomerik alt birim diğer alt birimlerden ayrılarak her bir rezidünün SASA değeri hesaplanmıştır. Bu değer kısaca $SASA_{i,mon}$ olarak adlandırılmıştır. Daha sonra aynı rezidülerin kompleks yapıdaki SASA değeri ($SASA_{i,comp}$) elde edilmiştir. Ardından her bir SASA değeri, o rezidü için belirlenmiş teorik SASA değerine bölünerek normalize edilmiş ve bu değer *bağıl SASA* ($rSASA$) olarak adlandırılmıştır. Teorik SASA değeri, her X rezidüsü için, *Gly-X-Gly* tripeptid oluşumundaki farklı konformasyonlar için hesaplanan en yüksek SASA değeri olarak belirlenmiştir (Tien vça., 2013).

$$rSASA_i = \left(\frac{SASA_i}{SASA_{max}} \right) \times 100 \quad (1)$$

Bağıl SASA değerleri, hem monomer haldeyken hem de kompleks yapı içinde belirlendikten sonra, farkları alınmıştır.

$$\Delta rSASA_i = rSASA_{i,mon} - rSASA_{i,comp} \quad (2)$$

Levy'nin çalışmasında verilmiş olan sınırlara göre her rezidünün proteinin hangi bölgesinde olduğu belirlenmiştir (Levy (2010)).

$$\begin{aligned}
 \text{İç Bölge:} & \quad (rSASA_{i,comp} < 25\%) \ \& \ (\Delta rSASA_i = 0) \\
 \text{Yüzey:} & \quad (rSASA_{i,comp} > 25\%) \ \& \ (\Delta rSASA_i = 0) \\
 \text{Ara Yüzey:} & \quad \Delta rSASA_i > 0
 \end{aligned} \tag{3}$$

Arayüzey ayrıca, aşağıda verilen koşullara göre üç farklı bölgeye daha ayrılmıştır:

$$\begin{aligned}
 \text{İç ara yüzey:} & \quad (\Delta rASA_i > 0) \ \& \ (rASA_{i,mon} > 0.25) \ \& \ (rASA_{i,comp} < 0.25) \\
 \text{Destek ara yüzey:} & \quad (\Delta rASA_i > 0) \ \& \ (rASA_{i,mon} < 0.25) \\
 \text{Kenar ara yüzey:} & \quad (\Delta rASA_i > 0) \ \& \ (rASA_{i,comp} > 0.25)
 \end{aligned} \tag{4}$$

Her 5 bölgenin “lipid” ve “periferik” olarak iki alt bölgeye ayrılması sonucu ortaya 10 farklı bölge çıkmıştır. Periferik bölgedeki arayüzeylerin özelliklerinin, membran komplekslerdeki arayüzey oluşumlardan farklı olması ve sitoplazmik komplekslerdeki arayüzey oluşumlarla ortak özelliklere sahip olması beklenen bir durumdur.

3.3. Farklı bölgeler için amino asit eğilimlerinin hesaplanması

Oluşturulan veri bankasındaki her protein kompleksi için, bir önceki iş paketinde yüzey alandaki değişime göre belirlenmiş bölgelerdeki amino asit dağılımı ya da, diğer bir deyişle *eğilimi* (dispozisyon), aşağıda verilen bağıntı ile hesaplanmıştır:

$$p_{i,j} = \frac{N_{i,j}}{\sum_{i=1}^{20} N_{i,j}} \tag{5}$$

Buradaki $N_{i,j}$, i rezidüsünün j bölgesindeki toplam sayısı olup, j bölgesindeki tüm rezidülere bölünerek bir frekans değeri, $p_{i,j}$ elde edilmiştir. Ancak bu değer, i rezidüsünün herhangi bir bölgeden bağımsız olacak şekilde proteinin her yerinde görülme oranına, p_i , bölünerek normalize edilmesi gerekmektedir;

$$p_i = \frac{N_i}{\sum_{i=1}^{20} N_i} \tag{6}$$

Normalize edilmiş amino asit dağılımı,

$$P_{i,j} = \frac{p_{i,j}}{p_i} \tag{7}$$

şeklinde ifade edilmiştir. Bu değer, 1'den büyük olması, i rezidüsünün, j bölgesinde bulunma *eğiliminin* ortalamaya göre yüksek olduğunu işaret etmektedir. Her protein yapısı için bu değer ayrı ayrı hesaplandıktan sonra, o kategoriye ait tüm protein kompleksleri için bu değerlerin bir ortalaması ve dağılım grafikleri elde edilmiştir. Dağılım grafikleri o amino

asidin *eğilim* değerinin kompleks yapılar arasında ne kadar değişken olduğu hakkında bilgilendirici olmaktadır.

Herhangi bir bölgedeki amino asit eğilimi (propansite), o amino asitlerin bölgedeki toplam sayısından ziyade, kapladıkları yüzey alan bilgisine dayalı olarak ta hesaplanmıştır. Bu sayede, her amino asidin dağılım değerleri, o bölgede kapladıkları farklı yüzey alanlarıyla, daha iyi ifade edilebilecek, diğer bir deyişle, SASA değeri büyük olan bir rezidünün eğilime olan katkısı daha fazla olacaktır.

$$\text{AraYüzey Propansite}_i (AP_i) = \left(\frac{SASA_{i,arayüzey}}{\sum_{i=1}^{20} SASA_{i,arayüzey}} \right) / \left(\frac{SASA_{i,yüzey}}{\sum_{i=1}^{20} SASA_{i,yüzey}} \right) \quad (8)$$

Buradaki arayüzey eğilimi kısaca AP_i (*Arayüzey Propansite*) olarak ifade edilmiş olup, $SASA_{i,arayüzey}$ arayüzey bölgedeki tüm i rezidülerin SASA değeri, $\sum_{i=1}^{20} SASA_{i,arayüzey}$ ise tüm 20 adet rezidünün arayüzeydeki toplam SASA değerleridir. Benzer şekilde, $SASA_{i,yüzey}$ yüzey bölgedeki i rezidülerin SASA değeri, $\sum_{i=1}^{20} SASA_{i,yüzey}$ ise tüm rezidülerin yüzeydeki toplam SASA değerleri olarak ifade edilmiştir. Bu değer 1'den büyük olması, o rezidünün yüzeyde bulunduğundan daha fazla arayüzeyde görüldüğünü işaret etmektedir. Buna benzer ifadeler, arayüzeyin kendi içinde de oluşturularak, arayüzeydeki üç farklı bölgenin (*çekirdek, destek ve kenar*) toplam arayüzeye göre ne kadar sıklıkta o rezidü tarafından tercih edildiği ortaya çıkartılmıştır.

$$\begin{aligned} \text{İç AraYüzey Propansite}_i &= \left(\frac{SASA_{i,iç\ arayüzey}}{\sum_{i=1}^{20} SASA_{i,iç\ arayüzey}} \right) / \left(\frac{SASA_{i,arayüzey}}{\sum_{i=1}^{20} SASA_{i,arayüzey}} \right) \\ \text{Destek AraYüzey Propansite}_i &= \left(\frac{SASA_{i,destek\ arayüzey}}{\sum_{i=1}^{20} SASA_{i,destek\ arayüzey}} \right) / \left(\frac{SASA_{i,arayüzey}}{\sum_{i=1}^{20} SASA_{i,arayüzey}} \right) \\ \text{Kenar AraYüzey Propansite}_i &= \left(\frac{SASA_{i,kenar\ arayüzey}}{\sum_{i=1}^{20} SASA_{i,kenar\ arayüzey}} \right) / \left(\frac{SASA_{i,arayüzey}}{\sum_{i=1}^{20} SASA_{i,arayüzey}} \right) \end{aligned} \quad (9)$$

Arayüzey propansite değerleri aynı zamanda, 2 tabanında logaritma değeri alındıktan sonra bir skor değerine dönüştürülerek kullanılmıştır. Örneğin, 8 no'lu eşitlikte verilmiş arayüzey propansite değerinin 0 olması, o rezidünün arayüzeyde ve yüzeyde eşit miktarlarda bulunduğunu, diğer bir deyişle, bir bölgeyi diğer bölgeye tercih etmediğini işaret etmektedir. Bu değer 1 ya da daha fazla olması ise, o rezidünün arayüzeyi yüzeye göre 2 ya da daha fazla kat tercih ettiğini göstermektedir.

3.4. Arayüzeyde etkileşen rezidülerin kontakt derecesi ve görülme sıklıklarının hesaplanması

Arayüzey bölgede bulunan rezidü çiftlerinin aralarındaki yakınlık (kontakt) derecesi, diğer bir deyişle iki monomer birimin hangi tür rezidü çiftleri aracılığıyla bir araya gelerek bir arayüz oluşturdukları, birimler arası etkileşimin derecesi ve türü hakkında önemli bilgiler içermektedir. Diğer taraftan, bu etkileşimlerin, membran ya da periferik (membran dışı) çevreden etkilenip etkilenmedikleri bilinmemektedir. Arayüzey bölgedeki iki rezidünün kontakt halde bulunabilmesi için uzaklık kriterleri iki ana atom arası 6 Å, ana atom ve yan atom arası 5 Å, iki yan atom arası 4 Å olarak belirlenmiştir ve iki rezidünün sahip olduğu ağır atomlardan (hidrojen atomu dışında kalan tüm atomlar) herhangi birinin bu uzaklık kriterlerinden herhangi birini sağlaması yeterlidir.

Her bir rezidü için öncelikle, kontakt derecesi (KD_i) hesaplanmıştır. Bu değer, arayüzey bölgede olup ta kontakt kuran i türündeki rezidülerin toplam sayısının (K_i), arayüzey bölgedeki tüm i rezidülerin toplam sayısına (N_i) oranı şeklinde belirlenmiştir:

$$KD_i = K_i/N_i \quad (10)$$

Bu değer daha sonra, arayüzeydeki ortalama kontakt derecesine bölünerek normalize edilmiştir. Ortalama kontakt derecesi arayüzeyde kontakt kuran toplam rezidü sayısının arayüzeydeki toplam rezidü sayısına oranı şeklinde belirlenmiştir. Normalize edilmiş KD_i aşağıdaki eşitlikle ifade edilmiştir:

$$NKD_i = \frac{K_i/N_i}{\sum_{i=1}^{20} K_i / \sum_{i=1}^{20} N_i} \quad (11)$$

Kontakt derecesinin yanı sıra, hangi tür i ve j amino asit arasında kontakt olduğunu belirlemek üzere, ij ikili amino asitlerin görülme sıklıkları (kontakt frekansı) hesaplanmış ve bir skor değerine dönüştürülmüştür. Aşağıda verilen olan 12 no'lu eşitlikte, arayüzeyde kontakt haldeki monomer A'ya ait i tipi rezidüler ile monomer B'deki j tipi rezidüler arasındaki kontakt sayısı ($K_{iA,jB}$), monomer B'deki i rezidüler ile monomer A'daki j rezidüler arasındaki kontakt sayısına ($K_{iB,jA}$) eklenere, arayüzeyde gözlemlenen toplam kontakt sayısına bölünmüştür. 12 no'lu eşitlikteki N_A ve N_B sırasıyla monomer A ve B'deki toplam rezidü sayısıdır. $K_{i,j}$ kontakt halde ise 1, değilse 0 olarak alınmıştır.

$$KP_{i,j} = \frac{K_{iA,jB} + K_{iB,jA}}{\sum_{i=1}^{N_A} \sum_{j=1}^{N_B} K_{i,j}} \quad (12)$$

Son adımda, gözlenen (*observed*) kontakt frekansı KP_{ij} , beklenen (*expected*) kontakt frekansına bölünerek ve 2 tabanında logaritmasında ifade edilerek aşağıda verilen eşitlikteki

gibi bir skor değerine dönüştürülmüştür. Bu değer, 0'dan farklı olması, arayüzeyde gözlemlenen ij rezidü çiftinin rastlantısal bir şekilde biraraya gelmediğine işaret etmektedir. Evrimsel açıdan bakıldığında, iki rezidünün kontakt kurmayı tercih etmelerinin (ya da etmemelerinin) daha kararlı arayüzey oluşumlarına olanak vermesi şeklinde bir sebep-sonuç ilişkisi çıkarılabilmektedir.

$$S_{i,j} = \log_2 \left(\frac{\text{gözlenen kontakt frekans}}{\text{beklenen kontakt frekans}} \right) = \log_2 \left(\frac{KP_{i,j}}{f_{i,A} \times f_{j,B} + f_{i,B} \times f_{j,A}} \right) \quad (14)$$

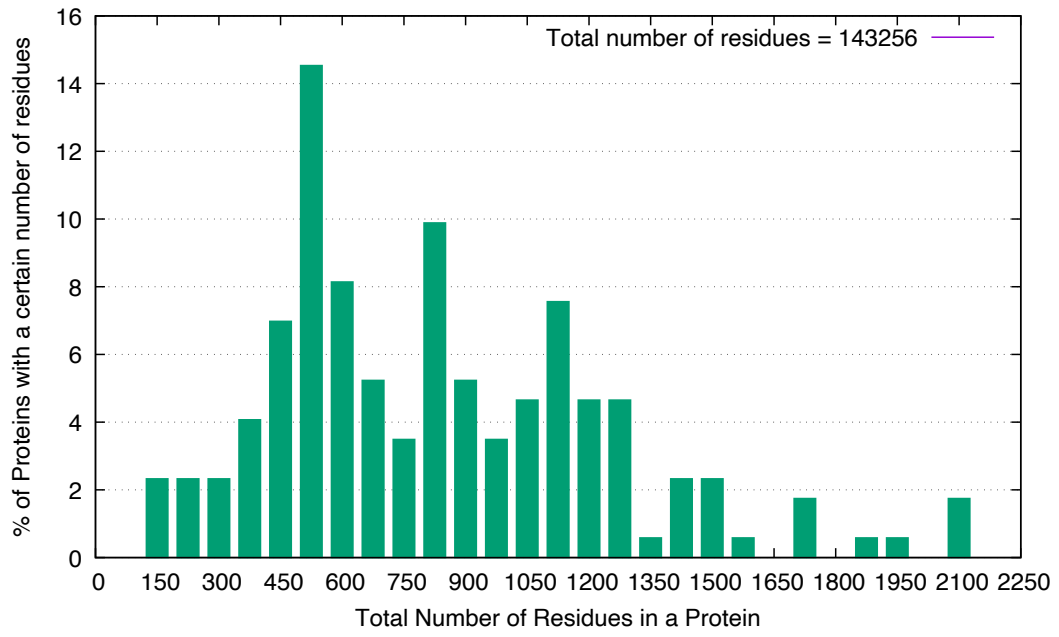
Buradaki $f_{i,A}$ ya da $f_{i,B}$ değerleri i rezidüsünün monomer A ve B'deki frekanslarıdır ve belirtilen monomerin arayüzeyindeki toplam i rezidüsünün yine o monomerin arayüzeyindeki toplam rezidü sayısına (ör. $N_{i,A}/N_A$) bölünerek elde edilmiştir. Benzer şekilde $f_{j,A}$ ve $f_{j,B}$ değerleri j rezidüsünün monomer A ve B'deki frekanslarıdır.

4. BULGULAR ve TARTIŞMA

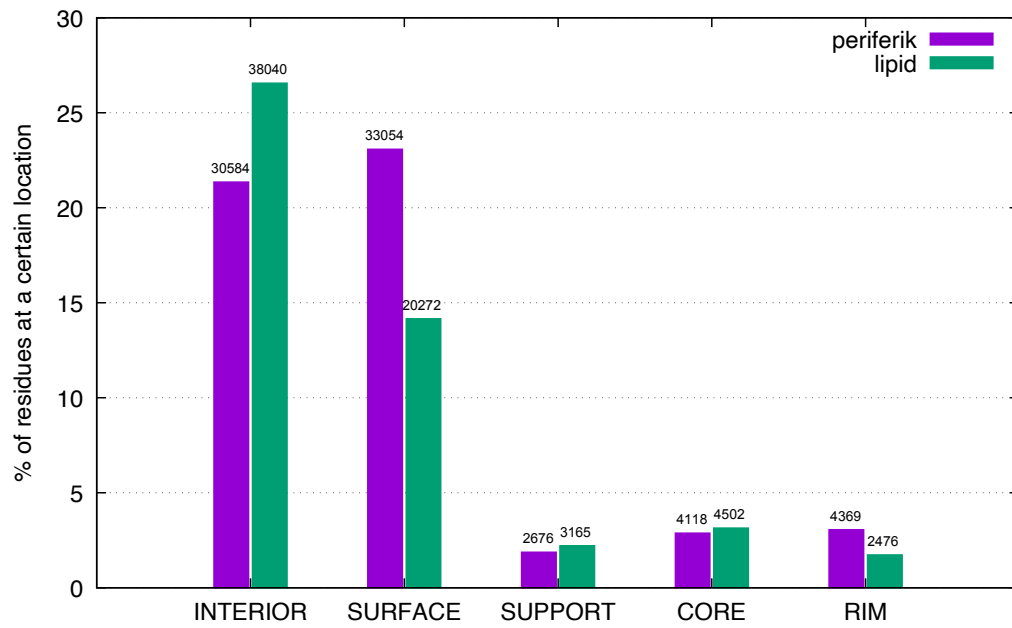
4.1. Alfa-sarmal membran protein komplekslerde ortalama amino asit miktarı

OPM veri bankasında çekilmiş ve yöntem kısmında anlatılan elemelerden geçen toplam 168 adet alfa-sarmal hücre zarı proteini için ilk aşamada proteindeki toplam rezidü sayısına göre bir dağılım grafiği elde edilmiştir. Şekil 4a'da gösterildiği gibi, 150-2250 arası rezidü barındıran farklı boyutlardaki yapılardan oluşmaktadır. Toplam 143,256 rezidü sayısı içeren veri setinde, rezidülerin buldukları farklı bölgelere göre dağılımları Şekil 4b'deki gibidir. Buna göre 38,040 rezidü ile en fazla rezidü içeren bölge lipid ortamdaki iç bölge olmuştur ve toplam rezidülerin %25'inden fazlasını oluşturmaktadır. İkinci sırada, periferik ortamda yüzey bölgedeki rezidülerin sayısı (toplam 33,054) gelmektedir. Arayüzey bölgelerde ise en fazla rezidü sayısı lipid ortamdaki çekirdek arayüzeye ait olup 4502 adettir.

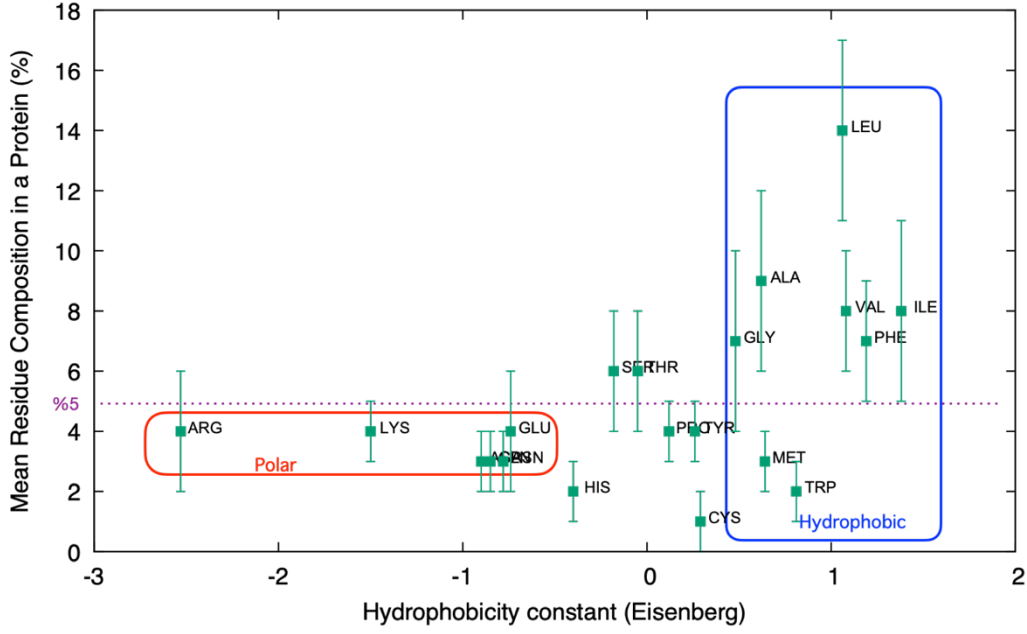
Her bir kompleksteki 20 rezidünün miktarları o proteindeki toplam rezidü sayısına bölünerek yüzde olarak ifade edilmiştir. Veri bankasındaki 168 protein için ortalama bir bileşim oranı belirlenmiş ve Şekil 4c'deki gibi hidrofobisite sabitlerine göre sıralanmıştır. Özellikle *Leu* gibi hidrofobik rezidülerin önemli bir kısmının polar rezidülere göre daha fazla sayıda oldukları görülmektedir. Tamamen rastgele dağılımın ifadesi olan %5'in üstünde olması da istatistiksel bir anlam kazandırmaktadır.



(a)



(b)

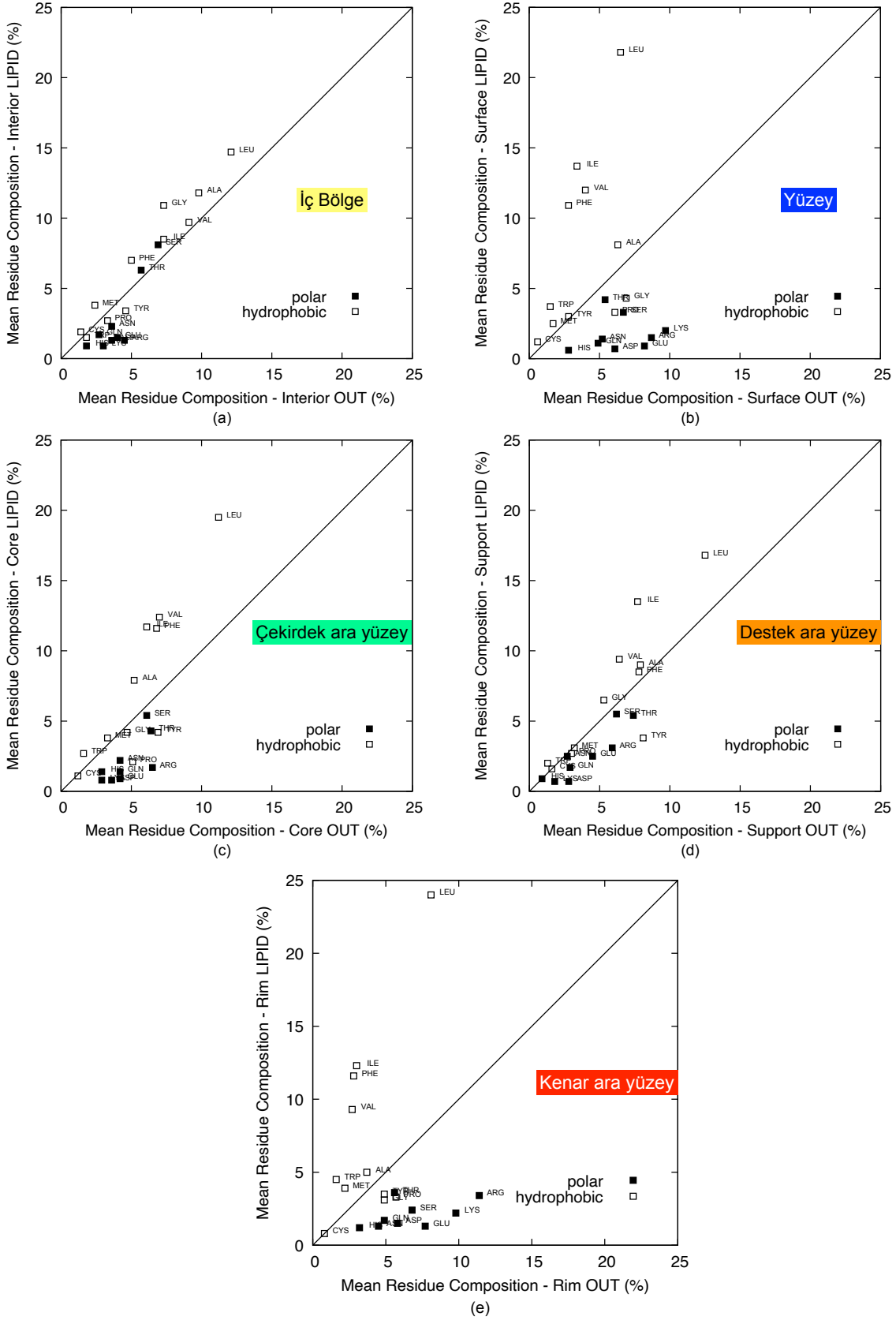


(c)

Şekil 4. 168 alfa-sarmal homodimerik protein kompleksinde rezidülerin (a) toplam sayısına ve (b) farklı bölgelerindeki dağılım grafikleri. (c) hidrofobisite sabitine (Eisenberg vd. (1984)) göre ortalama rezidü bileşimi.

Şekil 4'teki ortalama rezidü bileşim yüzdeleri tüm bölgeyi kapsamaktadır. Diğer taraftan, lipid ve periferik bölgelerin sundukları farklı kimyasal ortamların yanı sıra, yüzey ve arayüzey bölgeler birbirlerinden oldukça farklı yapıdadır. Dolayısıyla, hidrofobik ve polar rezidüler her bölgede farklı bileşim yüzdeleri gösterecektir. Şekil 5'te gösterildiği gibi, her bir farklı beş bölge için {1.iç, 2.yüzey, arayüzey (3.çekirdek, 4.destek, 5.kenar)} hem lipid hem de periferik kısımda olmak üzere iki farklı ortamdaki ortalama bileşim yüzdeleri elde edilmiştir. Buna göre, *Leu* lipid ortamdaki beş bölgede en çok sıklıkta görülen rezidü olmuştur. Periferik dış ortamda ise, sadece yüzey ve kenar arayüzey bölgelerde polar rezidülerin gerisinde durmaktadır. Hidrofobik rezidülerin polar rezidülere göre en çok görüldüğü bölge lipid ortamdaki yüzey ve kenar-arayüzey bölgeler olmuştur. Polar rezidüler ise beklenildiği üzere, en fazla sayıda periferik ortamdaki yüzey ve kenar arayüzey bölgelerde bulunmaktadır.

Şekil 5'te gösterilen diagonal çizgiden uzaklaşan noktalar proteindeki her bir beş farklı bölgenin periferik ve lipid ortamda farklı rezidü bileşimleri gösterdiğini işaret etmektedir. Bu farkın en çarpıcı olduğu bölgeler yüzey ve *kenar*-arayüzey olmuştur. Lipid moleküllerle çevrili bir yüzey oluşumunun, polar bir ortamdaki periferik yüzey oluşumundan farklı olması beklenen bir durumdur. Diğer taraftan, periferik ve lipid ortamdaki bileşim yüzdelerinin birbirine en yakın olduğu bölge iç bölge olmuştur. Bu durum iç bölgenin dış ortamdaki en az etkilenen bölge olduğunu açıkça göstermektedir.

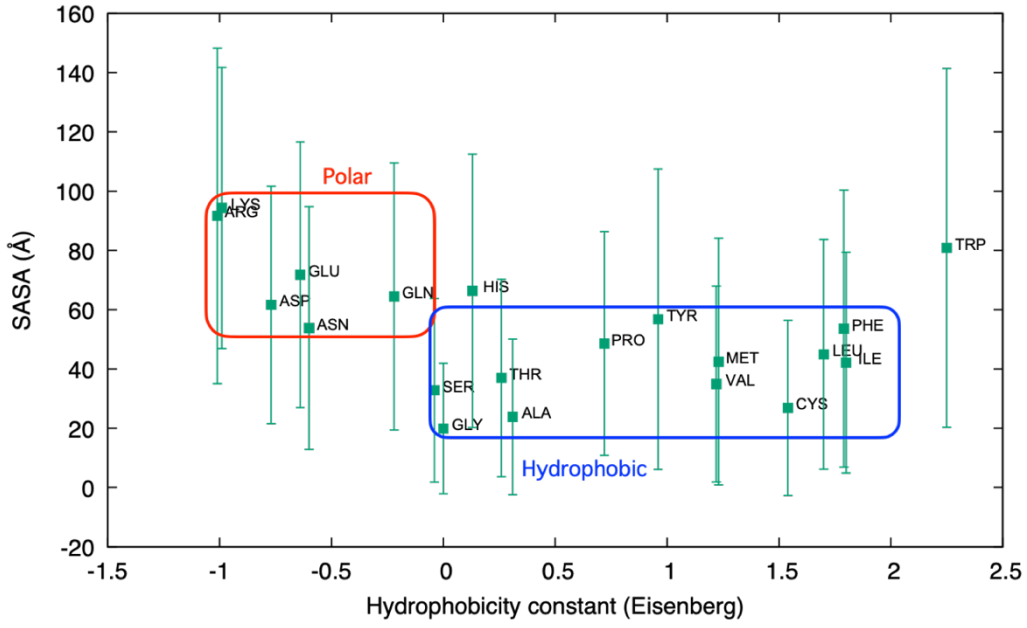


Şekil 5. 168 alfa-sarmal homodimerik protein kompleksinde hidrofobisite sabitine (Eisenberg vd. (1984)) göre farklı bölgelerdeki ortalama rezidü bileşimi. (a) İç bölge, (b) yüzey bölge, (c) çekirdek arayüzey, (d) destek arayüzey ve (e) kenar arayüzey.

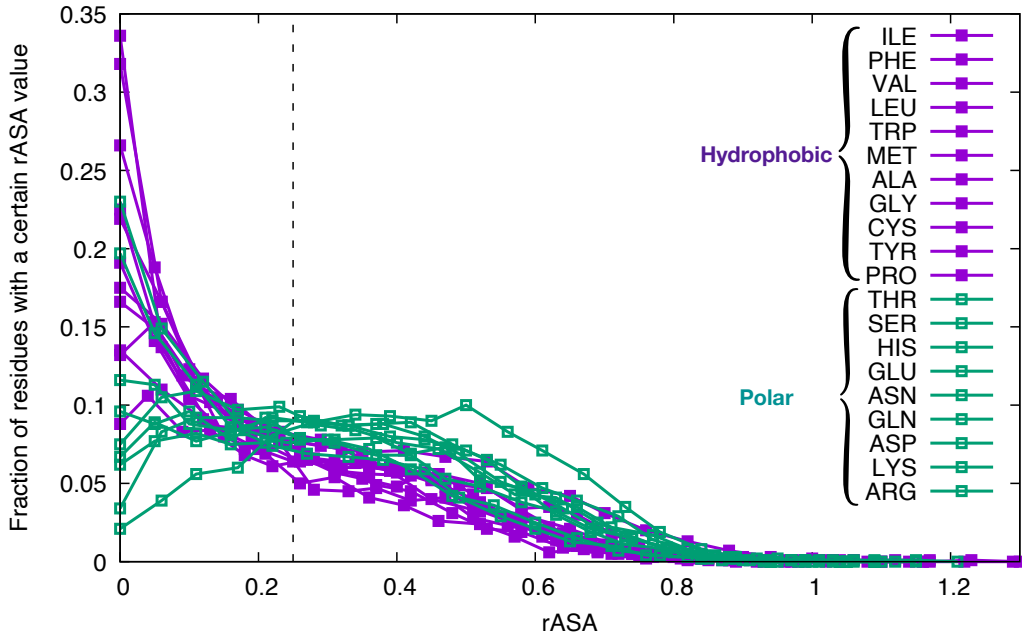
4.2. Alfa-sarmal membran protein komplekslerde ortalama SASA/rASA değeri:

Veri setindeki 168 alfa-sarmal membran protein kompleksi için, bileşim yüzdelerinin yanı sıra, *çözücü-erişir yüzey alana* sahip bölgelerde her bir rezidünün kapladığı ortalama yüzey alanları belirlenmiştir. Hidrofobisite sabitine göre Şekil 6a'daki gibi gösterilen SASA değerlerine bakılacak olursa, hidrofobik rezidülerin SASA değerleri 20-60 Å² civarında değişirken polar rezidü SASA değerleri 60-100 Å² arasında görülmektedir. Hidrofobik rezidülerden *Trp* bu kuralı bozan tek rezidüdür, zira polar rezidüler kadar yüksek SASA değerine sahiptir 80.8 ± 60.5 Å². Ancak standard sapma değerlerinden de görüldüğü üzere, veri bankasındaki proteinler büyük bir değişkenlik göstermektedir. Şekil 6b'de gösterilen rASA dağılım grafiğinde her rezidünün 0-1.3 aralığında farklı rASA değerlerine sahip oldukları görülmektedir. rASA=0.25 noktasında belirtilen dik noktalı çizgi, hidrofobik ve polar rezidülerin birbirlerine göre olan dağılımlarındaki değişimin en çarpıcı biçimde gerçekleştiği değeri belirtmektedir. Bu değerden büyük rASA değerine sahip rezidülerin yüzey ve kenar arayüzeyde oldukları düşünüldüğünde polar özellikle olanların çoğunlukla bu bölgelerde görülmeleri rastlantı değildir. Diğer taraftan hidrofobik rezidülerin büyük çoğunluğu rASA < 0.25 değerine sahip iç ve çekirdek/destek arayüzey bölgelerde görülmektedir.

SASA değerlerinin rezidülerin buldukları bölgeye göre karşılaştırmalı bir analizi Şekil 7'deki gibi verilmiştir. Yüzey ve kenar arayüzey bölgelerin diğer bölgelere göre en yüksek SASA değerlerine sahip oldukları, özellikle periferik taraftaki yüzey bölgelerin lipid taraftakilere göre daha yüksek SASA'lı rezidüler barındırdıkları görülmektedir. Diğer taraftan, çekirdek arayüzey ve destek arayüzeydeki rezidüler için lipid ya da periferik bölgede olmaları SASA değerleri üzerinde belirgin bir etki yapmamıştır. Polar rezidüler arasında *Arg* çekirdek arayüzeyde görülen en yüksek SASA değerine sahip rezidü olmuştur. Genelde, polar rezidüler arayüzeyde fazla sayıda olmamalarına rağmen SASA değeri bakımından hidrofobik rezidülere göre daha erişilabilir bir konformasyonda bulunmaktadırlar. Dolayısıyla, arayüzeydeki *Arg* rezidülerinin monomerler arası etkileşimde önemli bir rol oynaması kaçınılmazdır.

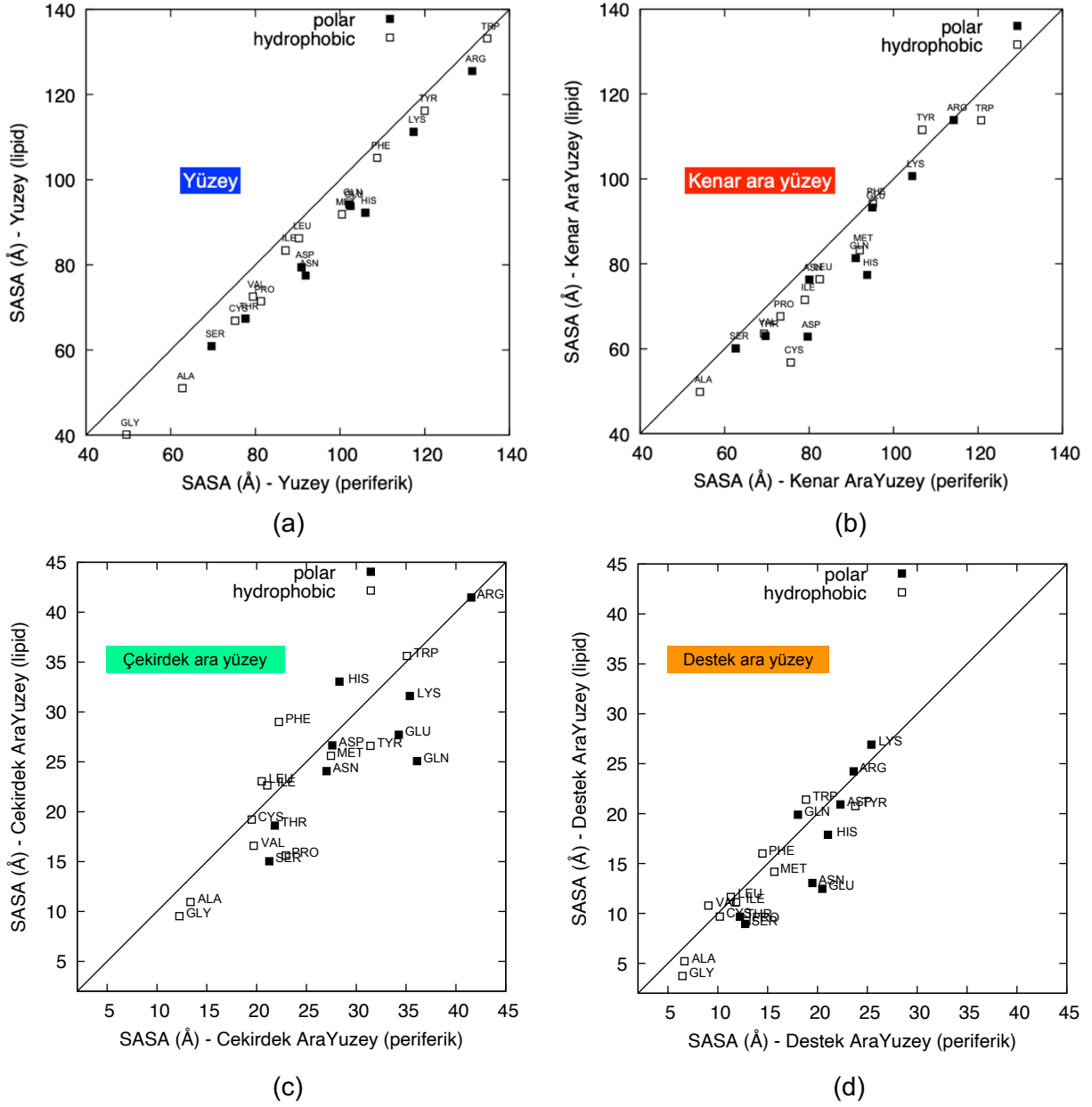


(a)



(b)

Şekil 6. 168 alfa-sarmal homodimerik protein kompleksinde (a) hidrofobisite sabitine (Eisenberg vd. (1984)) göre ortalama SASA değeri. (b) rASA dağılım grafiği.



Şekil 7. 168 alfa-sarmal homodimerik protein kompleksinde rezidülerin farklı bölgelerdeki ortalama SASA değerleri. (a) yüzey, (b) kenar arayüzey, (c) çekirdek arayüzey ve (d) destek arayüzey.

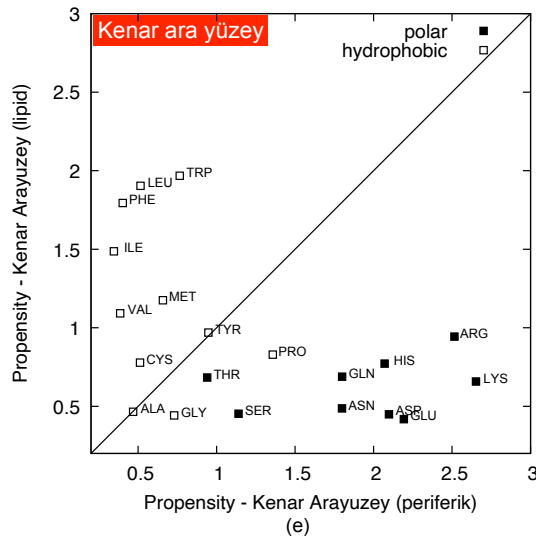
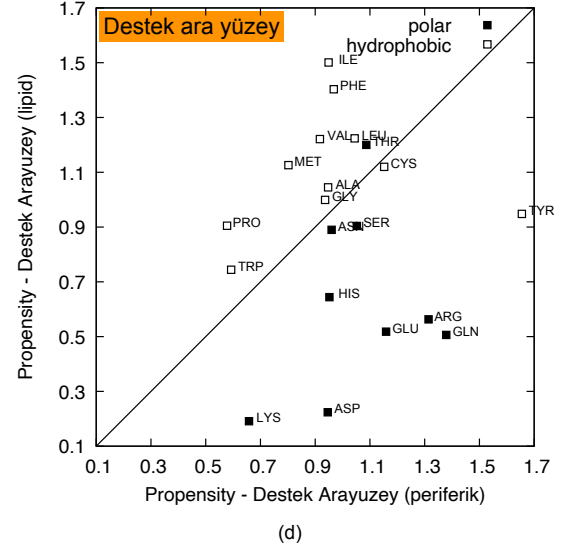
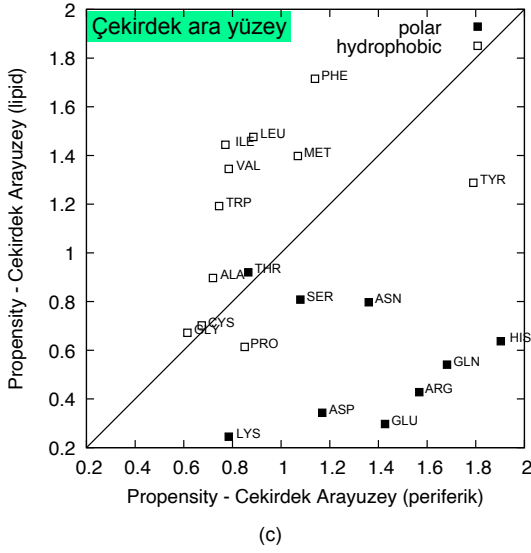
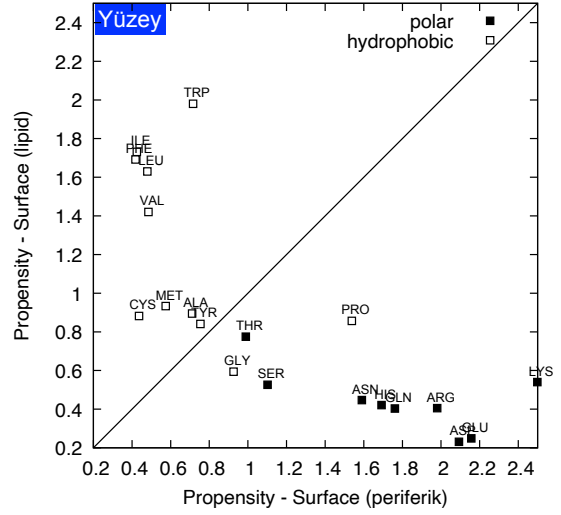
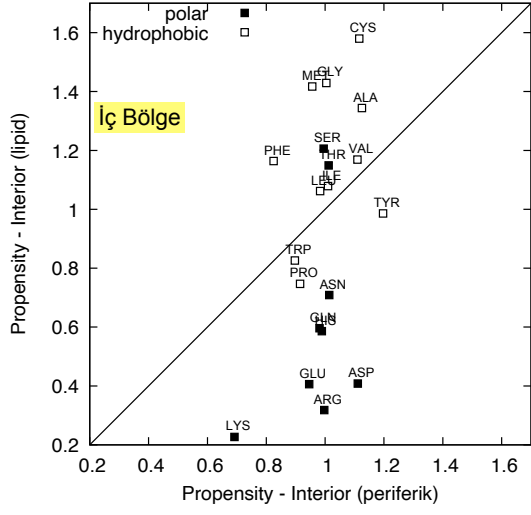
4.3. Alfa-sarmal membran protein komplekslerindeki rezidülerin ortalama eğilim (propansite) değerleri

Rezidülerin ortalamaya göre belirli bir bölgede daha fazla bulunması propansite olarak ta adlandırılan eğilim değerleriyle ifade edilmektedir (Bknz Yöntem). Bu değerın 1'e eşit olması, ortalama bileşim oranı ile aynı olduğunu göstermektedir ve belirtilen bölge için herhangi bir eğilimi bulunmamaktadır. 1'den büyük ya da küçük olması ise eğilimin ortalamadan büyük ya da küçük olduğunu işaret etmektedir. Şekil 8'de gösterildiği gibi, periferik ve lipid ortamlardaki rezidü dağılımları birbirlerinden oldukça farklıdır. Örneğin, periferik iç bölgelerde rezidülerin

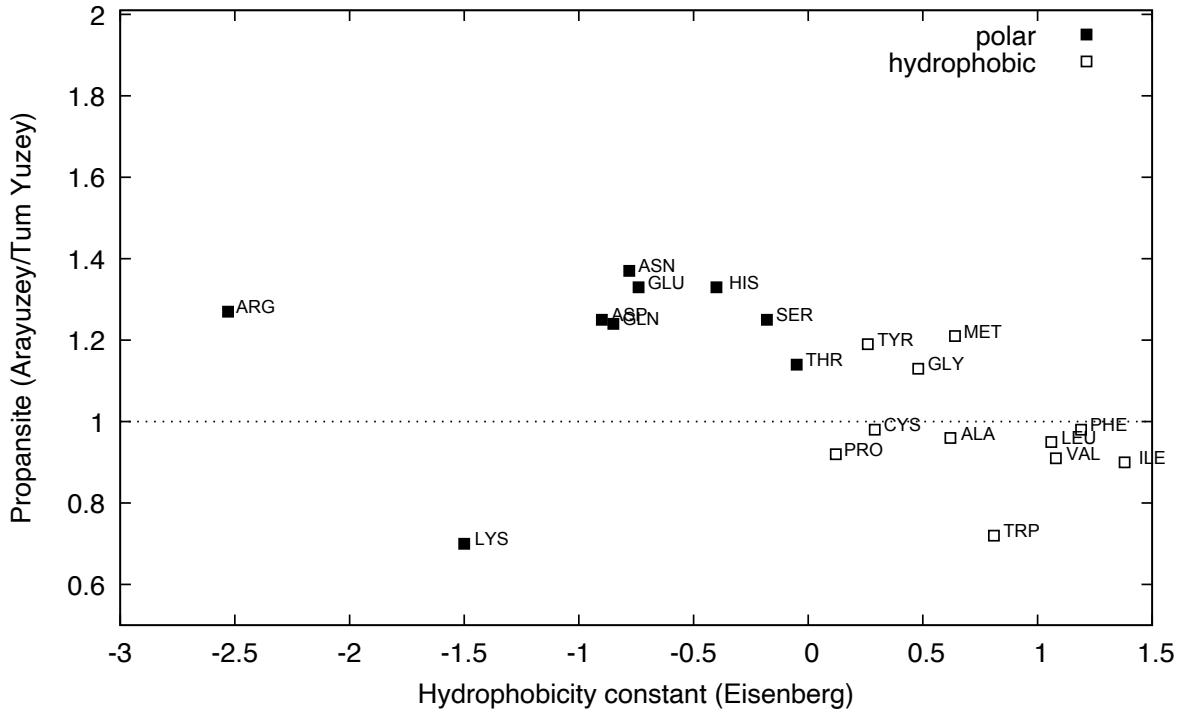
eğilim değerleri 1'e çok yakınken (iç bölge için herhangi bir tercihin söz konusu olmadığı bir durum), lipid iç bölge için hidrofobik rezidüler ortalamadan daha fazla, *Ser* ve *Thr* hariç polar rezidüler de daha az görülmektedir.

Yüzey ve kenar arayüzey bölgeler, lipid ortamdayken hidrofobik rezidüler tarafından, periferik ortamdayken polar rezidüler tarafından tercih edilmektedir. Dağılım grafiğinin en çarpıcı biçimde değiştiği bölgeler yüzey ve kenar arayüzeyler olmuştur. Bu durum, sitoplazmik protein komplekslerinin membran komplekslere göre ne kadar farklı bir rezidü profili sergileyeceğinin ilk işaretleridir. Hücre zarının üst ve alt bölgelerinde kalan periferik bölge, lipid ile dış ortam arasında bir geçiş bölgesi olduğundan membran ve sitoplazmik protein kompleksleri birbirinden ayırt etmek için son derece önemli ipuçları sunmaktadır. Önemli bir başka bulgu da, lipid ortamdaki kenar arayüzeylerdeki *Trp*, *Leu* ve *Phe* hidrofobik rezidülerin görülme sıklıklarının ortalamadan iki kat daha fazla olmasıdır. Hidrofobik rezidüler çekirdek ve destek arayüzey bölgelerine periferik ortamdayken belirli bir eğilim göstermezken, lipid taraftaki bu bölgeleri belirgin bir şekilde tercih etmektedirler. Diğer taraftan polar rezidüler lipid taraftaki çekirdek ve destek arayüzey bölgelerde ortalamadan daha az görülürken periferik ortamda aynı bölgelere daha fazla eğilim göstermektedirler, ancak bu eğilim periferik ortamdaki kenar arayüzey bölgelerde olduğu kadar belirgin değildir.

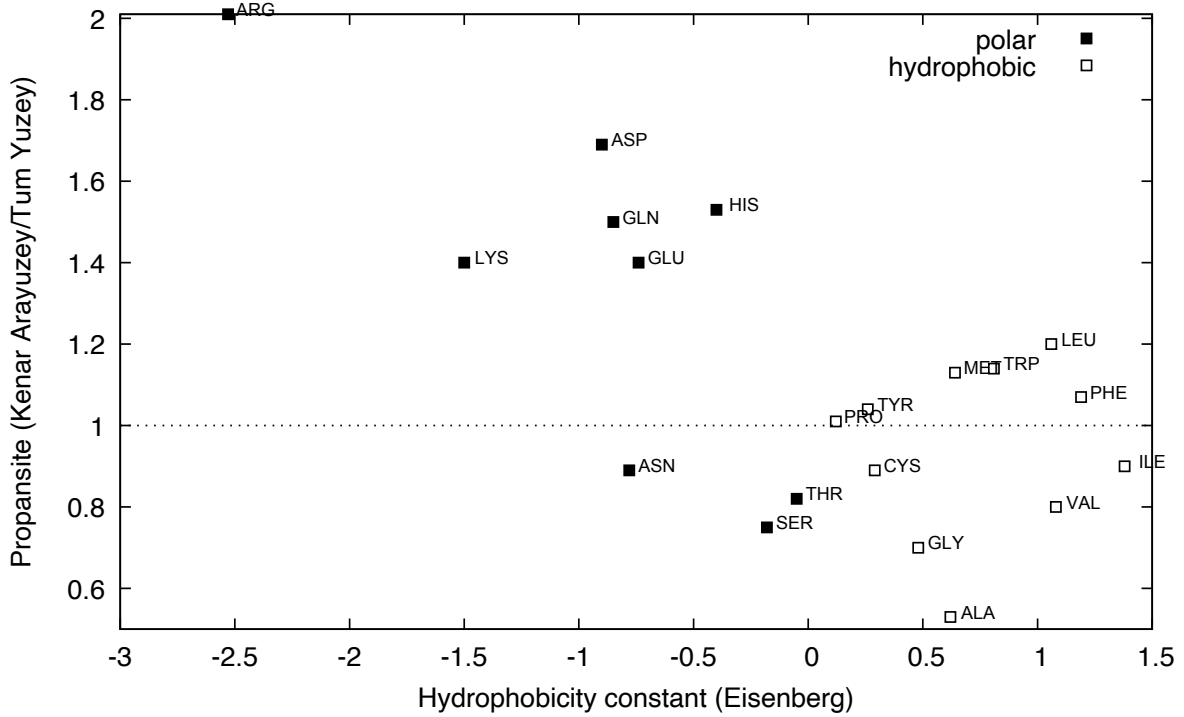
Hazırlık aşamasında olan bir sonraki projemizde geliştirmeyi planladığımız membran proteinlerde kompleks yapı tahmini algoritması için kullanılacak son derece önemli bir bulgu da, tüm lipid yüzey alana göre lipid arayüzey eğilim profilinden elde edilmiştir. Burada eğilim, tüm bölgelere göre değil, lipid bölgedeki arayüzey de dahil tüm yüzey alanına göre normalize edilmiştir (Bknz Yöntem, Denklem 8). Buna göre, polar rezidülerin arayüzeyi tüm lipid yüzey alana göre daha fazla tercih ettikleri açıkça görülmektedir. Diğer taraftan hidrofobik rezidülerden sadece *Tyr*, *Met* ve *Gly* polar rezidülere benzer bir özellik gösterirken, tam tersi bir durum sadece *Trp* için mevcuttur. Geri kalan diğer hidrofobik rezidüler ise tüm yüzey bölgede arayüzeyi ayırt edici bir şekilde tercih etmemektedir. Ancak, benzer bir analiz kenar arayüzey ile toplam yüzey alan için yapıldığında, polar rezidülerin kenar arayüzeyi daha belirgin bir şekilde tercih ettikleri görülmüştür (Bknz 9b). Özellikle *Arg*, 2.01'lik bir propansite değeriyle, kenar arayüzeyi en fazla tercih eden rezidü olmuştur. Hidrofobik rezidülerden ise sadece *Leu* polar rezidülere benzer bir eğilim göstermektedir, *Ala*, *Gly*, *Val* ise kenar arayüzey dışındaki diğer yüzey bölgelerde bulunmayı yeğlemektedirler. Bu sonuçlar, kompleks yapısı bilinmeyen ve arayüzey tahmini yapılacak sistemlerde kullanılabilir önemli bir parametreler sunmaktadır. Tüm yüzey alanı taranan bir monomer yapıda, olası arayüzey oluşturabilecek polar rezidüler (özellikle *Arg*) taranabilir ve sayıca en fazla yoğunluk gösteren bölgeler olası arayüzey bölgesi olarak öne sürülebilir.



Şekil 8. 168 alfa-sarmal homodimerik protein kompleksinde rezidülerin farklı bölgeler için propansite (eğilim) değerleri. (a) yüzey, (b) kenar arayüzey, (c) çekirdek arayüzey ve (d) destek arayüzey.



(a)

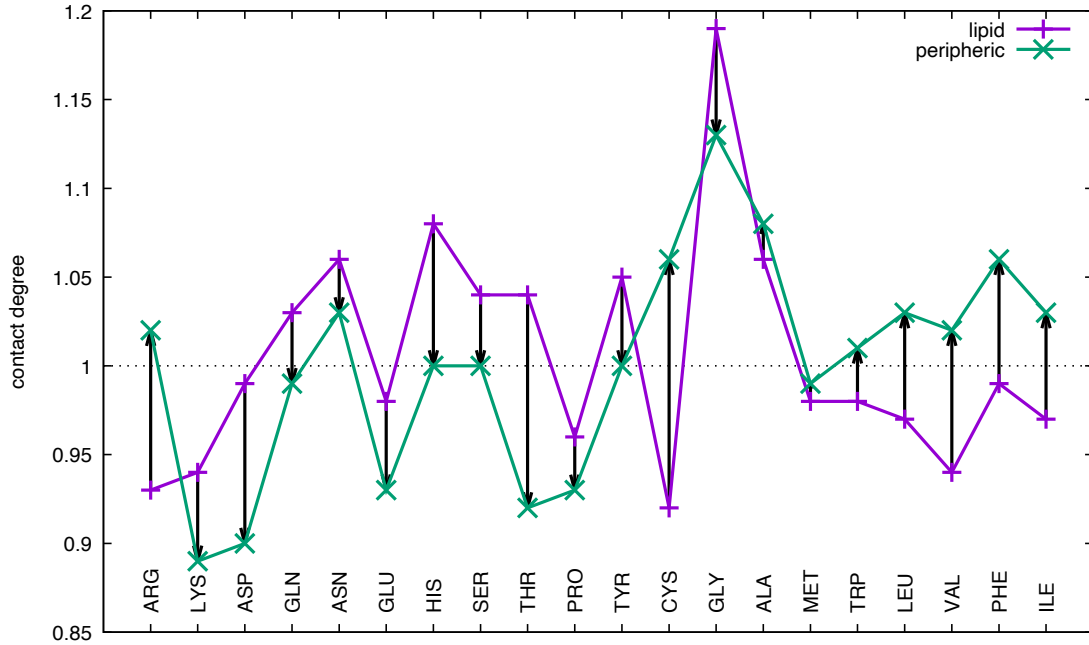


(b)

Şekil 9. 168 alfa-sarmal homodimerik protein kompleksinde hidrofobisite sabitine (Eisenberg vd. (1984)) göre tüm yüzeye göre (a) arayüzey ve (b) kenar arayüzey propansite değerleri.

4.4. Alfa-sarmal membran protein komplekslerdeki arayüzey bölgelerde kontakt rezidü çiftlerinin belirlenmesi

Her rezidü türü için yöntem kısmında anlatıldığı gibi belirlenen kontakt derecesi Şekil 10'daki gibi gösterilmiştir. X-ekseni üzerinde belirtilen rezidüler Eisenberg hidrofobisite sabitine göre sıralanmıştır, diğer bir deyişle, negatif değerdeki sabitlerle ifade edilen polar rezidüler şeklinde solunda, hidrofobik rezidüler ise sağında yer almaktadır. Her iki ortam (lipid ve periferik) için ayrı ayrı hesaplanan kontakt derecelerinin birbirlerinden oldukça farklı oldukları belirgin bir şekilde göze çarpmaktadır. Bir başka ilgi çekici sonuç ta, lipid ortamdaki periferik ortama geçildiğinde kontakt derecesinin polar rezidüler için azalması, hidrofobik rezidüler içinse artmasıdır. Şayet lipid ortamdaki bir arayüzeyden bahsediliyorsa, arayüzeyde bulunan polar rezidüler diğer monomerdeki rezidülerle ortalamanın üstünde bir etkileşim içine girmektedir. Hidrofobik rezidülerde ise benzer bir durum periferik arayüzeylerde gerçekleşmektedir. Polar rezidüler lipid arayüzeylerde hidrofobik rezidülere göre çok daha az sayıda görülmesine rağmen, monomer-monomer etkileşimlerinde hidrofobik rezidülere göre daha fazla katkı sağlamaktadır. Periferik arayüzeylerde polar rezidülere göre daha az sayıda olmalarına rağmen hidrofobik rezidüler arayüzey etkileşimlerinde daha fazla rol oynamaktadırlar, zira polar rezidüler periferik ortamda dışarıyla temas haldeki yüzey bölgelerde bulunmayı tercih etmektedirler.



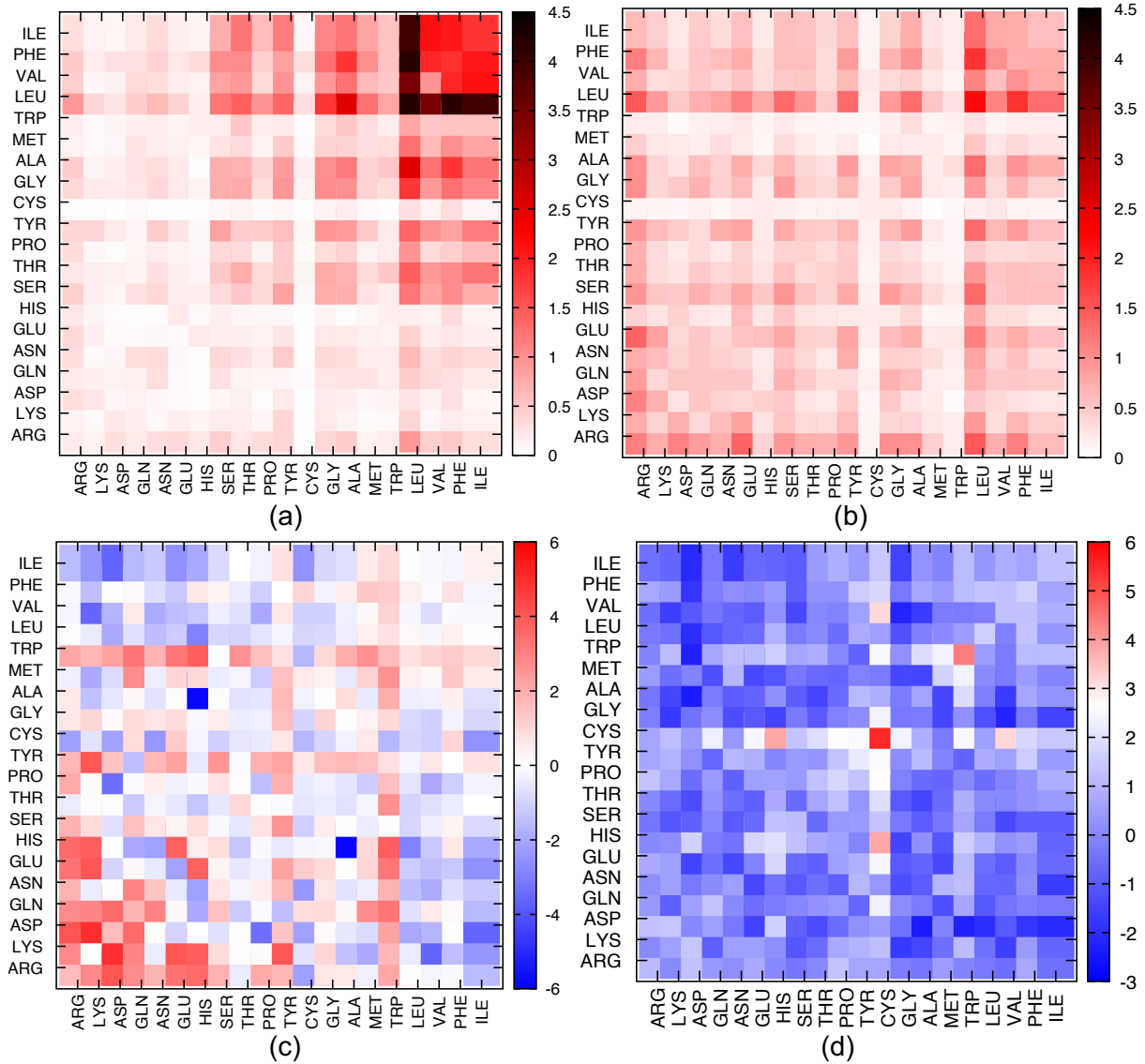
Şekil 10. Rezidü türlerinin lipid ve periferik ortamdaki arayüzey bölgelerdeki kontakt derecesi.

Yan zincir yerine bir Hidrojen atomuna sahip en küçük rezidü olan *Gly* arayüzeyde ortalamanın üstünde bir kontakt kurma eğilimi göstermektedir. Bu durum hem lipid hem de periferik ortamlar için gözlemlenmektedir. Arayüzeyden ziyade iç bölgelerde bulunmayı tercih eden *Gly* arayüzeyde olduğu zaman monomerler arası etkileşimde önemli bir rol oynamaktadır. Lipid ortamdaki arayüzeylerde *Gly*'den sonra en yüksek kontak derecesine sahip rezidü *His*'dir. *His* bileşim oranı en düşük olan (%2-3) rezidülerden olup özellikle iç ve arayüzeylerde nadir görülmektedir. Ama kontakt derecesi bakımından ikinci sırada bulunması dikkate değerdir, zira arayüzeyde az ama etkili etkileşimlerde bulunduğunu işaret etmektedir. Bu bulgu, arayüzey tahmin algoritmaları için dikkate değerdir, zira yüzey bölge olası arayüzey tahmini için tarandığında bir *His* rezidüsüne nadir de olsa rastlanması o bölgenin arayüzey olma olasılığını arttıracak yönde olacaktır. Dikkate değer bir diğer rezidü ise *Cys*'dir. *Cys* %2 gibi çok az miktarlarda bulunmasına rağmen çoğunlukla iç bölgelerde bulunma eğilimindedir, zira tuz köprülerinin kurulmasında rol oynamaktadır. Ancak Şekil 10'da görüldüğü gibi kontakt derecesi arayüzey lipid ortamda olduğunda oldukça düşük, periferik ortamda ise yüksek bulunmuştur.

Arayüzeyde etkileşim içindeki *ij* rezidü çiftleri yöntem kısmında anlatıldığı gibi gözlenen (*observed*) kontakt frekansı KP_{ij} cinsinden 12 no'lu denklemde verildiği gibi hesaplanmıştır. Buna göre arayüzeydeki kontakt kuran toplam rezidü çiftlerinin sayısına orantılanarak yüzde cinsinden ifade edilmiştir. Şekil 11'de hem lipid hem de periferik ortamlardaki arayüzeyler için (a) ve (b)'deki gibi gösterilmiştir. Lipid ortamda, toplam kontak sayısının yaklaşık %20'si *Leu-Leu*, *Leu-Phe*, *Leu-Ile*, *Leu-Val* ve *Leu-Ala*'den oluşmaktadır. Hidrofobik rezidüler arası oluşan etkileşimlerin lipid ortamdaki arayüzeylerde polar rezidüleri göre oldukça baskın olduğu görülmektedir. Periferik arayüzeylerde ise, polar ya da hidrofobik ayrımı yapmak güçleşmiştir. Buna rağmen, lipid ortamda en fazla görülen bu beş hidrofobik rezidü çiftinin periferik ortamdaki arayüzeyde toplam kontak sayısının %7-8'ini oluşturduğu görülmüştür. Periferik arayüzeyde en fazla sıklıkta görülen polar rezidü *Arg* olmuştur. *Arg* ya bir hidrofobik rezidü olan *Leu* ile ya da polar *Glu* ile kontakt kurmuştur, ancak bunların sayısı toplam kontakt sayısının en fazla %1'ini oluşturmaktadır.

Yöntem kısmında son adımda anlatıldığı gibi (*observed*) kontakt frekansı KP_{ij} , beklenen (*expected*) kontakt frekansına bölünerek ve 2 tabanında logaritması alınıp 2 ile çarpılarak bir skor değerine dönüştürülmüştür. Bu değer, 0'dan farklı olması, arayüzeyde gözlemlenen *ij* rezidü çiftinin rastlantısal bir şekilde biraraya gelmediğini, pozitif değerlerin iki rezidünün kontakt kurmayı tercih ettiklerini, negatif değerler ise etmediklerini işaret etmektedir. Bunun sonucunda daha kararlı bir arayüzey oluşumuna olanak verilmiş ve evrimsel açıdan bakıldığında organizma için avantajlı olmuştur. Skor değeri ($2 \times \log_2$) cinsinden ifade

edildiğinden bu değerin örneğin 2, 4, ya da 6 olması, ij rezidülerin rastlantısal bir oluşumun iki, dört ya da sekiz kat üstünde bir olasılığa sahip olduğunu ifade etmektedir. Şekil 11c ve 11d'de gösterildiği gibi lipid ve periferik arayüzeyler için skor değerleri ayrı ayrı hesaplanmıştır. Buna göre lipid arayüzeylerde, polar-polar rezidü çiftlerinin etkileşme olasılıklarının rastlantısal bir durumdan 6 kat fazla olduğu görülmüştür. Bu durum polar rezidülerin arayüzeylerde az sayıda olmalarına rağmen etkileşim için o bölgede bulduklarını açıkça göstermektedir. Diğer taraftan ilginç olan durum, hidrofobik rezidülerin kontak frekansları yüksek olmasına rağmen rastlantısal bir modelden farksız davranmalarıdır. Sayıca çok olmaları, her hidrofobik rezidüsünün etkileşim içinde rol oynayacağı anlamına gelmemektedir. Bu durum, Şekil 10'da verilen kontak derecelerini gösteren profilden de anlaşılmaktadır.



Şekil 11. (a),(c) lipid ve (b),(d) periferik ortamlardaki arayüzeyde gözlenen (*observed*) kontak frekansı KP_{ij} ve skor değerleri

Farklı bir durum, Şekil 11d'de gösterilen periferik arayüzey için olan skor değerleri haritasında mevcuttur. Neredeyse tüm rezidüler rastlantısal modelin de altında nadir görülen etkileşimler sergilemektedirler. Diğer taraftan, *Cys* rezidüsü diğerlerinden oldukça farklı bir profile sahiptir. Kendi kendisiyle olan skor değeri (5.5) maksimum değer olup, *His*, *Val*, *Pro*, *Tyr*, *Trp*, *Glu*, gibi çoğunlukla hidrofobik olan rezidülerle yüksek etkileşim skor değerlerine sahiptir. *Cys*'den sonra, kendi kendisiyle yüksek etkileşim içinde olan rezidü *Trp*'dir (skor değeri: 4.3).

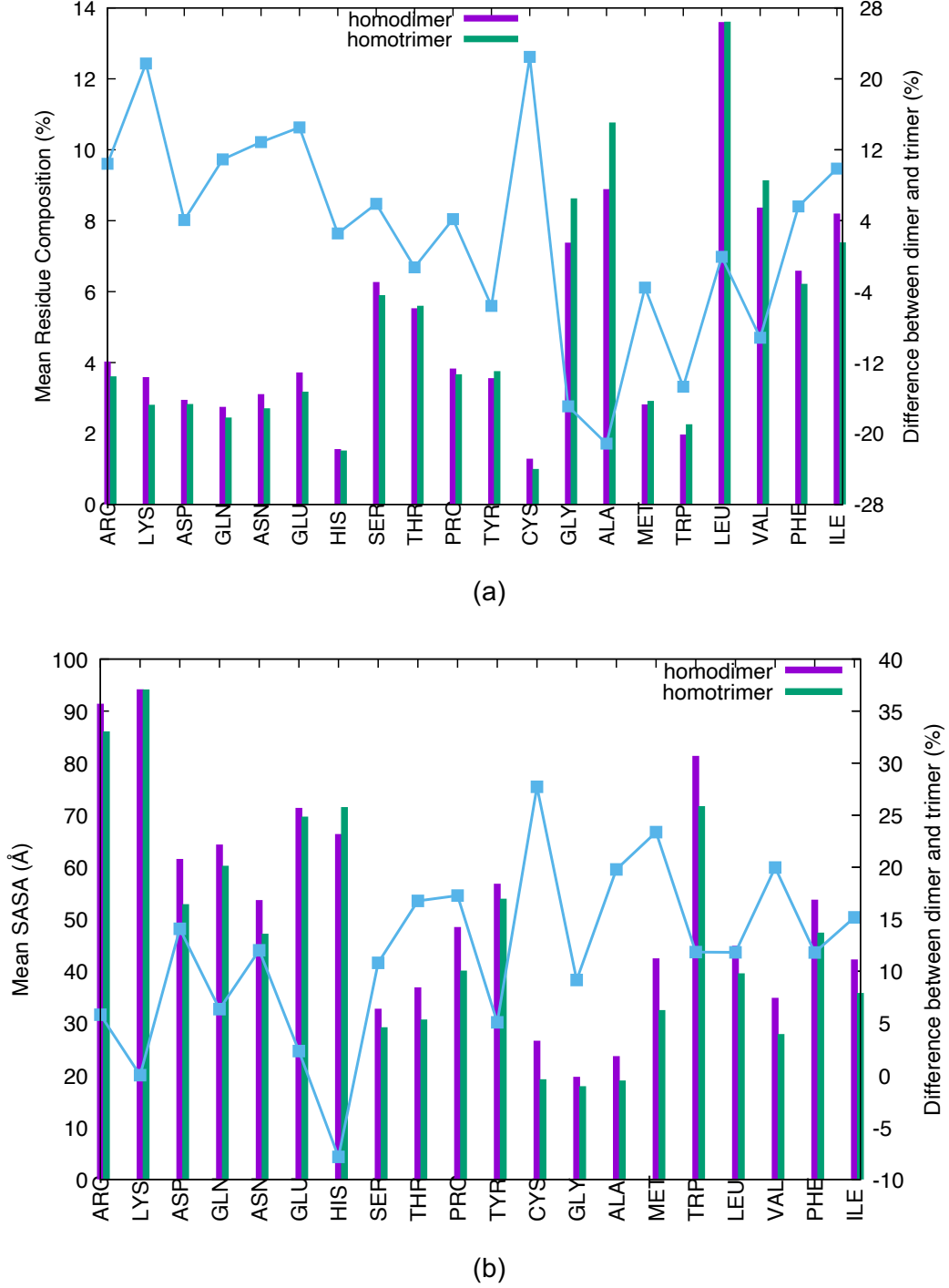
Lipid arayüzey bölgede en yüksek skor değerine sahip rezidü *Trp* olmuştur ve *Trp* rezidüsünün geçtiği *Trp-Met*, *Trp-Thr*, *Trp-Trp*, *Trp-Ile*, *Phe-Cys*, *Trp-Phe*, *Ile-Ile*, *Phe-Phe* ve *Trp-Val* çiftleri belirlenmiştir. *Trp* aynı zamanda lipid ortamdaki arayüzeyde pozitif skor değerine sahip tek rezidüdür. Hidrofobik rezidülerden hemen her yönden farklı bir konumdadır. *Trp* miktar olarak proteinlerde *Cys* ve *His* rezidüleri kadar nadir görülmektedir (%2 bileşim miktarı). Diğer taraftan hidrofobik rezidüler arasında kenar ve arayüzey kenar bölgelerde en yüksek bulunma eğilimi (propansite) gösteren rezidüdür. Bu önemli bulgu lipid arayüzey tahminlerinde kullanılacak niteliktedir zira, *Trp* arayüzeyde az görülmesine rağmen tamamının kontakt halde olması (Bknz. Şekil 10; kontakt derecesi=1.01) ve yüksek etkileşim skor değerleri bu rezidünün lipid yüzeyde bulunması halinde, bulunduğu bölgenin arayüzey olma ihtimalini artırıcı yönde olacaktır.

4.5. Homotrimerik ve Homodimerik Alfa-sarmal membran protein komplekslerin karşılaştırmalı analizi

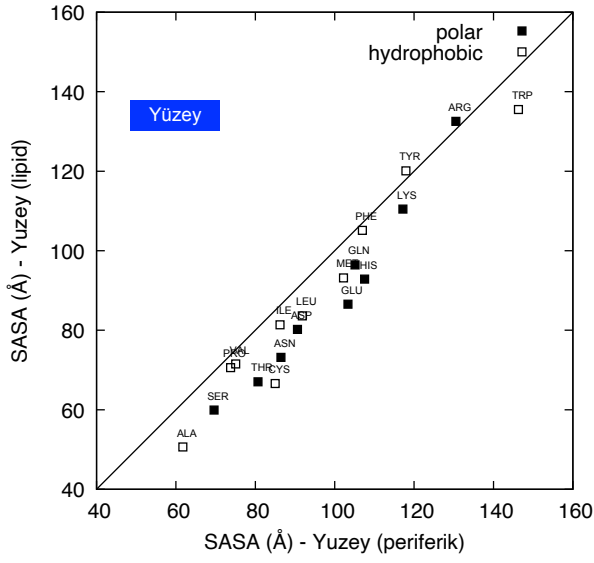
Bu projede, kompleks yapıyı oluşturan monomerik birim sayısı amino asit dağılımını ve etkileşimini etkileyen faktörlerden biri olarak sunulmuştur. Dolayısıyla, dimerik yapıların yanı sıra toplam 60 alfa-sarmal homotrimerik yapı kendi içinde benzer bir analize tabii tutulmuş ve dimerik yapı sonuçları ile karşılaştırılarak gösterilmiştir. İlk sonuçlar monomerik birim sayısının amino asit dağılımına olan etkisini açıkça göstermektedir, zira Şekil 12a'daki ortalama rezidü bileşim yüzdesi ve Şekil 12b'deki ortalama SASA değerlerindeki farklılıklar son derece belirgindir. Bileşim yüzdelere bakıldığında polar rezidülerin dimerik yapılarda trimerlere göre %10-20 daha fazla olduğu, *Ala* ve *Trp* gibi hidrofobik rezidülerin ise trimerik komplekslerde daha fazla görüldüğü tespit edilmiştir. Diğer taraftan ortalama SASA değerlerinin dimerik yapılarda trimerik yapılara göre %10-30 mertebelerinde daha fazla olduğu görülmektedir.

SASA değerleri kompleks yapının farklı bölgelerine göre incelendiğinde dimerik yapılarda olduğu gibi periferik ve lipid ortamlarda belirgin farklılıklar olduğu görülmektedir (Şekil 13). Bu farklılıklar özellikle çekirdek ve destek arayüzeylerde yüzey ve kenar arayüzey bölgelere göre

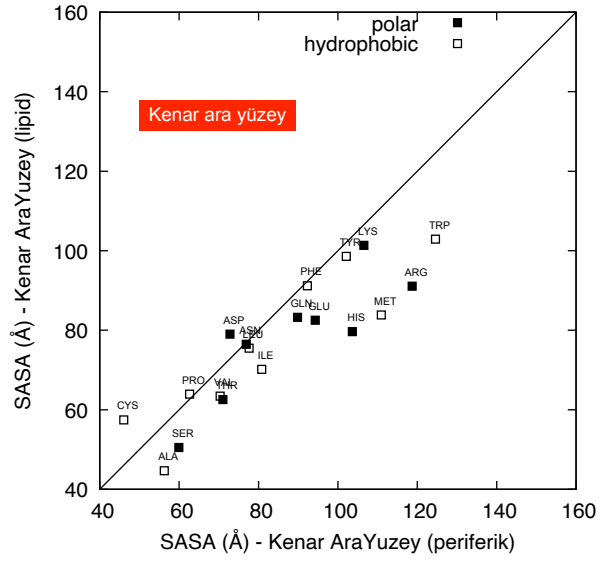
daha çarpıcıdır. Periferik ortamdaki çekirdek ve destek arayüzeylerdeki çok sayıda rezidünün ortalama SASA değerlerine bakıldığında, lipid ortamdaki benzer bölgelere karşılık gelen aynı rezidülerin SASA değerlerinden daha yüksek oldukları belirgin bir şekilde görülmektedir (Şekil 13c ve 13d). Ayrıca bu farklılığın trimerik kompleks yapılarda dimerik yapılara göre daha fazla olduğu belirlenmiştir.



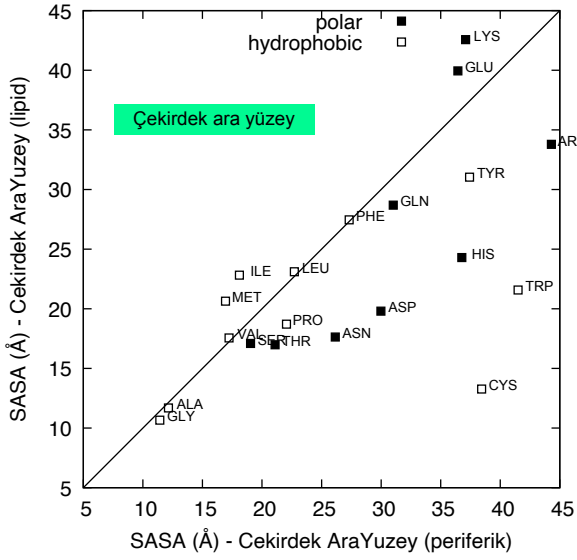
Şekil 12. Dimerik ve trimerik alfa sarmal yapılarda rezidülerin ortalama (a) bileşim yüzdeleri ve (b) SASA değerleri



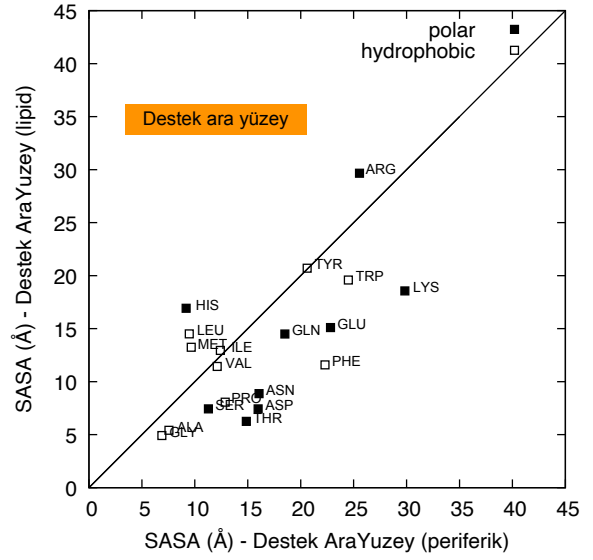
(a)



(b)



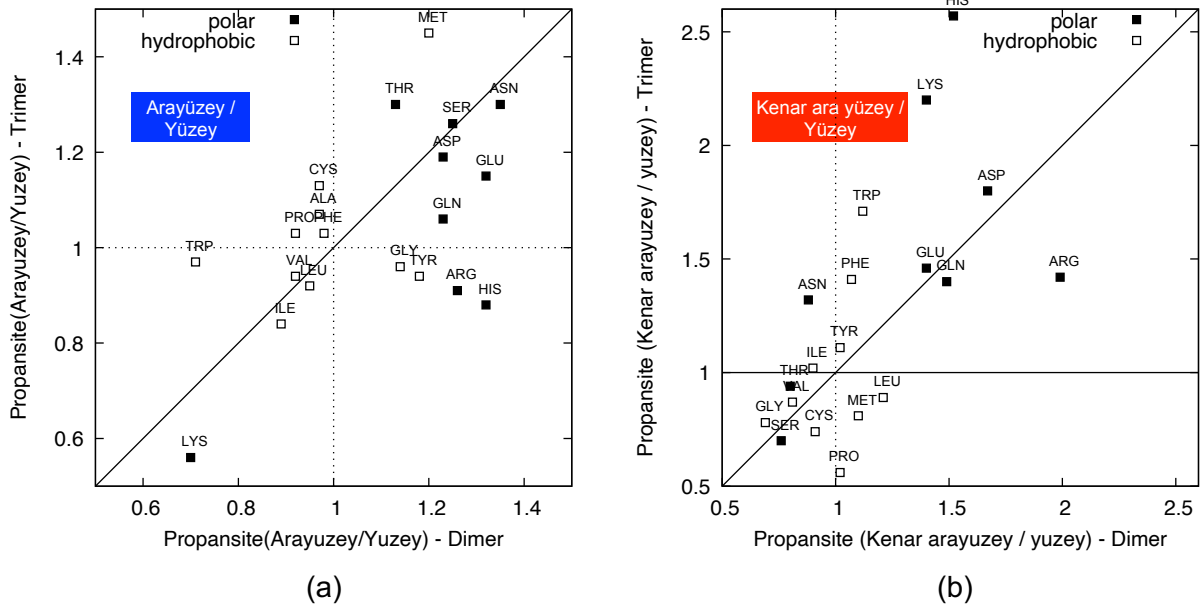
(c)



(d)

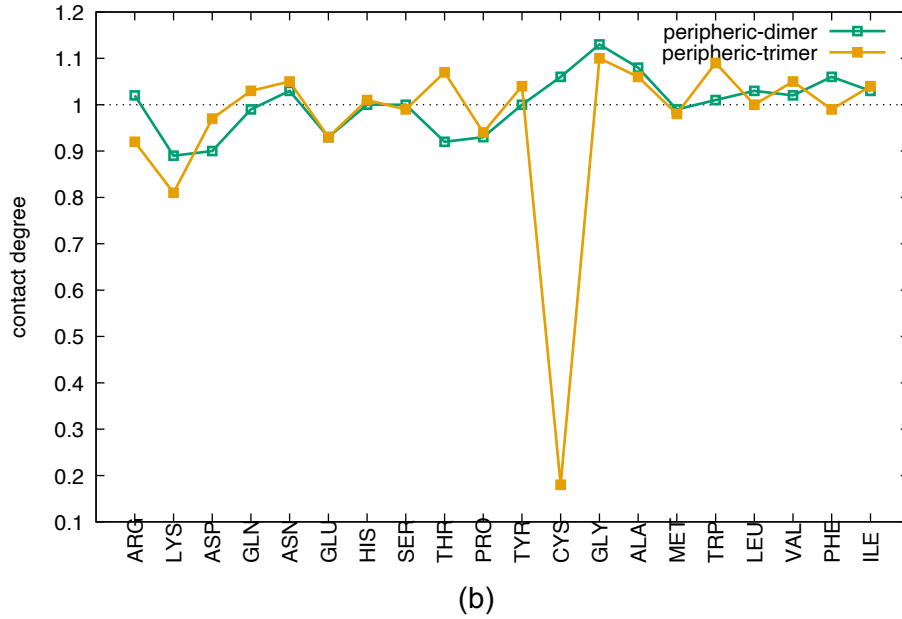
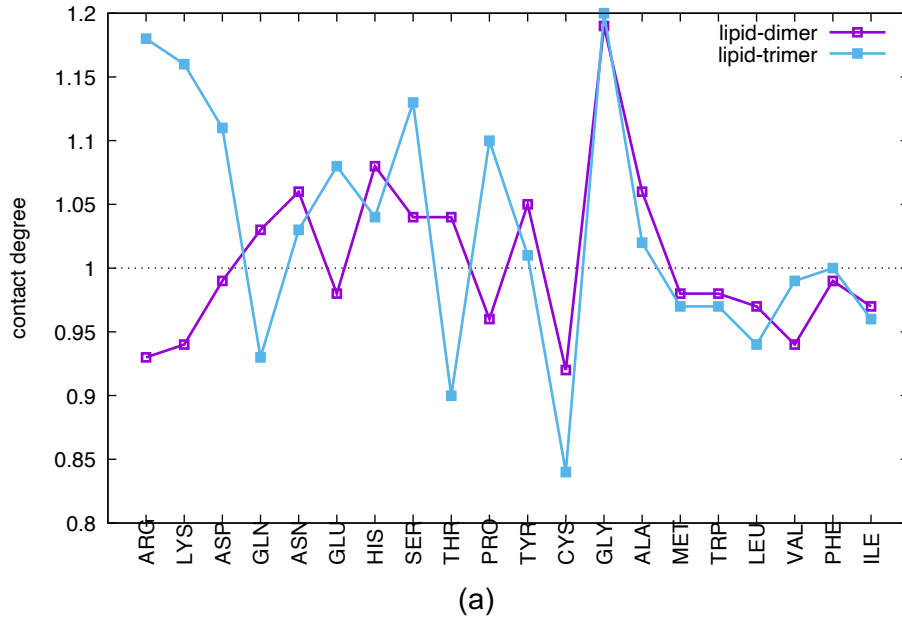
Şekil 13. 60 alfa-sarmal homotrimerik protein kompleksinde rezidülerin farklı bölgelerdeki ortalama SASA değerleri. (a) yüzey, (b) kenar arayüzey, (c) çekirdek arayüzey ve (d) destek arayüzey.

Propansite eğilim değerleri dimerik yapılarla karşılaştırıldığında, özellikle polar rezidüler için belirgin farklılıklar ortaya çıkmaktadır. Şekil 14a'da, arayüzey bölgenin tüm lipid yüzey bölgeye göre çoğunlukla polar rezidüler tarafından daha fazla tercih edildiği, bu eğilimin dimerik yapılarda trimerik yapılara göre daha fazla olduğu gösterilmiştir. Diğer taraftan, kenar arayüzey bölgeler özellikle polar rezidülerin diğer yüzey bölgelere göre daha fazla yoğunlukta olduğu bölgeler olmuştur. Diğer taraftan, *Phe* ve *Trp* hariç diğer hidrofobik rezidüler için yüzey bölgeler kenar arayüzeye göre daha fazla tercih edilmiştir.



Şekil 14. 60 alfa-sarmal homotrimerik protein kompleksinde rezidülerin farklı bölgeler için propansite (eğilim) değerleri. (a) arayüzey / yüzey, (b) kenar arayüzey / yüzey.

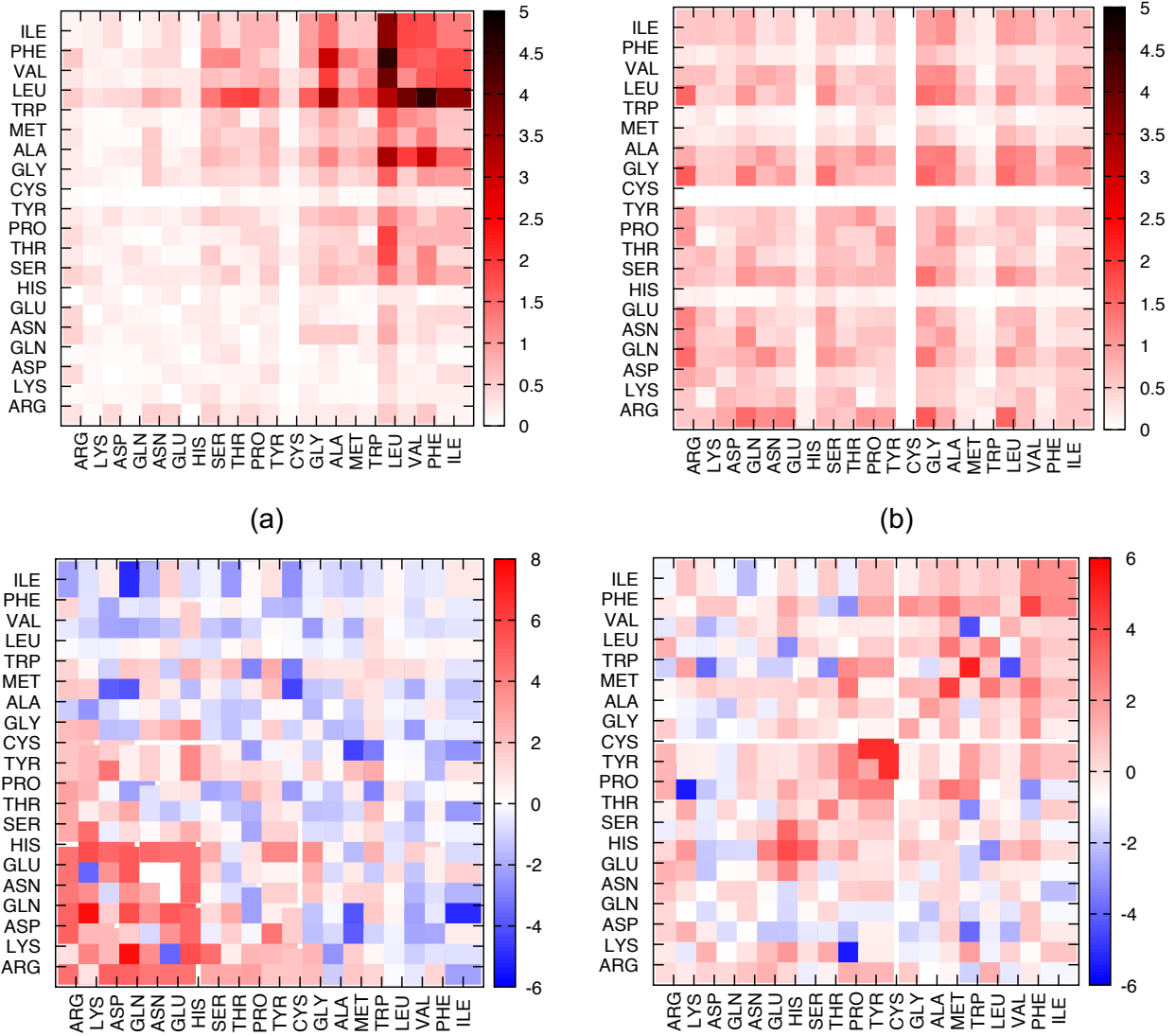
Alfa-sarmal homo dimerik ve trimerik yapılardaki önemli bir diğer fark kontakt derecelerinde belirlenmiştir. Şekil 15'te lipid ve periferik ortamlar için gösterildiği gibi, farklılık en fazla lipid ortamdaki arayüzey bölge için görülmektedir. Özellikle *Arg* ve *Lys* gibi uzun yan zincirlere sahip polar rezidülerin kontakt dereceleri dimerik yapıda düşükken, trimerik yapıda belirgin derecede artmıştır (Bknz Şekil 15a). Benzer bir artış, *Glu*, *Ser* ve *Pro* için de görülmektedir. Önemli bir bulgu da, her iki tür kompleks yapı için de en yüksek kontakt derecesine sahip rezidünün *Gly*, en düşük kontakt derecesine sahip olanın ise *Cys* olmasıdır. Diğer taraftan, periferik ortamdaki duruma bakıldığında, farklılık derecesi düşüktür, ancak *Cys* trimerik komplekslerdeki arayüzey bölgelerde neredeyse yok denecek kadar az miktarda etkileşim içinde görülmektedir.



Şekil 15. Dimerik ve trimerik yapılar, rezidülerin (a) lipid ve (b) periferik ortamda bulunan arayüzey bölgelerdeki kontakt dereceleri.

Önemli bir diğer sonuç Şekil 16'da gösterilmiş olan lipid ve periferik ortamlardaki arayüzeylerde gözlemlenen (*observed*) kontakt frekansı $KP_{i,j}$ ve skor değerleridir. En büyük farklılık trimerik yapılar *Cys* rezidüsünün gözlemlenmiş bir etkileşimi olmadığından diğer rezidülerle kontakt frekansı sıfırdır. Daha da önemlisi sıfır kontakt frekansı herhangi bir skor değerine dönüştürülemediğinden tanımsızdır, dolayısıyla Şekil 16c ve 16d'de ince beyaz çizgilerle gösterilmektedir. Diğer taraftan trimerik yapıların dimerik yapılarla bazı benzerlikleri görülmektedir, örneğin, $KP_{i,j}$ değerlerini gösteren grafiklerde hidrofobik rezidüler lipid ortamdaki arayüzeylerde daha fazla sayıda etkileşime girerken, polar rezidülerin kontakt

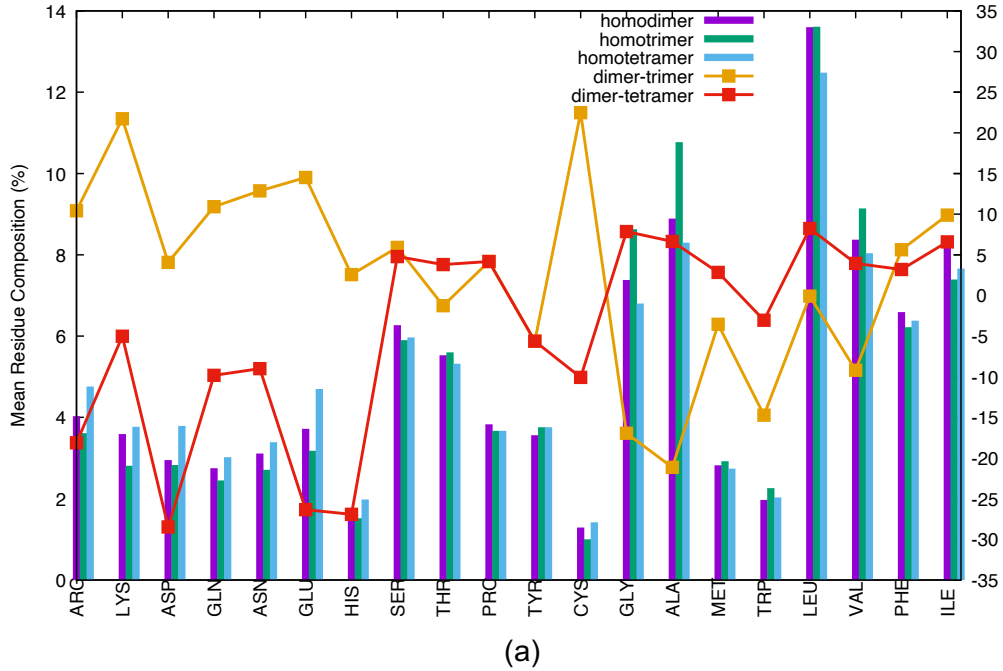
frekansları sıfıra yakındır. Periferik ortamdaki arayüzeylerde *His*, *Cys*, *Trp* ve *Met* rezidüleri dimerik yapılarda olduğu gibi az sayıda bulunurken farklı olarak *Phe* rezidüsü trimerik yapılarda az sayıda etkileşim içinde görülmektedir. İlginç bir sonuç ta, *Cys* rezidüsü içindir, zira arayüzeylerde oldukça az sayıda olmalarına rağmen, periferik ortamdaki arayüzeylerde bulunduğu kendisiyle ve *Tyr* rezidüsüyle oldukça yüksek skor değerleriyle etkileşime girmektedir. Benzer bir durum dimerik yapılarda da mevcuttur, ama *Tyr* yerini *His* almıştır. Dimer ve trimer yapılarıdaki en büyük farklılık periferik arayüzeylerdeki skor değerlerinde görülmüştür. Şekil 16d ve 11d karşılaştırıldığında, trimerik yapılarda, *Cys* dışında *Trp* ve bunun gibi birçok hidrofobik ve polar rezidü etkileşim içinde görülmektedir. Diğer taraftan, *Cys* dışındaki diğer tüm rezidüer dimerik yapılarda rastlantısal modelin çok altında bir sıklıkta etkileşmektedir.

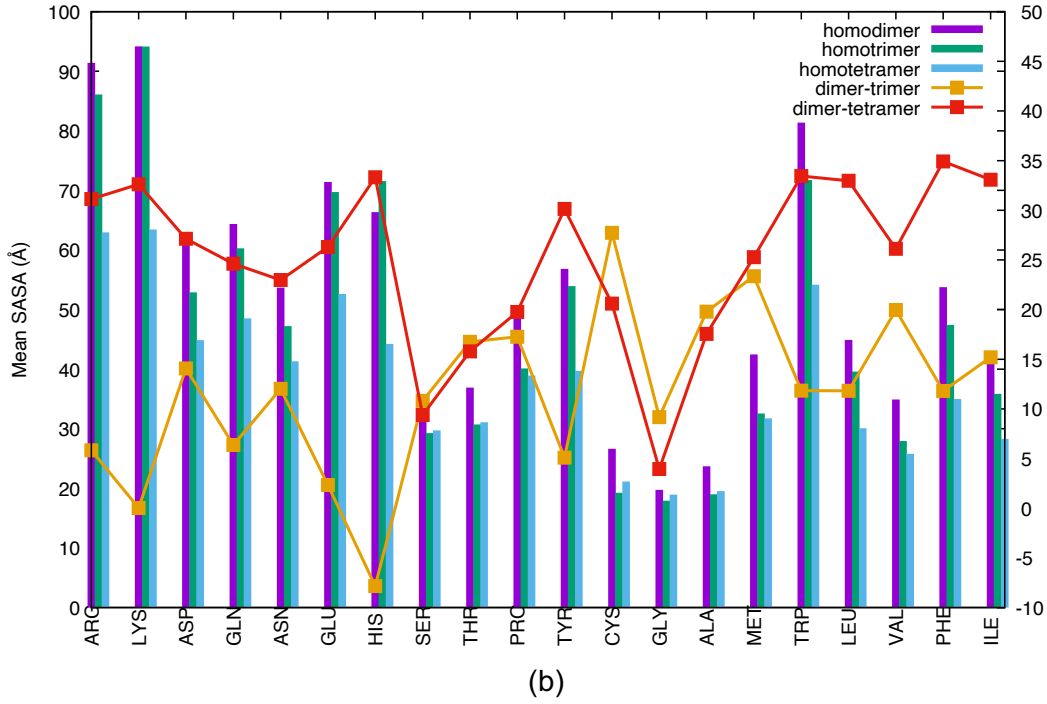


Şekil 16. Alfa-sarmal trimerik yapılarda (a),(c) lipid ve (b),(d) periferik ortamlardaki arayüzeyde gözlenen (*observed*) kontakt frekansı $KP_{i,j}$ ve skor değerleri

4.6. Homotetramerik, Homotrimerik ve Homodimerik Alfa-sarmal membran protein komplekslerinin karşılaştırmalı analizi

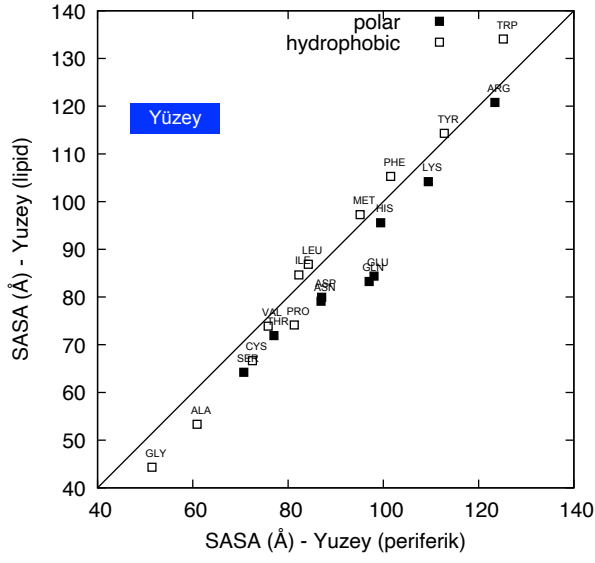
Membran protein komplekslerindeki bir diğer farklı oluşum homotetramerik yapılarda görülmektedir. Veri setimizde bulunan toplam 99 tetramerik yapıya uygulanan benzer bir analiz sonucunda farklılıklar, dimerik ve trimerik yapılarla karşılaştırılmıştır. Şekil 17a'da gösterildiği gibi bileşim yüzdelerindeki farklılık en fazla polar rezidülerde görülmektedir. Polar rezidüler en fazla tetramer yapılarda görülürken en az sayıda trimer yapılarda görülmektedir. Diğer taraftan hidrofobik rezidüler polar rezidülere yüksek ancak monomerik oluşumlar arasında en düşük seviyede tetramerik yapılarda bulunmaktadır. Şekil 17b'de SASA değerleri karşılaştırıldığında tetramerik yapılarda belirgin bir düşüş gözlemlenmektedir. Özellikle dimer ve tetramerik yapılarda arasındaki fark belirgin düzeyde yüksektir.



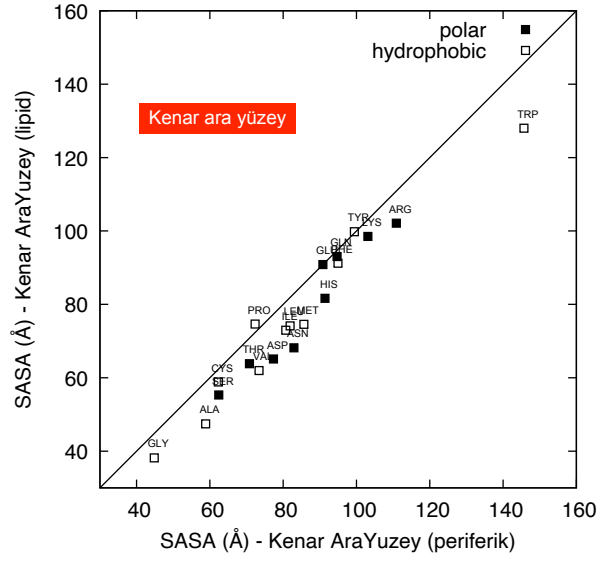


Şekil 17. Dimerik, trimerik ve tetramerik alfa sarmal yapılarında rezidülerin ortalama (a) bileşim yüzdeleri ve (b) SASA değerleri

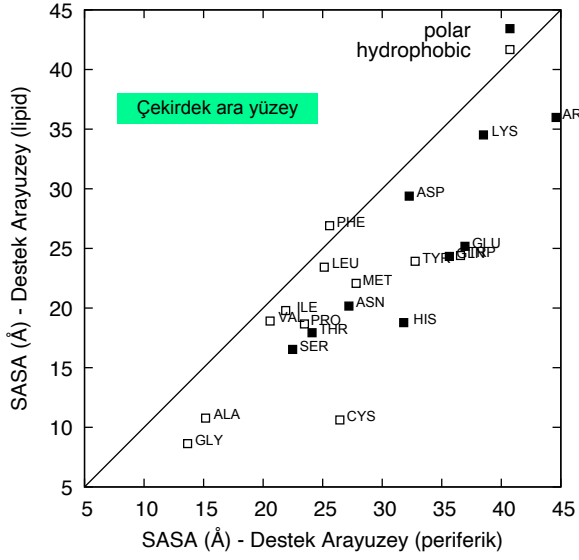
Periferik dış ortam ve lipid ortamdaki arayüzeylerde görülen yapısal farklılıklar tetramerik yapılarında da görülmektedir. Şekil 18'de gösterildiği gibi özellikle periferik ortamdaki çekirdek ve destek arayüzeylerde rezidülerin ortalama SASA değerlerinin lipid ortamdakine göre daha fazla olduğu ortaya çıkmıştır. Benzer bir durum Şekil 19'da gösterilen propansite değerlerinde de bulunmaktadır. Arayüzey bölgenin yüzey bölgeye göre daha fazla tercih edilmesini ifade eden 1'den büyük propansite değerleri polar rezidüler için periferik ortamda lipid ortama göre daha yüksek seviyelerdedir. Aynı şekilde kenara arayüzeylerin yüzey bölgelere göre polar rezidüler tarafından tercih ediliyor olması da, periferik ortama girince daha belirgin hale gelmektedir. Genel olarak polar rezidüler yüzey bölgelerde az sayıda olsalar da bulduklarında bu bölgenin bir arayüzey oluşumu olma ihtimali yükselmektedir. Dolayısıyla propansite değerleri, arayüzey tahmin algoritmalarında kullanılacak önemli bilgi kaynakları haline gelmektedir.



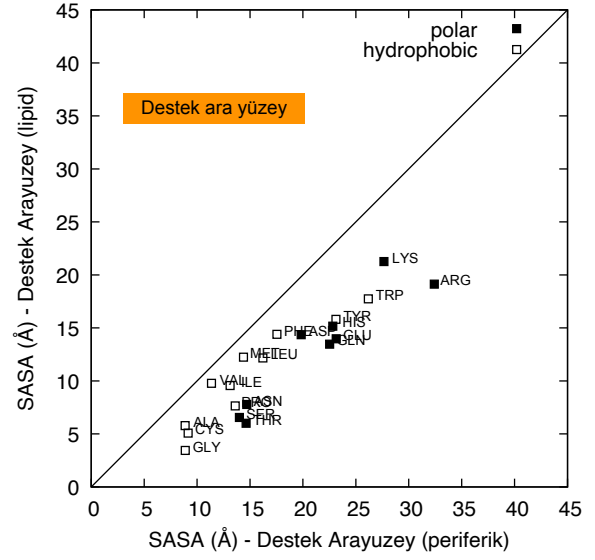
(a)



(b)

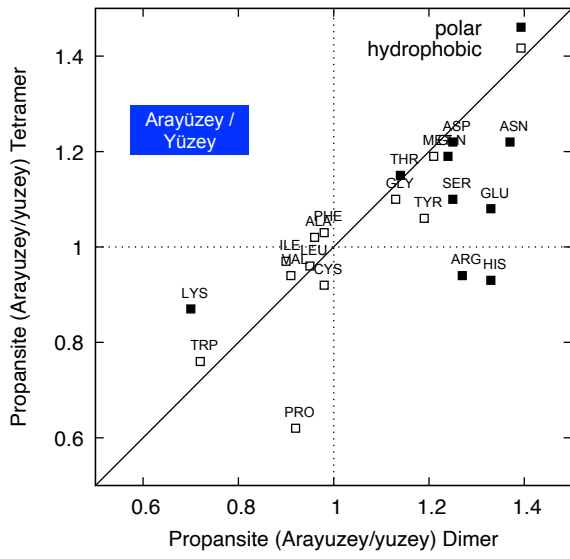


(c)

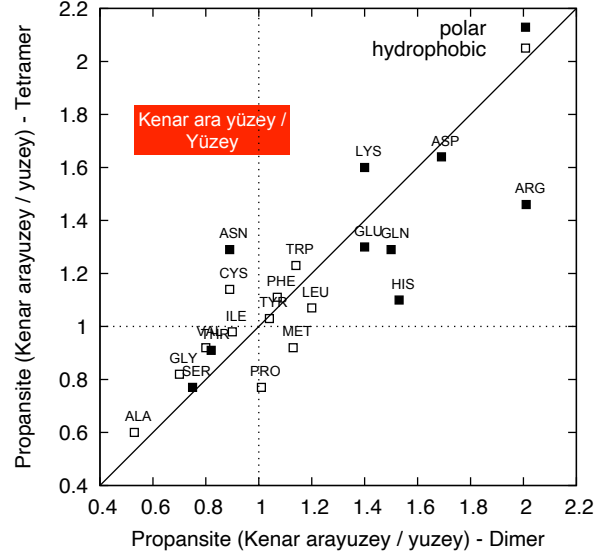


(d)

Şekil 18. 99 alfa-sarmal homotetramerik protein kompleksinde rezidülerin farklı bölgelerdeki ortalama SASA değerleri. (a) yüzey, (b) kenar arayüzey, (c) çekirdek arayüzey ve (d) destek arayüzey.



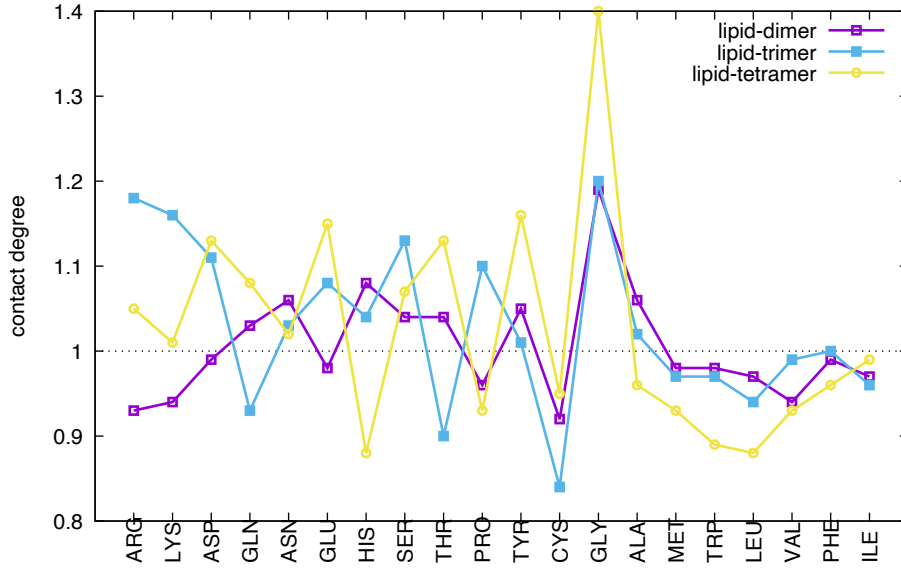
(a)



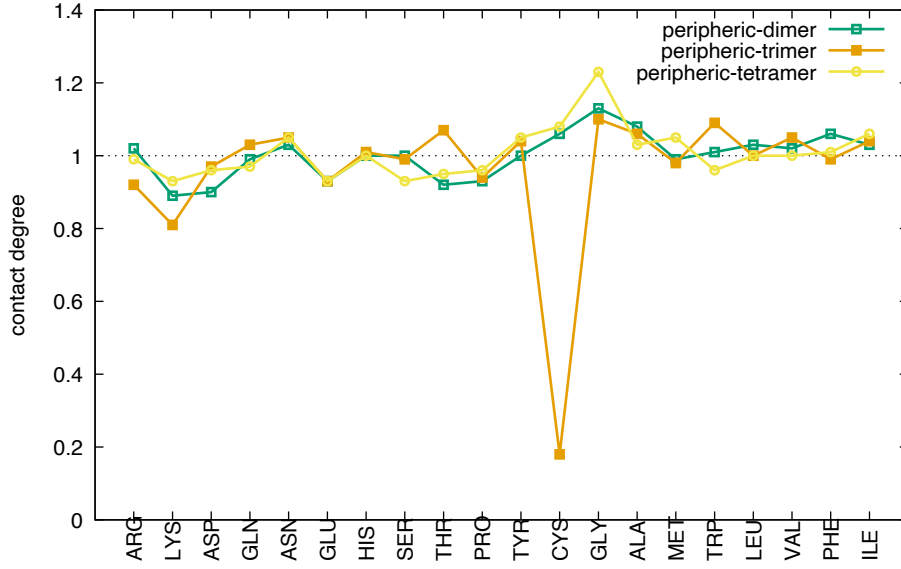
(b)

Şekil 19. 99 alfa-sarmal homotetramerik protein kompleksinde rezidülerin farklı bölgeler için propansite (eğilim) değerlerinin dimerik yapılarıdaki propansite değerleriyle karşılaştırılması. (a) arayüzey / yüzey, (b) kenar arayüzey / yüzey.

Kontakt dereceleri tetramerik yapılarda bakıldığında Şekil 20'deki gibi farklılıklar ortaya çıkmıştır. Lipid ortamdaki arayüzeylerde Gly rezidüsü 1.4'lük bir değerle kategoriler arasında görülen en yüksek kontakt derecesi olmuştur. Diğer taraftan hidrofofik rezidülerin kontakt dereceleri tetramerik yapılarda, dimer ve trimerlere göre az da olsa bir miktar daha düşük seviyelerdedir. Diğer taraftan polar rezidüler (His, Pro ve Cys gibi özel statüdeki rezidüler hariç) 1'in üstünde değerler sergilemektedir. Periferik arayüzeylerde ise özellikle dimer yapılarla benzerlik taşımaktadır ve lipid ortamdakine benzer şekilde Gly periferik ortamda görülen en yüksek kontakt derecesine sahiptir.



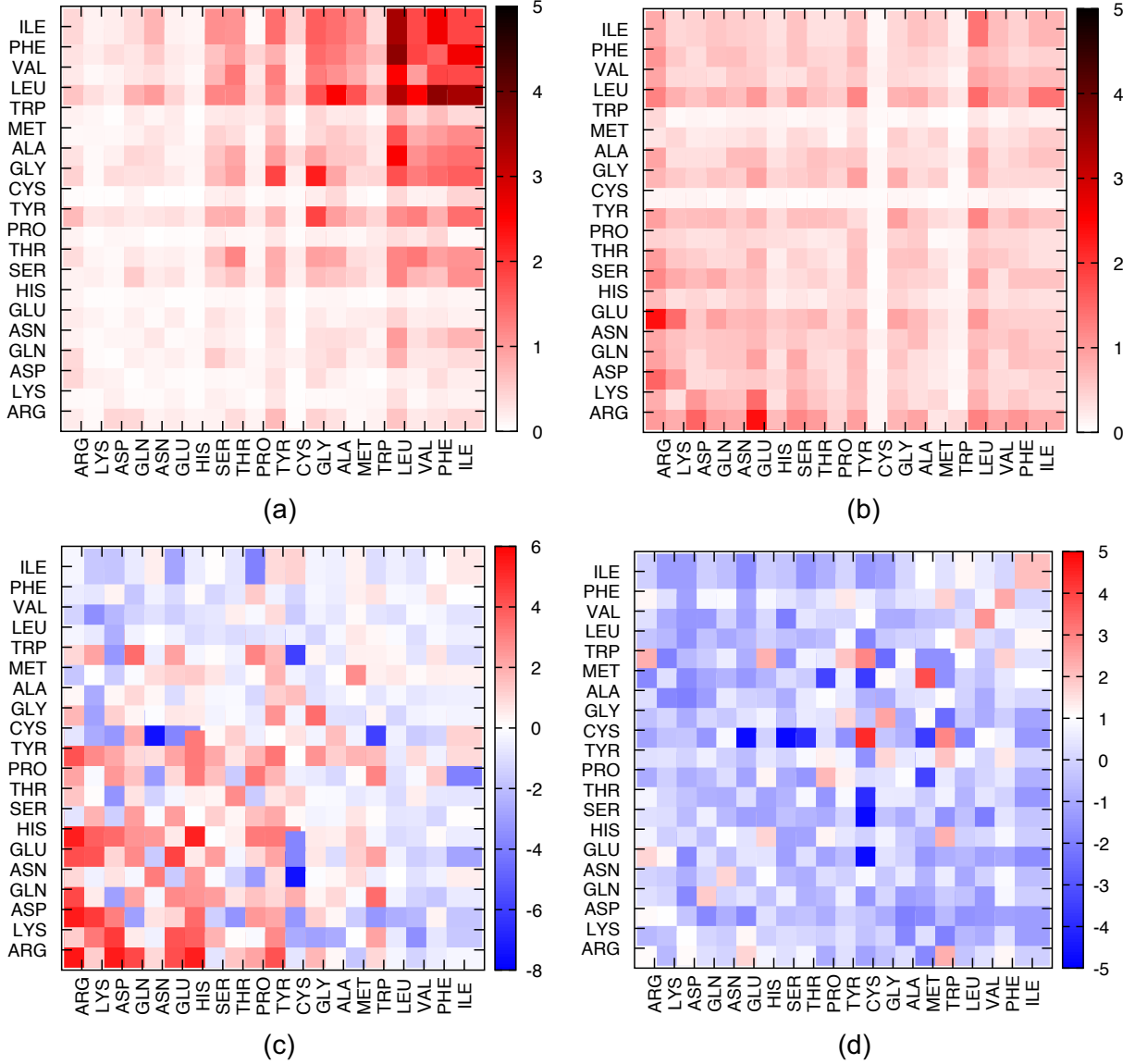
(a)



(b)

Şekil 20. Dimerik, trimerik ve tetramerik yapılar da, rezidülerin (a) lipid ve (b) periferik ortamda bulunan arayüzey bölgelerdeki kontakt dereceleri.

Son olarak tetramerik yapılar da lipid ve periferik ortamlar da arayüzeylerde gözlemlenen (*observed*) kontakt frekansı $KP_{i,j}$ ve skor değerlerine bakıldığında dimerik yapıdakilerle büyük bir benzerlik taşıdığı görülmektedir. Kontakt frekansının lipid arayüzeylerde hidrofobik rezidüler arasında yüksek, polar rezidüler arasında düşük olması, periferik ortamda ise homojen bir dağılım göstermesi göze çarpan ilk benzerliktir. Skor değerlerine bakıldığında lipid arayüzeylerde polar rezidüler arasında skor değerlerinin yüksek olması dimerik yapılar da görülmüştür. Periferik arayüzeylerdeki skor değerleri tetramerik yapılar da da düşüktür, ancak dimerik yapılar da kadar düşük negatif değerler değildir.

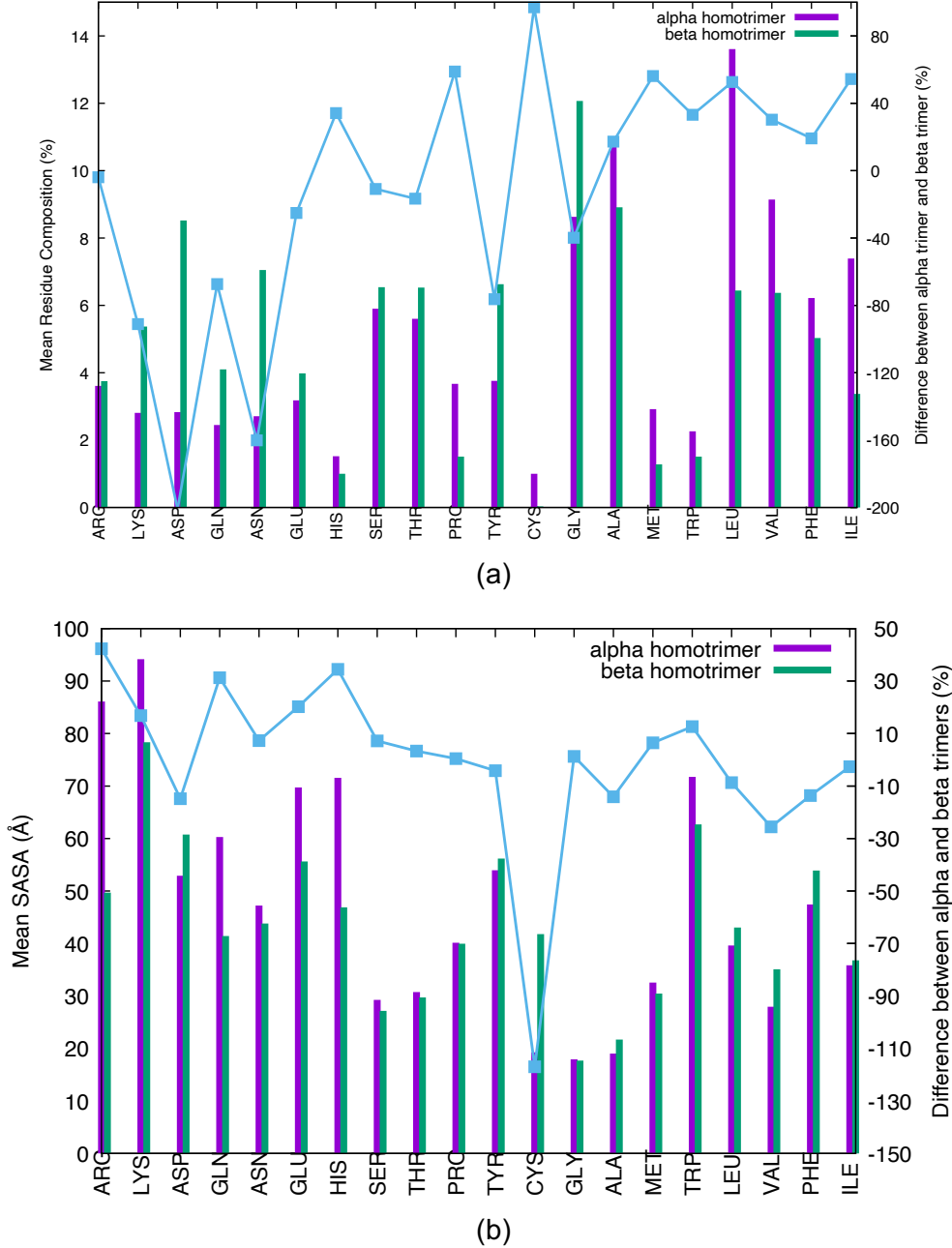


Şekil 21. Tetramerik yapıarda (a),(c) lipid ve (b),(d) periferik ortamlardaki arayüzeyde gözlenen (*observed*) kontakt frekansı $KP_{i,j}$ ve skor değerleri

4.7. Beta-silindir ve Alfa-sarmal membran protein komplekslerin karşılaştırmalı analizi

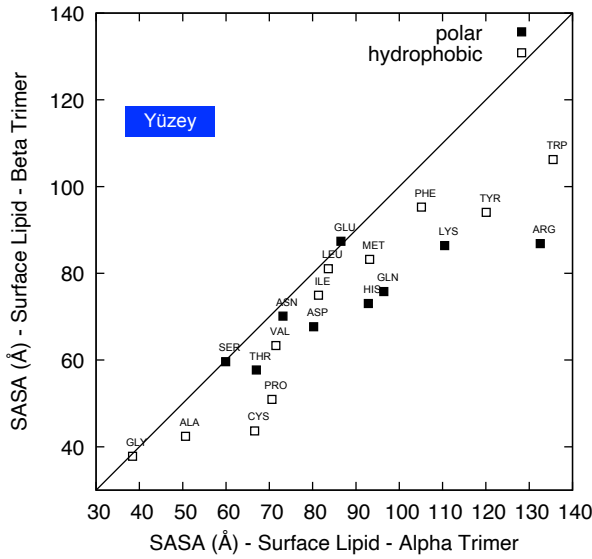
Son olarak sekonder yapının arayüzey bölgesine ve genel kompleks yapıdaki rezidü dağılımına olan etkisini ortaya çıkartmak için veri setinde az sayıda bulunan 39 beta-silindir yapı benzer bir analize tabii tutulmuştur. OPM veribankasında bulunan toplam 317 beta-silindir yapı elemelerden sonra 39'a düşmüştür, bunlardan 31'i homotrimer, 2'si heterodimer ve 6'sı da homodimer yapıdır. Az sayıda bulunmalarından dolayı bu kategoriler bir araya getirilerek incelenmiştir. Şekil 22'de gösterildiği gibi, alfa-sarmal trimer yapılarla

karşılaştırılmıştır, zira setin büyük kısmı trimer yapılardan oluşmaktadır. Rezidü bileşim yüzdeleri karşılaştırıldığında ilk göze çarpan farklılık polar rezidülerde olmuştur. Beta-silindir yapılarda polar rezidüler miktarı ciddi şekilde alfa-sarmaldakilerden iki üç kat fazladır. Diğer taraftan hidrofobik rezidülerin alfa-sarmal yapılarda daha fazla olduğu görülmektedir. Ortalama SASA değerleri karşılaştırıldığında, polar rezidülerin SASA değerlerinin alfa-sarmal yapılarda %20-30 daha fazla, hidrofobik rezidülerin ise %10 mertebelerinde düşük olduğu ortaya çıkmıştır.

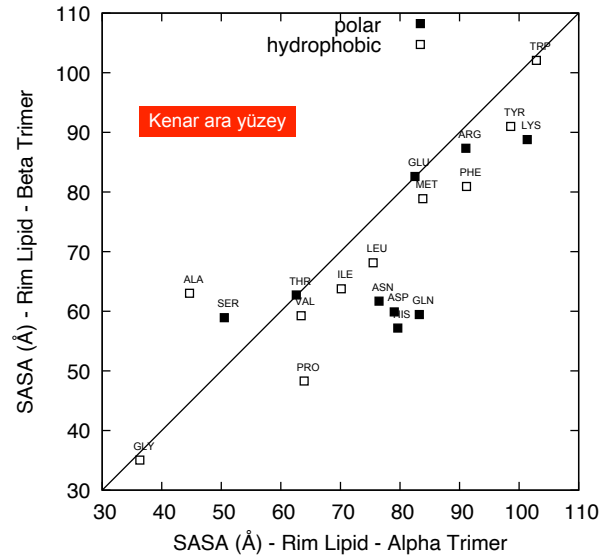


Şekil 22. Alfa-sarmal ve beta-silindir yapılarda rezidülerin ortalama (a) bileşim yüzdeleri ve (b) SASA değerleri

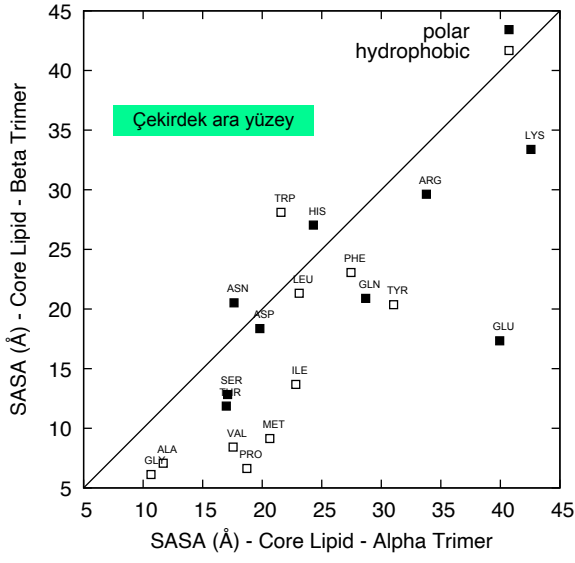
Rezidülerin ortalama SASA değerleri bölgelere göre bakıldığında, alfa-sarmal yapılarda özellikle polar rezidülerin beta-silindirlere göre daha yüksek SASA değerlerine sahip oldukları görülmüştür (Şekil 23). Bu fark yüzey ve kenar arayüzey bölgeler için daha da belirgindir. Destek arayüzeyler karşılaştırıldığında, *Lys* rezidüsü beta-silindir yapılarda görülen en yüksek SASA değerine sahipken alfa-sarmal yapılarda ortalama bir SASA değerine sahiptir. Diğer taraftan çekirdek arayüzeyde hidrofobik rezidülerin SASA değerleri polar rezidüleri göre belirgin derecede daha düşüktür. Diğer taraftan, Şekil 24'teki propansite değerleri karşılaştırıldığında, polar rezidülerin beta-silindir yapılarda alfa-sarmal yapıdakilere göre bir miktar daha fazla arayüzeyde olma eğilimi gösterdikleri ortaya çıkmıştır. Beta-silindir yapılarda bir başka göze çarpan özellik polar rezidülerin hidrofobiklere göre daha yüksek propansite değerlerine sahip olduğudur. Neredeyse tamamının propansite değerleri 1'den büyüktür. Hidrofobik rezidülerde ise tam tersi bir durum görülmektedir. Kenar arayüzey eğilim oranlarına bakıldığında daha dağınık bir profil ortaya çıkmıştır. Beta-silindir yapılarda 0-3 aralığında propansite değerlerine sahip rezidüler görülmektedir. Diğer taraftan polar rezidüler alfa-sarmal yapılarda çoğunlukla 1'den büyük değerlere sahiptir.



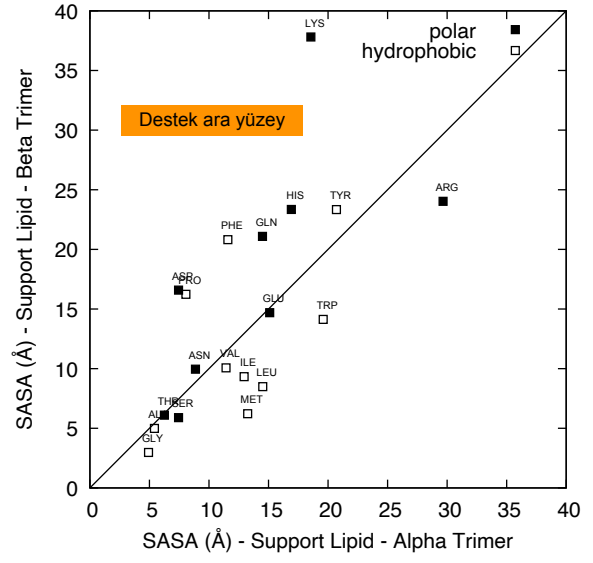
(a)



(b)

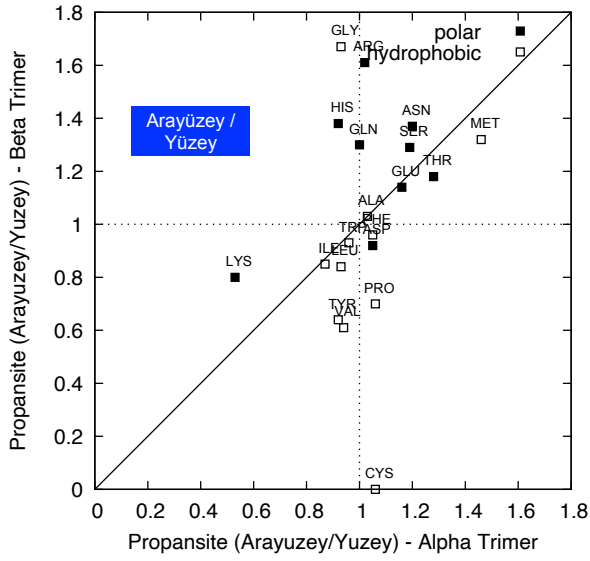


(c)

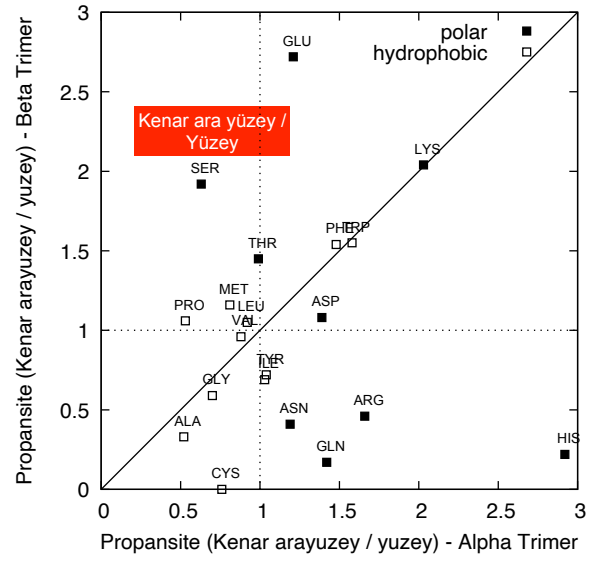


(d)

Şekil 23. 39 beta-silindir kompleks yapıdaki rezidülerin farklı bölgelerdeki ortalama SASA değerlerinin alfa-sarmal yapıdakilerle karşılaştırılması. (a) yüzey, (b) kenar arayüzey, (c) çekirdek arayüzey ve (d) destek arayüzey.



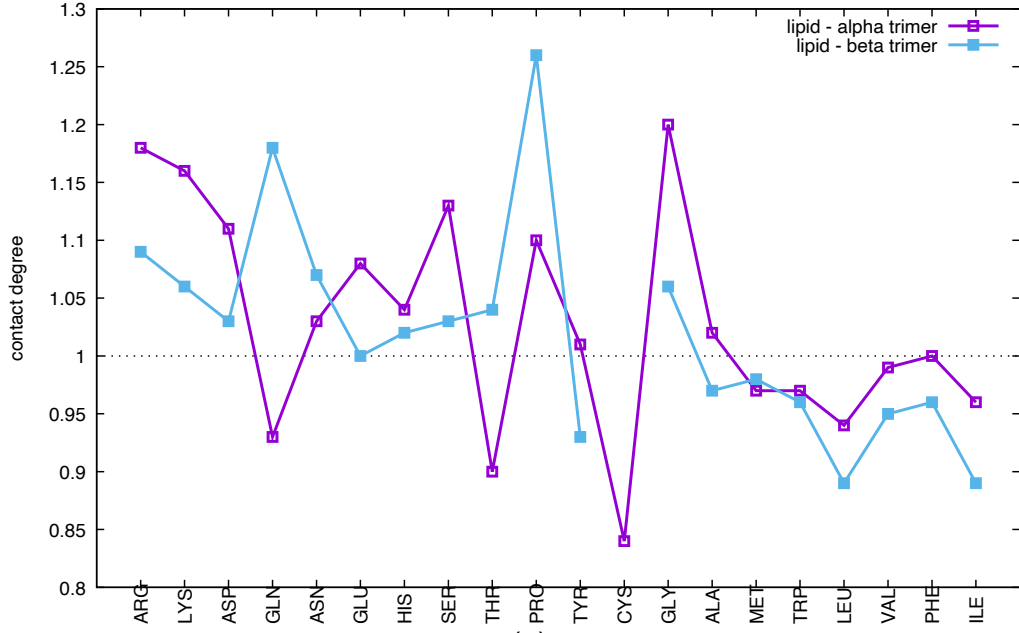
(a)



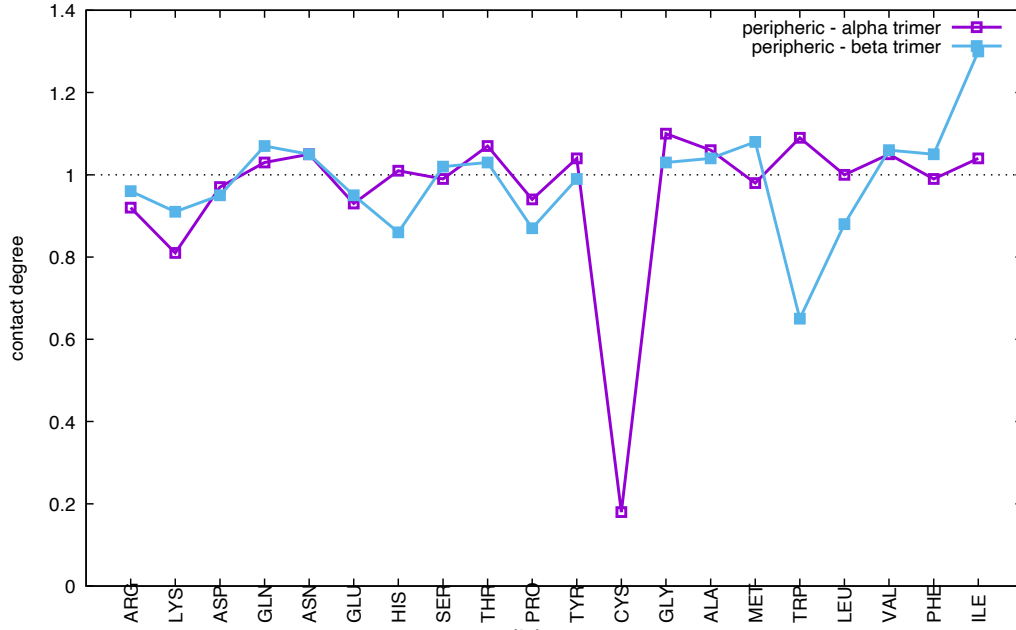
(b)

Şekil 24. 39 beta-silindir kompleks yapıdaki rezidülerin farklı bölgeler için propansite (eğilim) değerlerinin alfa-sarmal yapıdaki propansite değerleriyle karşılaştırılması. (a) arayüzey / yüzey, (b) kenar arayüzey / yüzey.

Kontakt dereceleri alfa-sarmal ve beta-silindir yapılar arasında karşılaştırıldığında farklılıklar en fazla lipid ortamdaki arayüzeyler için görülmektedir. Beta-silindir yapılarında polar rezidülerin tamamı 1'den büyük kontakt derecesine sahip olurken, hidrofobik rezidülerin tamamının kontakt dereceleri 1'den küçüktür. Periferik arayüzey ortamında, en büyük farklılık *Trp* rezidüsü için görülmüştür; alfa-sarmal yapılarında 1'den büyük bir kontakt derecesi beta-silindir yapılarında 0.6'a kadar düşmüştür.



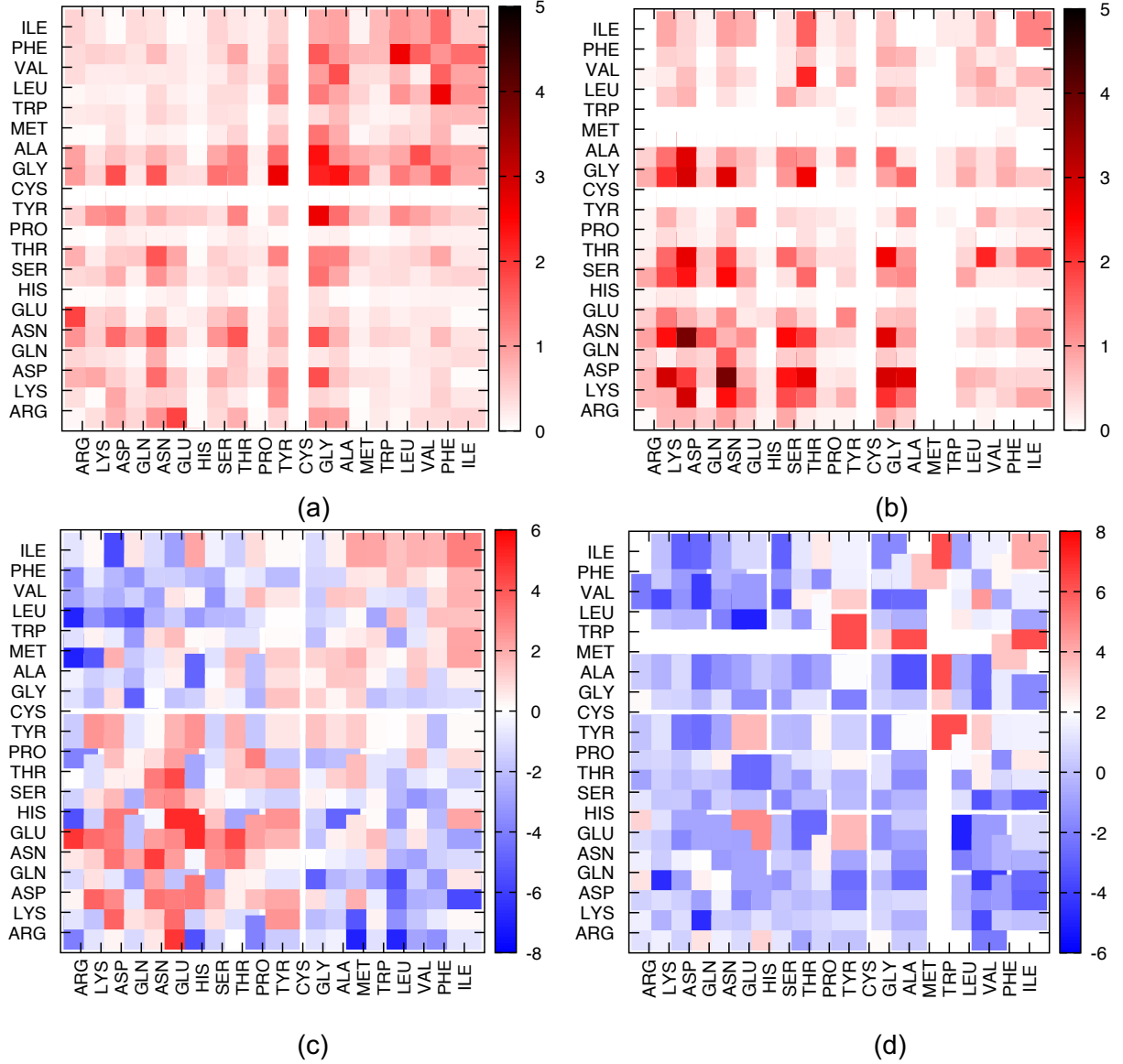
(a)



(b)

Şekil 25. Alfa-sarmal ve beta-silindir yapılarında rezidülerin (a) lipid ve (b) periferik ortamda bulunan arayüzey bölgelerdeki kontakt dereceleri.

Son olarak gözlemlenen kontakt frekansları ve skor değerlerine bakıldığında alfa-sarmal yapılardan oldukça farklı oldukları açıkça görülebilmektedir. Lipid ortamdaki kontakt frekansları hidrofobik rezidüler arasında yüksektir, ama alfa-sarmaldakiler kadar belirgin değildir. Diğer taraftan periferik ortamdaki arayüzeylerde polar rezidüler arası kontakt frekansları alfa-sarmal yapılara belirgin düzeyde yüksektir. Skor değerlerine bakıldığında Cys rezidüsü herhangi bir rezidü çiftiyle gözlemlenmediğinden tanımsızdır ve Şekil 26c ve 26d'de ince beyaz çizgilerle gösterilmiştir. En belirgin farklılık ta, periferik ortamdaki arayüzeyler için belirlenen skor değerlerindedir. Hidrofobik rezidü çiftleri arasında belirgin düzeyde yüksek olup pozitif değerlere sahiptir. Diğer taraftan alfa-sarmal yapılarda belirgin bir gruplaşma yoktur, ve pozitif değerler homojen bir şekilde dağılım göstermektedir.



Şekil 26. Beta-silindir yapılarda (a),(c) lipid ve (b),(d) periferik ortamlardaki arayüzeyde gözlenen (*observed*) kontakt frekansı KP_{ij} ve skor değerleri

5. Sonuç

Bu projede kompleks proteinlerin farklı bölgelerindeki amino asit dağılımını ve monomerik birimler arası etkileşimi etkileyen faktörler ve sonuçları karşılaştırmalı bir analizle incelenmiştir. Bu faktörler sırasıyla, monomerik birim sayısı, hücre içi (lipid) ve hücre dışı (periferik) ortam ve sekonder yapıdır. Bu doğrultuda OPM veri bankasından elde edilen 2216 alfa-sarmal ve 317 beta-silindir yapı titiz bir elemeye geçirilerek, eksik rezidülü olanlar ve kompleks oluşumları biyolojik olmak yerine kristalografik olan yapılar elenmiştir. Ayrıca, çoklu dizi hizalaması sonucunda sekansları %90'nın üstünde benzerlik gösteren yapılar elenmiştir. Bu sayede "artık" bilgiler veri setinden çıkartılarak her biri farklı amino asit dizilerinden oluşan özgün yapılar elde edilmiştir. Eleme sonucunda, alfa-sarmal yapılardan 168 homodimer, 29 heterodimer, 60 homotrimer, 99 homotetramer ve 7 heterotetramer olmak üzere toplam 363, beta-silindir yapılardan ise 6 homodimer, 2 heterodimer ve 31 homotrimer olmak üzere toplam 39 adet kompleks yapı elde edilmiştir. Beta-silindir yapılar az sayıda oldukları için 39 yapı bir arada ele alınmıştır. Diğer taraftan alfa-sarmal yapılardan 168 homodimer, 60 homotrimer ve 99 homotetramer ayrı kategorilerde analiz edilmiştir. Heterotetramer yapılar az sayıda oldukları ve heterotrimer yapı bulunmadığı için bu çalışmada sadece homokompleks versiyonlar ele alınmıştır.

Toplam 168 adet alfa-sarmal homodimer yapı üzerinde gerçekleştirilen ilk analizler sonucunda, *Met* ve *Trp* dışındaki diğer hidrofobik rezidülerin ciddi oranda polar rezidülere göre daha fazla olduğu belirlenmiştir. *Ser* ve *Thr* bileşimi ise %5 oranında ortalama bir değerde kalmıştır. *Cys* ve *His* ise %2 civarında en az bileşim oranına sahip rezidüler olmuştur. Benzer bir durum trimerik ve tetramerik yapılarda da gözlemlenmiştir. Diğer taraftan en dikkat çekici fark polar rezidülerin tetramerik yapılardaki yoğunluğunda görülmüştür. Rezidülerin ortalama SASA değerlerine bakıldığında *Trp* hariç diğer hidrofobik rezidülerin SASA değerleri 20-40 Å² civarındayken polar rezidülerinki 50-90 Å² arası değişmektedir. Kompleks yapıdaki birim sayısına göre bir karşılaştırma yapılacak olursa, tetramerik yapılardaki SASA değerleri polar rezidüler için 40-60 Å², hidrofobik rezidüler içinse 20-30 Å² gibi daha düşük bir profil sergilemektedir. Diğer taraftan, 39 beta-silindir yapı için benzer bir çalışma yapıldığında polar ve hidrofobik rezidülerin bileşim oranları arasındaki farkın ortadan kalktığı görülmektedir. *Cys* rezidüsü eser miktarda görülürken *His*, *Pro*, *Met* ve *Trp* %1 civarında bir oranla en düşük bileşime sahip rezidüler olmuşlardır. SASA değerlerine bakıldığında *Lys* hariç polar rezidülerin SASA değerlerinde alfa-sarmal yapılara göre ciddi bir düşüş (%25-30) gözlemlenmiştir.

Lipid ve periferik etkiyi ortaya çıkartmak için her kategori için hem lipid hem de periferik ortamlarda bulunan arayüzeylerdeki SASA, propansite, kontakt derecesi ve skor değerlerine

bakılarak farklılıklar gösterilmiştir. Periferik ortam, lipid ortama göre daha polar bir çevre sunduğundan, rezidülerin ortalama SASA değerlerinde bir miktar artış görülmüştür. Bu artış özellikle yüzey ve kenar arayüzey bölgelerde daha belirgin olmuştur. En büyük farklılıklar propansite değerlerinde görülmüştür. İç bölge de dahil olmak üzere tüm yüzey ve arayüzey bölgelerdeki propansite değerleri lipid ve periferik ortamlarda ciddi farklılıklar sergilemektedir. Örneğin alfa-sarmal homodimerik yapılarda polar rezidüler periferik ortamlardaki yüzey bölgelere lipid ortamdaki yüzey bölgelerden daha yüksek bir eğilim göstermektedir. Kontak derecelerine bakıldığında lipid ortamda görülen değerler periferik ortamdaki değerlerden tamamen farklı bir profil sergilemektedir. Örneğin, hidrofobik rezidülerin çoğu, dimer, trimer ve tetramerlerin tamamında 1'den düşük bir kontakt derecesi sergilerken, periferik ortama geçildiğinde bu değer birçok rezidü için 1'in üstüne çıkmaktadır. Polar rezidüler lipid ortamda daha karışık bir özellik sergilerken, farklı birim sayılarına sahip kompleks yapılarda belirgin bir ayırım yapılamamaktadır.

Skor değerlerine bakıldığında lipid ve periferik ortamdaki farklılık tüm kompleks kategorileri için belirgin bir şekilde ortaya çıkmıştır. Lipid ortamdaki gözlemlenen kontakt frekansı hidrofobik rezidü çiftleri için belirgin seviyede yüksekken, periferik ortama geçildiğinde frekans yoğunluğu tüm rezidü türleri arasında eşit bir dağılım göstermektedir. Diğer taraftan, lipid ortamdaki skor değerlerine bakıldığında pozitif ve yüksek olanlar polar rezidü çiftleri arasında görülürken, periferik ortama geçildiğinde her bir kategori için farklı dağılımlar ortaya çıkmaktadır. Örneğin dimerik yapılarda neredeyse rezidülerin tamamına yakını negatif skor değerlerine sahipken Cys-Cys 5.4 gibi yüksek bir skor değeri sergilemektedir. Trimerik yapılardaki periferik ortama bakıldığında, polar-polar ve hidrofobik-hidrofobik rezidü çiftlerinde daha yüksek pozitif değerler görülürken, negatif skorların sayısı oldukça az sayıdadır. Diğer taraftan, tetramerik yapılar dimerik yapılarla benzer bir dağılım gösterirken, beta-sarmal yapılar tamamen kendine özgü bir skor dağılımına sahiptir. Hidrofobik-hidrofobik rezidü çiftleri beta-sarmal yapılarda belirgin bir şekilde yüksek pozitif skor değerlerine sahiptir. Bu farklılıklar, her kategorinin kendi içinde analiz edilerek kendi içinde çalışılması ve bir takım istatistiksel skor değerleri belirlenirken bu ayırımın göz önünde bulundurulması gerektiğini açıkça göstermektedir.

6. Referanslar

- Ben-Naima, A. 2006. "On the driving forces for protein-protein association", *J Chem Phys*, 125, 24901.
- Berchanski, A., Shapira, B., Eisenstein, M. 2004. "Hydrophobic complementarity in protein-protein docking", *Proteins*, 56, 130–142.
- Eisenberg, D., Schwarz, E., Komaromy, M., Wall, R. 1984. "Analysis of Membrane and Surface Protein Sequences with the Hydrophobic Moment Plot", *J Mol Biol*, 179, 125–142.
- Gerstein, M., Lesk, A.M., Chothia, C. 1994. "Structural mechanisms for domain movements in proteins", *Biochemistry*, 33, 6739–6749.
- Glaser, F., Steinberg, D.M., Vakser, I.A., Ben-Tal, N. 2001. "Residue Frequencies and Pairing Preferences at Protein–Protein Interfaces", *Proteins*, 43, 89–102.
- Jayasinghe, S., Hristova, K., White, S.H. 2001. "MPtopo: a database of membrane protein topology", *Protein Sci*, 10(2), 455–458.
- Keskin, O., Mab, B., Nussinov, R. 2005. "Hot regions in protein–protein interactions: the organization and contribution of structurally conserved hot spot residues", *J Mol Biol*, 345, 1281–1294.
- Kysilka, J., Vondrasek, J. 2012. "Towards a better understanding of the specificity of protein-protein interaction", *J Molec Recognit*, 25, 604-615.
- Levy, E.D. 2010. "A Simple Definition of Structural Regions in Proteins and Its Use in Analyzing Interface Evolution", *J Mol Biol*, 403, 660-670.
- Orientations of proteins in membrane, OPM. <https://opm.phar.umich.edu>, Son erişim tarihi : 1 Eylül 2020.
- Protein Data Bank. <http://www.rcsb.org>, Son erişim tarihi : 1 Ekim 2020.
- Ulmschneider, M.B., Sansom, M.S.P, 2001. "Amino acid distributions in integral membrane protein structures", *Biochimica et Biophysica Acta*, 1512(1), 1-14.
- Ulmschneider, M.B., Sansom, M.S., Di Nola, A. 2005. "Properties of integral membrane protein structures: derivation of an implicit membrane potential", *Proteins*, 59, 252–265.
- Xu, Y., Xu, D., Liang, J. 2007. "Computational Methods for Protein Structure Prediction and Modeling: Structure Prediction of Membrane Proteins", Springer, 71 – 108.
- Yan, C., Feihong, W., Jernigan, R.L., Dobbs, D., Honavar, V. 2008. "Characterization of protein-protein interfaces", *Protein J*, 27(1), 59-70.



Young, L., Jernigan, R.L., Covell, D.G. 1994. "A role for surface hydrophobicity in protein-protein recognition", *Protein Sci*, 3, 717–729.

TÜBİTAK
PROJE ÖZET BİLGİ FORMU

Proje Yürütücüsü:	Doç. Dr. EBRU DEMET AKDOĞAN
Proje No:	119M188
Proje Başlığı:	Kompleks Proteinlerin Farklı Bölgelerindeki Amino Asit Dağılımı Ve Etkileşimini Etkileyen Çevresel Ve Yapısal Faktörlerin İncelenmesi: Hücre Zarı Ortamı, Sekonder Yapı Ve Kompleks Yapıdaki Monomer Sayısı
Proje Türü:	1002 - Hızlı Destek
Proje Süresi:	12
Araştırmacılar:	
Danışmanlar:	
Projenin Yürütüldüğü Kuruluş ve Adresi:	KADİR HAS Ü. MÜHENDİSLİK VE DOĞA BİLİMLERİ F.
Projenin Başlangıç ve Bitiş Tarihleri:	01/09/2019 - 01/09/2020
Onaylanan Bütçe:	45000.0
Harcanan Bütçe:	44325.25
Öz:	<p>Bu projede kompleks yapılardaki amino asit türleri, dağılımları, görülme sıklıkları, çözücü erşilir yüzey alanı (SASA), arayüzeyde etkileşen amino asit çiftlerinin görülme sıklıkları ve kontakt derecelerini etkileyen yapısal ve çevresel faktörler ele alınarak, bu değişimler ayrıntılı bir şekilde ortaya konmuştur. Bu faktörlerden biri olan kompleks yapının içinde bulunduğu ortam ele alındığında, hücre zarı proteinin hücre dışına taşan periferik ortamdaki arayüzey bölgesinin, hücre içi lipid ortamdaki arayüzey bölgesinden her yönden farklı olduğu gösterilmiştir. Bir diğer faktör de kompleks yapıyı oluşturan monomerik birim sayısıdır. Alfa-sarmal kategorisinde yer alan dimerik, trimerik ve tetramerik yapılar amino asitlerin tür ve dağılımları, SASA, propansite ve arayüzey kontakt dereceleri, frekansları ve skor değerlerine göre karşılaştırıldığında her yönden dikkat çekici farklılıklar ortaya çıkmıştır. Son olarak sekonder yapıdaki değişikliğin ne derece etkili bir faktör olacağını göstermek amacıyla, beta-silindir yapılar kendi içlerinde incelenmiş ve en büyük farklılık ta bu kategoride ortaya çıkmıştır. Bu bulgular doğrultusunda bir sonraki daha kapsamlı proje önerimizde geliştirilmesi planlanan arayüzey tahmin algoritması için kritik öneme sahip farklılıklar ortaya çıktığından, her kategorinin kendi içinde ele alınması gerektiği ve faktörlerden en fazla hücre içi/dışı ortam ile sekonder yapının dikkate alınması gerektiği vurgulanmıştır.</p>
Anahtar Kelimeler:	protein-protein arayüzey, sekonder yapı, hücre zarı ortamı, lipid, periferik, monomerik birimler
Fikri Ürün Bildirim Formu Sunuldu Mu?:	Hayır