

T.C.
NAMIK KEMAL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

**ÖZÜNDE DÜZENSİZ YAPILI PROTEİNLER ARASINDAN SEÇİLEN BİR GRUP
PROTEİNİN REZİDÜLERİNİN YERLEŞME ÖZELLİKLERİNİN HESAPLAMALI
YÖNTEMLERLE İNCELENMESİ**

Rasim Murat AYDINKAL

BİYOLOJİ ANABİLİM DALI

DANIŞMAN: DOÇ. DR. EVREN CABİ

TEKİRDAĞ-2016

Doç. Dr. Evren Cabi danışmanlığında, Rasim Murat Aydınkal tarafından hazırlanan “Özünde düzensiz yapılı proteinler arasından seçilen bir grup proteinin rezidülerinin yerleşme özelliklerinin hesaplamalı yöntemlerle incelenmesi” isimli bu çalışma aşağıdaki jüri tarafından Biyoloji Anabilim Dalı’nda Yüksek Lisans tezi olarak oy birliği ile kabul edilmiştir.

Jüri Başkanı : Prof. Dr. Serbülent YILDIRIM

İmza:

Üye : Doç. Dr. Ersin KARABACAK

İmza:

Üye : Doç. Dr. Evren CABI

İmza:

Fen Bilimleri Enstitüsü Yönetim Kurulu adına

Prof. Dr. Fatih KONUKCU

Enstitü Müdürü

ÖZET

Yüksek Lisans Tezi

ÖZÜNDE DÜZENSİZ YAPILI PROTEİNLER ARASINDAN SEÇİLEN BİR GRUP PROTEİNİN REZİDÜLERİNİN YERLEŞME ÖZELLİKLERİNİN HESAPLAMALI YÖNTEMLERLE İNCELENMESİ

Rasim Murat AYDINKAL

Namık Kemal Üniversitesi

Fen Bilimleri Enstitüsü

Biyoloji Anabilim Dalı

Danışman: Doç. Dr. Evren CABİ

Bu çalışmada özünde düzensiz yapıli proteinlerin (ÖDYP) rezidü yerleşme düzenleri hesaplamalı yöntemlerle incelenmiştir. Kullanılan ölçüt, proteinlerin alfa karbonlarının koordinatlarından Delaunay Teselasyonu ile elde edilen tetrahedronların küresellik değerleridir. Kontrol grubu olarak seçilen ÖDYP olmayan proteinlerin (alfa proteinler, beta proteinler) rezidü yerleşme düzenlerini incelemek için oluşturulan tetrahedronların küresellik değerlerinin dağılımlarının heterojen oldukları saptanmıştır. Bu sebeple kontrol grubu iki ayrı grup olarak alfa proteinler ve beta proteinlerden oluşmaktadır. Kontrol grubu proteinlerin ÖDYP grubu proteinler ile Kolmogorov-Smirnov testi ile karşılaştırılmaları sonucu ÖDYP grubu proteinlerin hem alfa, hem de beta proteinlerden küresellik ölçütü kullanıldığında farklılık gösterdiği bulunmuştur. Bu tez çalışmasında elde edilen bulgulara göre ÖDYP'ler birbirleri ile bir grup oluşturmak yerine kendilerine özgü bir yapı sergilemektedirler.

Anahtar kelimeler: Delaunay Teselasyonu, küresellik, özünde düzensiz yapıli proteinler, rezidü yerleşim düzeni

2016, 75 sayfa

ABSTRACT

MSc. Thesis

COMPUTATIONAL INVESTIGATION OF THE EXTENT OF REGULARITIES IN RESIDUE PACKING IN A REPRESENTATIVE GROUP OF INTRINSICALLY DISORDERED PROTEINS

Rasim Murat AYDINKAL

Namık Kemal University

Institute of Natural and Applied Sciences

Department of Biology

Supervisor: Dr. Evren CABİ

In this thesis, the extent of regularities in residue packing of intrinsically disordered proteins (IDPs) were investigated by computational methods. The criterion is the sphericity of the tetrahedra obtained by Delaunay Tessellation of the C-alpha coordinates of proteins. It is observed that the distributions of the sphericity values of the tetrahedra of the control group proteins which are non-IDPs (alpha proteins, beta proteins) are heterogeneous. Therefore, the control group consists of two separate groups as alpha proteins and beta proteins. It has been found that the IDPs are distinct from both alpha proteins and beta proteins after the comparison of the two sample Kolmogorov-Smirnov test results of proteins in control group, and IDPs, according to the sphericity criterion. As an interesting result of this study, each IDP present as a unique structure instead of belonging to the same structural group.

Keywords: Delaunay Tessellation, residue packing, intrinsically disordered proteins, sphericity

2016, 75 pages

İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET	i
ABSTRACT	ii
İÇİNDEKİLER	iii
ÇİZELGE DİZİNİ	vi
ŞEKİL DİZİNİ	ix
KISALTMALAR	x
ÖNSÖZ	xi
1. GİRİŞ	1
1.1. Proteinlerin Genel Özellikleri.....	1
1.2. Proteinlerin Yapısı.....	3
1.2.1. Primer yapı: amino asitler.....	3
1.2.2. Sekonder yapı	4
1.2.2.1. Alfa sarmal	5
1.2.2.2. Beta iplikçığı	6
1.2.3. Tersiyer yapı: 3-boyutlu yapı	8
1.2.4. Kuaterner yapı: kompleks oluşturmuş protein molekülleri	9
1.3. Proteinlerin Katlanması	10
1.3.1. Termodinamik açıdan bakış.....	10
1.3.2. Protein katlanma kinetiği.....	12
1.4. Veri Tabanları.....	12
1.4.1. DISPROT veri tabanı.....	13
1.4.2. CATH protein sınıflandırma veri tabanı.....	13

1.4.3. RCSB-PDB: protein bilgi bankası	15
1.5. Özünde Düzensiz Yapılı Proteinler (ÖDYP).....	15
1.5.1. Özünde düzensiz yapılı proteinlerin doğal çokluğu	19
1.5.2. Özünde düzensizlik tayini	20
1.5.2.1. X-ışını kristalografisi.....	21
1.5.2.2. Nükleer manyetik rezonans spektroskopisi	21
1.6. Küresellik Kavramı.....	21
1.7. Delaunay Teselasyonu	22
1.8. Kolmogorov-Smirnov Testi.....	23
2. MATERYAL ve YÖNTEM.....	25
2.1. Kontrol Grubu Proteinleri ve Yapısal Sınıflandırmaları	25
2.2. Özünde Düzensiz Yapılı Proteinlerin Listesinin Hazırlanması	38
2.3. Irregular Tetrahedron Rezidü Kümelerinin Elde Edilmesi.....	43
2.4. ÖDYP'lerin Küresellik Değerlerinin Hesaplanması	43
3. BULGULAR ve TARTIŞMA	44
3.1. Proteinlerin Küresellik Hesapları	44
3.1.1. Kontrol grubu proteinler.....	44
3.1.1.1. Alfa proteinler.....	44
3.1.1.2. Beta proteinler	45
3.1.2. Özünde düzensiz yapılı proteinler	45
3.2. İki Ayrı Proteinde Tetrahedronların Küresellik Dağılımlarının Kolmogorov-Smirnov İki Örnekleme Testi İle Karşılaştırılması	46
3.2.1. Alfa proteinler.....	46
3.2.2. Beta proteinler	49
3.2.3. Alfa proteinler ve beta proteinler.....	50

3.2.4. Alfa proteinler ve özünde düzensiz yapılı proteinler.....	51
3.2.5. Beta proteinler ve özünde düzensiz yapılı proteinler	51
3.2.6. Özünde düzensiz yapılı proteinler	52
3.2.7. Kolmogorov-Smirnov testi sonuçlarının değerlendirilmesi	53
4. SONUÇ ve ÖNERİLER	54
5. KAYNAKLAR.....	55
EKLER	60
ÖZGEÇMİŞ	61

ÇİZELGE DİZİNİ

Sayfa

Çizelge 1.1 : Amino asitlerin özellikleri ve standart kodları.....	2
Çizelge 1.2 : Kristalografik analizler sonucu belirlenemeyen amino asitler için dejeneresite kodları	2
Çizelge 1.3 : CATH hiyerarşik sınıflandırma kodları ve kümeleme kriterleri.....	14
Çizelge 1.4 : Bazı 3-boyutlu geometrik şekiller ve küresellik değerleri	22
Çizelge 2.1 : Kontrol grubu protein yapıları	25
Çizelge 2.2 : 1LMB kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları.....	26
Çizelge 2.3 : 1B0N kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları.....	26
Çizelge 2.4 : 1M45 kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları.....	26
Çizelge 2.5 : 1VRK kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları.....	27
Çizelge 2.6 : 1A7W kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları.....	27
Çizelge 2.7 : 1BGF kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları	27
Çizelge 2.8 : 1DK8 kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları	28
Çizelge 2.9 : 1G33 kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları	28
Çizelge 2.10 : 1G4I kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları	28
Çizelge 2.11 : 1GV2 kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları	29
Çizelge 2.12 : 1HBK kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları.....	29
Çizelge 2.13 : 1I2T kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları	29
Çizelge 2.14 : 1I8O kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları.....	30
Çizelge 2.15 : 1L9L kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları	30
Çizelge 2.16 : 1NOG kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları.....	30
Çizelge 2.17 : 1C9O kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları.....	31

Çizelge 2.18 : 1HOE kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları	31
Çizelge 2.19 : 1I0C kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları	31
Çizelge 2.20 : 1M9Z kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları	32
Çizelge 2.21 : 1O5U kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları	32
Çizelge 2.22 : 1OH4 kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları	32
Çizelge 2.23 : 1PMH kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları.....	33
Çizelge 2.24 : 2FCB kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları.....	33
Çizelge 2.25 : 1ROC kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları	33
Çizelge 2.26 : 3EZM kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları.....	34
Çizelge 2.27 : 1TUD kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları	34
Çizelge 2.28 : 1UMH kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları	34
Çizelge 2.29 : 1V05 kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları	35
Çizelge 2.30 : 1V6P kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları	35
Çizelge 2.31 : 1Y0M kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları.....	35
Çizelge 2.32 : İncelenecek olan özünde düzensiz yapılı proteinlerin listesi	38
Çizelge 2.33 : DP00143 kodlu özünde düzensiz yapılı proteinin künyesi	39
Çizelge 2.34 : DP00145 kodlu özünde düzensiz yapılı proteinin künyesi	39
Çizelge 2.35 : DP00347 kodlu özünde düzensiz yapılı proteinin künyesi	40
Çizelge 2.36 : DP00387 kodlu özünde düzensiz yapılı proteinin künyesi	41
Çizelge 2.37 : DP00705 kodlu özünde düzensiz yapılı proteinin künyesi	42
Çizelge 2.38 : DP00748 kodlu özünde düzensiz yapılı proteinin künyesi	42
Çizelge 3.1 : Alfa proteinlerin küresellik hesaplarının Kolmogorov-Smirnov testi sonucunda elde edilen P değerleri	48
Çizelge 3.2 : Beta proteinlerin küresellik hesaplarının Kolmogorov-Smirnov testi sonucunda elde edilen P değerleri	49

Çizelge 3.3	: Alfa proteinler ile Beta proteinlerin küresellik hesaplarının Kolmogorov-Smirnov testi sonucunda elde edilen P değerleri	50
Çizelge 3.4	: Alfa proteinler ile özünde düzensiz yapılı proteinlerin küresellik hesaplarının Kolmogorov-Smirnov testi sonucunda elde edilen P değerleri	51
Çizelge 3.5	: Beta proteinler ile özünde düzensiz yapılı proteinlerin küresellik hesaplarının Kolmogorov-Smirnov testi sonucunda elde edilen P değerleri	52
Çizelge 3.6	: Özünde düzensiz yapılı proteinlerin küresellik hesaplarının Kolmogorov-Smirnov testi sonucunda elde edilen P değerleri	52
Çizelge 3.7	: Kolmogorov-Smirnov testi sonuçlarındaki oranlar	53

ŞEKİL DİZİNİ

	<u>Sayfa</u>
Şekil 1.1 : Amino asit yapısı	4
Şekil 1.2 : Ramachandran grafiği.....	5
Şekil 1.3 : Alfa sarmalda hidrojen bağı oluşumu.....	6
Şekil 1.4 : Beta iplikçığı gösterimi	7
Şekil 1.5 : Beta iplikçiklerinin Ribbon diyagramında gösterimi	8
Şekil 1.6 : Miyogloblin'in 3-boyutlu yapısı.....	9
Şekil 1.7 : Protein katlanması enerji diyagramı şematik örneği	11
Şekil 1.8 : Özünde düzensiz yapılı protein örneği, C2H2-tipi çinko finger solusyon yapılı proteini	15
Şekil 1.9 : Protein yapı-fonksiyon paradigması sonrası oluşmuş bazı modern bilim dalları	16
Şekil 1.10 : Özünde düzensiz yapılı proteinlerin keşfi ile ortaya çıkan çalışma alanlarından bazıları.....	18
Şekil 1.11 : Makromoleküle bağlanmış ÖDYP'nin mevcut 3 boyutlu yapıya adapte olması	18
Şekil 1.12 : Çeşitli proteomlardaki özünde düzensizlik dağılımları	20
Şekil 1.13 : Voronoi diyagramı ve Delaunay Teselasyonu örneği	23
Şekil 2.1 : Alfa proteinlerin 3-boyutlu yapıları.....	36
Şekil 2.2 : Beta proteinlerin 3-boyutlu yapıları	37
Şekil 2.3 : Özünde düzensiz yapılı proteinlerin 3-boyutlu yapıları	43
Şekil 3.1 : Alfa proteinlerdeki tetrahedronların küresellik histogramları	44
Şekil 3.2 : Beta proteinlerdeki tetrahedronların küresellik histogramları.....	45
Şekil 3.3 : Özünde düzensiz yapılı proteinlerdeki tetrahedronların küresellik histogramları	46

KISALTMALAR

ALS	: Amyotrofik Lateral Skleroz
CATH	: Class, Architecture, Topology, Homologous superfamily
DisProt	: The Database of Protein Disorder
ER	: Endoplazmik Retikulum
IDP	: Intrinsically Disordered Protein
JCBN	: Joint Commission on Biochemical Nomenclature
K-S	: Kolmogorov-Smirnov Test
MATLAB	: Matrix Laboratory
MAPK8	: Mitogen-Activated Protein Kinase 8
NCBI	: National Center for Biotechnology Information
NIH	: National Institutes of Health
NMR	: Nuclear Magnetic Resonance
ÖDYP	: Özünde Düzensiz Yapılı Protein
PDB	: Protein Data Bank
PMC	: PudMed Center
PONDR	: Predictor Of Naturally Disordered Regions
RCSB	: Research Collabatory of Structural Biology
SCOP	: Structural Classification of Proteins
SwissProt	: Swiss Prot Protein Sequence Database
TrEMBL	: Translated European Molecular Biology Laboratory Nucleotide Sequence Database
UniProt	: Universal Protein Resource

ÖNSÖZ

Bu çalışma Namık Kemal Üniversitesi Fen Bilimleri Enstitüsü Biyoloji Anabilim dalında yüksek lisans tezi olarak hazırlanmıştır.

Çalışmada özünde düzensiz yapılı proteinler (IDP) arasından seçilen bir grup proteinin alfa karbonlarının koordinatlarından Delaunay Teselasyonu ile elde edilen tetrahedronların küresellik değerleri ile, kontrol grubu olarak seçilen yapısı iyi tanımlanmış alfa proteinler ve beta proteinlerin alfa karbonlarının koordinatlarından Delaunay Teselasyonu ile elde edilen tetrahedronların küresellik değerlerinin dağılımları karşılaştırılarak, özünde düzensiz yapılı proteinlerin rezidü yerleşim özellikleri araştırılmıştır. Küresellik değerleri Kolmogorov-Smirnov testi ile karşılaştırıldığında, özünde düzensiz yapılı proteinler hem alfa proteinlerden, hem de beta proteinlerden küresellik ölçütü bakımından farklılık göstermiştir. Elde edilen bulgular, özünde düzensiz yapılı proteinlerin birbirleri ile bir grup oluşturmak yerine, kendilerine özgü bir yapı sergilediğini göstermesi açısından önemlidir.

Tez çalışmamda “bilim insanı modeli” olarak örnek aldığım değerli danışmanlarım Doç. Dr. Evren CABİ ve Doç. Dr. Elife Zerrin BAĞCI'ya yönlendirme ve bilgilendirmeleriyle çalışmamı bilimsel temeller ışığında sabırla şekillendirdikleri için, Ali Nihat Gökyiğit Vakfı'nın kıymetli yöneticileri Uzm. Salih Sercan Kanoğlu ve Prof. Dr. Adil GÜNER'e yüksek lisans öğrenimim süresince gösterdikleri müsamaha için, çalışmalarımın MATLAB programındaki ilk aşamalarında beni teşvik ederek cesaretlendiren Uzm. Furkan ELİBOL'a yardımseverliği ve arkadaşlığı için teşekkür ederim.

Mayıs 2016

Rasim Murat Aydınkal
Biyolog

1. GİRİŞ

1.1 Proteinlerin Genel Özellikleri

Proteinler, 20 farklı çeşitteki amino asitten meydana gelen polimerlerdir ve her birinin farklı kimyasal özellikleri bulunmaktadır. Her bir amino asit, komşusu olan amino aside kovalent bağ yaparak bağlanarak ve sonucunda uzun peptit zinciri oluşturarak bir protein molekülünü meydana getirmektedir. Bu sebeple proteinlere polipeptit denilmektedir. Tüm proteinlerin kendilerine özgü amino asit dizisi vardır, böylece farklı amino asit dizileri de farklı proteinleri oluşturmaktadır.

Proteinin 3 boyutlu yapısının, doğrudan o proteinin amino asit dizisindeki atomlar arasındaki fiziksel etkileşimlerin karmaşık bir etkileşimi sonucunda ortaya çıktığı belirlenmektedir (Rose ve ark. 2006).

Sulu çözelti içerisinde fizyolojik sıcaklıklarda, proteinlerin polipeptit zincirleri genellikle kompakt (küresel) formda katlanırlar. Proteinlerdeki farklı amino asitlerin dizisi primer yapıyı oluşturmaktadır. Bu da proteinlerin üst seviyedeki yapılarına nasıl katlanacağını belirler.

Polipeptit zincirinin sekonder yapısı, C-H ve C=O grupları arasındaki etkileşimlerle alfa sarmalı veya beta iplikçığı oluşturabilir. Küresel formdaki proteinler tersiyer yapılarına katlanarak ulaşırlar. Çoğu protein birden fazla polipeptitin katlanmış zincirlerinin bağlantısıyla oluşmaktadır; bu da proteinin kuarterner yapısını belirler. Bir polipeptitin protein olarak işlev görmesi için, fizyolojik şartlarda uygun tersiyer yapısına ulaşmış olmalıdır. Yanlış katlanan proteinler, yardımcı şaperon proteinleri ile kontrol edilirler. Yanlış katlanma veya katlanmama durumlarının Alzheimer, Parkinson, Amiyotrofik lateral skleroz (ALS) gibi nörodejeneratif hastalıklara sebep olduğu bilinmektedir (Lin ve Beal 2006, Uversky ve ark. 2008, Uversky ve ark. 2014).

Çizelge 1.1 ile gösterilen 3-harfli ve 1-harfli amino asit koduna ek olarak, kimyasal veya kristalografik analizler sonucunda bir proteinin belli bir pozisyonundaki amino asit kesin olarak belirlenemediğinde, Çizelge 1.2'deki kısaltmalar da kullanılabilir.

Çizelge 1.1. Amino asitlerin özellikleri ve standart kodları (IUPAC-IUB JCBM 1984)

Amino asit	3-Harf	1-Harf	Yan zincir polaritesi	Yan zincir asidite veya bazidite
Alanin	ALA	A	Apolar	Nötür
Arjinin	ARG	R	Polar	Kuvvetli bazik
Asparajin	ASN	N	Polar	Nötür
Aspartik asit	ASP	D	Polar	Asidik
Sistein	CYS	C	Polar	Nötür
Glütamik asit	GLU	E	Polar	Asidik
Glütamin	GLN	Q	Polar	Nötür
Glisin	GLY	G	Apolar	Nötür
Histidin	HIS	H	Polar	Hafif bazik
İzolösin	ILE	I	Apolar	Nötür
Lösin	LEU	L	Apolar	Nötür
Lizin	LYS	K	Polar	Bazik
Metiyonin	MET	M	Apolar	Nötür
Fenilalanin	PHE	F	Apolar	Nötür
Prolin	PRO	P	Apolar	Nötür
Serin	SER	S	Polar	Nötür
Treonin	THR	T	Polar	Nötür
Triptofan	TRP	W	Apolar	Nötür
Tirozin	TYR	Y	Polar	Nötür
Valin	VAL	V	Apolar	Nötür

Çizelge 1.2. Kristalografik analizler sonucu belirlenemeyen amino asitler için dejeneresite kodları (Stothard 2000)

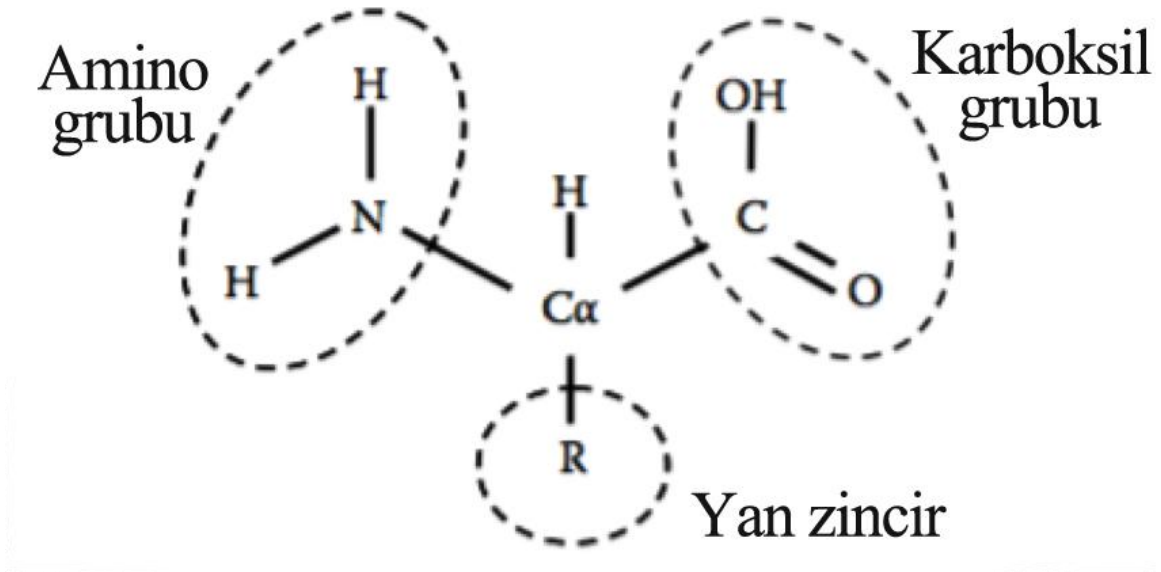
Dejeneresite kodu	3-Harf	1-Harf
Asparajin veya Aspartat	ASX	B
Glütamin veya Glütamat	GLX	Z
Herhangi biri veya bilinmeyen	UNK	X

1.2. Proteinlerin Yapısı

1.2.1 Primer yapı: amino asitler

Birincil yapı, bir polipeptit zincirin amino asit dizisidir. Birincil yapı peptit bağlarla oluşur, protein sentezi sırasında meydana gelir. Polipeptit zincirin iki ucu, uçlardaki serbest grubun cinsine bağlı olarak karboksil uç (veya C-terminal) ve amino uç (veya N-terminal) olarak adlandırılır. Amino asit rezidüleri N-ucundan başlar, buradaki amino grubu (NH₂-grup) bir peptit bağı içinde yer almaz. Organik moleküllerde fonksiyonel grubun bağlı olduğu ilk karbon atomuna alfa karbon (C α), ikinci karbon atomuna ise beta karbon (C β) denilmektedir. Bu isimlendirme Yunan alfabesindeki sıraya göre devam etmektedir. Canlılarda bir proteinin birincil yapısı o proteini kodlayan gene bağlıdır. DNA'daki belli bir nükleotit dizisi (bir gen) okunarak bir mRNA sentezlenir, mRNA da, translasyon olarak adlandırılan bir süreç sonucu, ribozomlar tarafından okunup, genin ürünü olan proteinin sentezlenmesini sağlar.

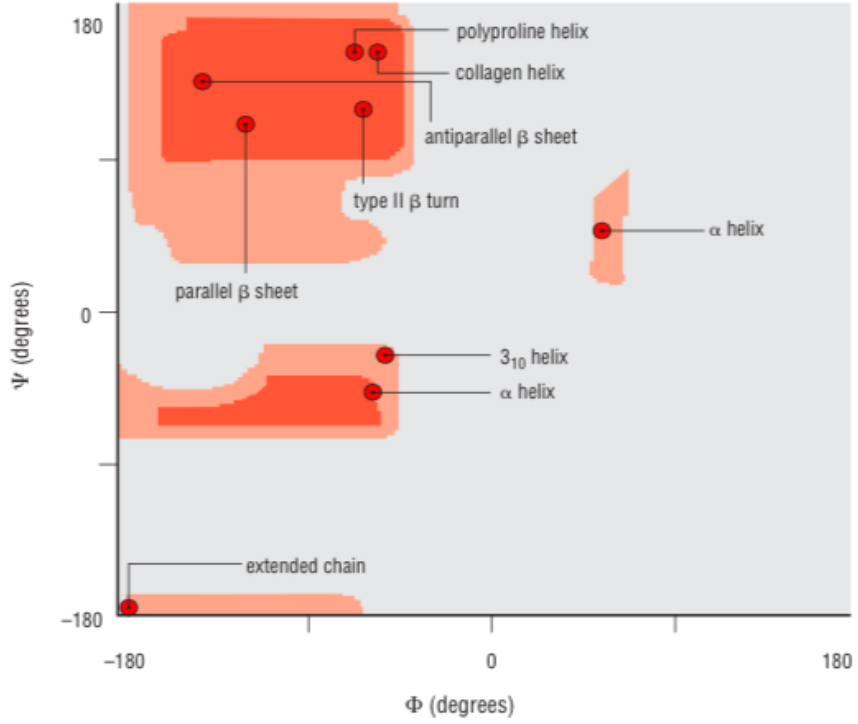
Proteinin dizisi o proteine spesifiktir, proteinin yapısı ve fonksiyonunu belirler. Bir proteinin dizisi Edman yıkımı veya kütle spektrometresi gibi yöntemlerle belirlenebilir. Ancak, çoğu zaman, o proteini kodlayan genin dizisi doğrudan okunarak (bunun için genetik kod kullanılır) proteinin dizisi belirlenir. Translasyon sonrası değişimler, disülfid oluşumu, fosforilasyon ve glikozilasyonlar gibi, genelde birincil yapının parçası sayılırlar ama gen dizisi okunarak bunların varlığı gösterilemez (Wikipedia 2015). Şekil 1.1'de amino asit yapısı gösterilmektedir.



Şekil 1.1. Amino asit yapısı

1.2.2 Sekonder yapı

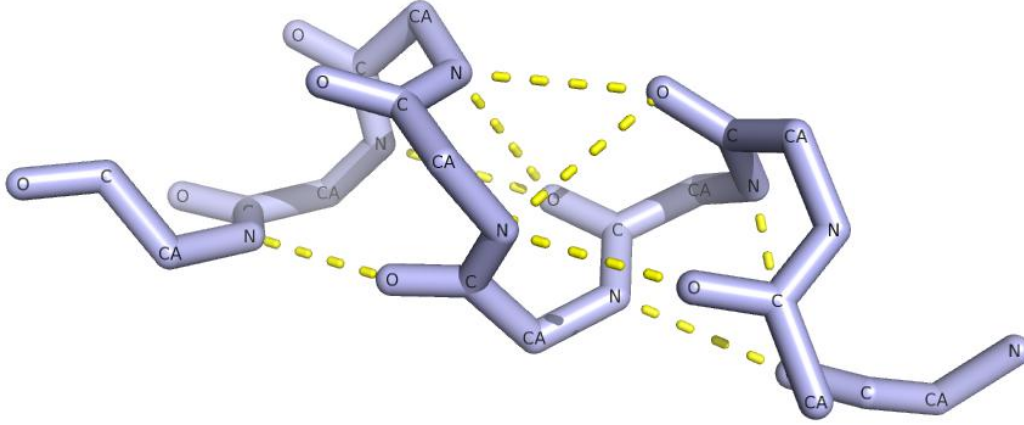
Proteinlerde sekonder yapı, çok düzenli yerel alt yapılarıdır. Alfa sarmal ve beta iplikçığı olarak adlandırılan iki ana sekonder yapı tipi, 1951'de Linus Pauling ve arkadaşları tarafından önerilmiştir (Pauling ve ark. 1951). Bu sekonder yapılar, protein omurgasındaki atomlar arasındaki hidrojen bağı örüntüleri ile tanımlanır. Düzenli bir geometriye sahiptir, Şekil 1.2.'deki Ramachandran grafiğinde belli ψ ve ϕ açıları ile kısıtlanmıştır (Ramachandran ve ark. 1963). Alfa sarmalı ve beta iplikçikleri, peptit omurga üzerindeki hidrojen bağı verici ve alıcılarının oluşmasını sağlayan yerel yapılarıdır. Katlanmış protein yapısında bazı kısımlar düzenli değildir. Bu kısımlar rastgele sarım (*random coil*) ile karıştırılmamalıdır; rastgele sarım, belli bir sabit üç boyutlu yapıya sahip olmayan, katlanmamış polipeptit zinciridir. Birbirini izleyen birkaç ikincil yapı bir "süperikincil birim" (*supersecondary structure*) oluşturabilir (Chiang ve ark. 2007).



Şekil 1.2. Ramachandran grafiği (Petsko ve Ringe 2008)

1.2.2.1 Alfa sarmal

Alfa sarmalları, proteinin ana zinciri hidrojen bağları ile kararlı hale gelmiş bir sarmal konformasyona kavuştuğunda oluşur. Şekil 1.3 hidrojen bağlarının oluşumunu örnek olarak göstermektedir. Bir hidrojen bağında n'inci rezidünün karbonil oksijeni (n+4)'üncü rezidünün azot donörünün hidrojen akseptörüdür. Alfa sarmalı şu şekilde ayırt edilir: n'inci (n+4)'üncü rezidü ile bir hidrojen bağı yolu ile etkileşir (Pauling ve ark. 1951).



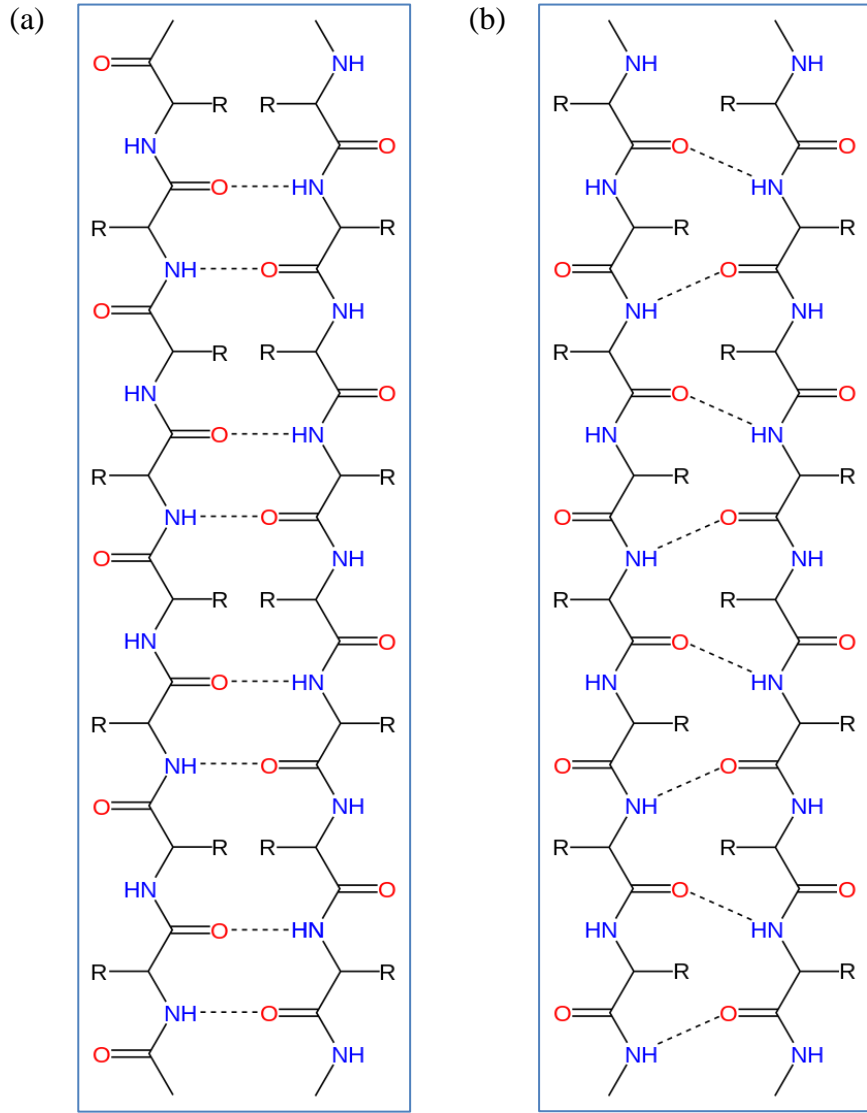
Şekil 1.3. Alfa sarmalda hidrojen bağı oluşumu (Sarı kesik çizgiler hidrojen bağlarını temsil etmektedir.)

Alfa sarmalında bir dönüş 3.6 rezidüye karşılık gelir (Pauling ve ark. 1951). Bu yüzden sarmalın eksenine dik bir perspektiften bakıldığında, birbiri ardına gelen karbon atomlarının arasında aşağı yukarı 100° 'lik bir açısal farklılık görülür (Richmond ve Richards 1978, Murzin ve Finkelstein 1988, Barlow ve Thornton 1988).

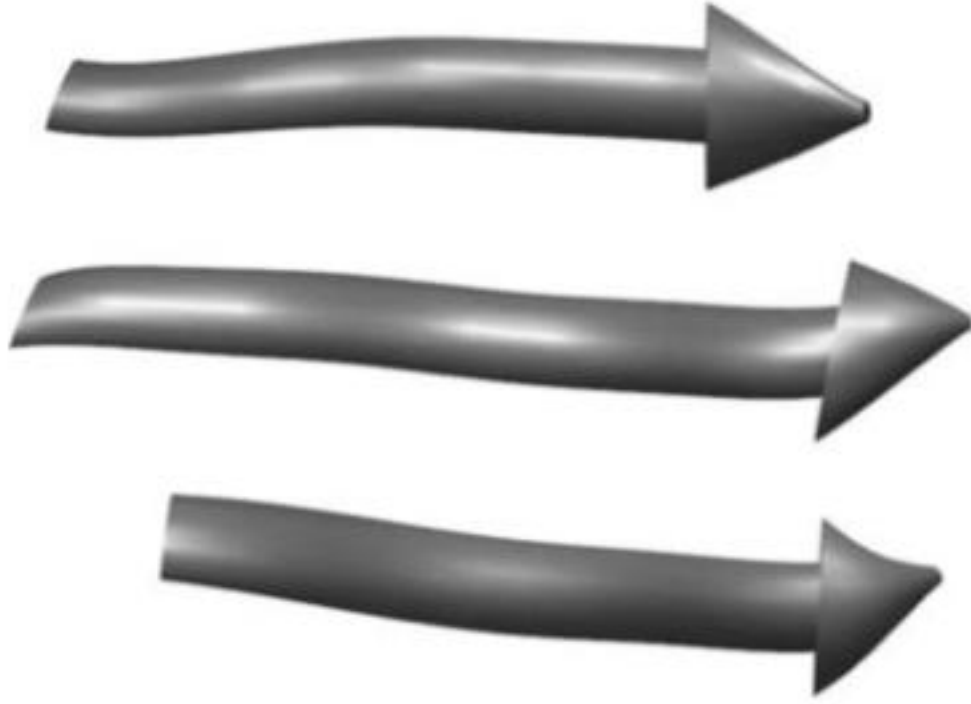
1.2.2.2 Beta iplikçığı

Alfa sarmallarında olduğu gibi, beta iplikçikleri de hidrojen bağları ile oluşan ikincil yapılardır. Ancak, bu durumda, birincil yapıda birbirinden çok uzakta olabilen iki amino asit dizisi etkileşir. Hidrojen bağları iki amino asit zincir parçası arasında iki farklı oryantasyona sahip bağlantı oluşturabilir (Chothia 1973, Salemme ve Weatherford 1981):

- Paralel beta iplikçiklerinde birbirlerini takip eden hidrojen bağlarında her iki iplikçığın amino asit endeksi azalır,
- Antiparalel beta iplikçiklerinde ise birbirlerini takip eden hidrojen bağlarında bir iplikçikteki amino asit endeksi artarken, diğerinin amino asit endeksi azalır (Salemme ve Weatherford 1981).



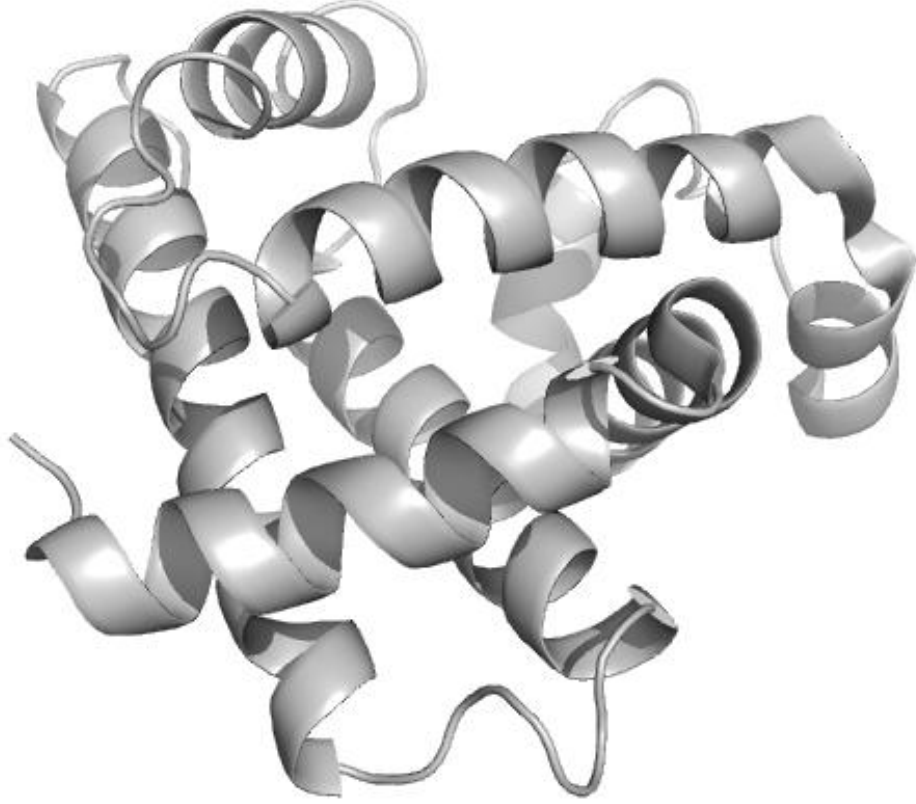
Şekil 1.4. Beta iplikçığı gösterimi (a) antiparalel beta iplikçikleri, (b) paralel beta iplikçikleri (Wikipedia 2015)



Şekil 1.5. Beta iplikçiklerinin Ribbon diyagramında gösterimi

1.2.3. Tersiyer yapı: 3-boyutlu yapı

Üçüncül yapı, tek bir protein molekülünün üç boyutlu yapısıdır. Alfa sarmal ve beta iplikler kompakt bir yapı oluşturacak şekilde katlanırlar. Bu katlanma hidrofobik etkileşimler (hidrofobik rezidülerin sudan uzaklaşması) tarafından yönlendirilir ama yapının stabil olabilmesi için spesifik üçüncül etkileşimlerle (tuz köprüleri, hidrojen bağları, disülfid bağları ve yan zincirlerin sıkı istiflenmesi gibi) bir protein bölgesinin kısımlarının yerinin stabil olması gerekir. Şekil 1.6'da miyoglobin için 1MBN örneği bulunmaktadır.



Şekil 1.6. Miyoglobin'in 3-boyutlu yapısı (PDB kodu: 1MBN)

1.2.4. Kuaterner yapı: kompleks oluşturmuş protein molekülleri

Dördüncül yapı birkaç protein veya polipeptit zincirinin (bu bağlamda bunlara protein altbirimleri denir) bir araya gelmesinden meydana gelen büyük bir kompleks yapıdır. Dördüncül yapı, üçüncül yapıyı stabilize eden, kovalent olmayan bağlar ve disülfid bağları tarafından stabilize edilir. İki ve daha çok (yani çoklu) polipeptitten oluşan kompleksler mültimer olarak adlandırılır. Daha spesifik olarak, böylesi bir yapı eğer iki altbirimden oluşuyorsa dimer, üç altbirimden oluşuyorsa trimer, ve dört altbirimden oluşuyorsa tetramer olarak adlandırılır. Altbirimler genelde birbirlerine simetri işlemleri ile ilişkilidir, örneğin bir dimerde iki katlı bir simetri eksenine sahiptir. Birbirinin aynı altbirimlerden oluşan mültimerlere değinirken "homo-" öneki kullanılır (örneğin bir homotetramer), farklı altbirimlerden oluşanlar için ise "hetero-" öneki kullanılır (örneğin bir heterotetramer, iki alfa ve iki beta zinciri olan hemoglobinden bahsederken).

1.3. Proteinlerin Katlanması

Proteinler konformasyonel olarak dinamik moleküllerdir ve su içinde sürekli olarak hareket halindedirler. Bir proteinin fonksiyonunu gerçekleştirebilmesi için, 3-boyutlu şeklinin de o fonksiyon için uygun olması gerekmektedir (Levinthal 1969). Fonksiyonel şekillerini mikrosaniye gibi çok kısa bir sürede katlanarak alırlar. Katlanma işlemi geri dönüşü olabilen bir işlemdir, genellikle yapıları bozulsa da eski şekillerini geri alabilirler.

Biyolojik olarak uygun şartlarda proteinler en düşük enerjiyi tercih ederler. Proteinler ribozomlarda doğal polipeptit olarak sentezlenir ve endoplazmik retikulum (ER) lumeninde katlanırlar. Bu işlemin amacı proteinin 3-boyutlu yapısını almasıdır.

Moleküler şaperonlar proteinlerin katlanma sürecinde yardımcı proteinlerdir. Şaperonlar katlanmamış polipeptit zincirin hidrofobik bölgelerine geçici olarak bağlanır, onları çözücülerine karşı perdeler. Kümeleşmeyi (agregasyon) önleyerek proteinlerin uygun sekonder yapılarını kazanmalarına, hatalı katlanan proteinlerin ise yıkımlanmasına (degredasyon) yardımcı olurlar (Anfinsen 1973). ER'de gerçekleşen bu yıkımlanma olayı sonucu protein fonksiyonunu kaybeder. Endoplazmik retikulum kalite kontrolündeki kusurlar hatalı katlanmış proteinlerin burada birikimine ve kronik ER stresine yol açarak apoptoza neden olur.

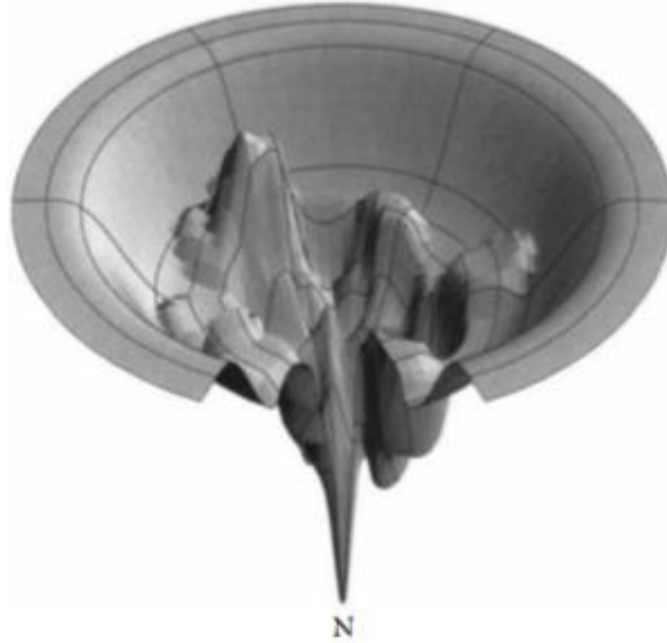
1.3.1. Termodinamik açıdan bakış

Bir proteinin katlanması, su molekülleri dahil olmak üzere katlanmış ve katlanmamış durumlar arasındaki Gibbs serbest enerjisi farkına göre eğilimini belirleyen çok sayıda zayıf etkileşmeye bağlıdır. Entalpi H , Kelvin sıcaklığı T , entropi S iken Gibbs serbest enerjisi (1.1) bağıntısında verilmiştir.

$$\Delta G_{toplam} = \Delta H_{zincir} + \Delta H_{çözücü} - T\Delta S_{zincir} - T\Delta S_{çözücü} \quad (1.1)$$

Bireysel bileşenler katlanmış ve katlanmamış durumlar arasındaki farklılıklardan kaynaklanmaktadır (Garrett ve Grisham 2007). Katlanmış yapı çok düzenlidir; bu sebepten

dolayı denklemdeki $-T\Delta S_{zincir}$ değeri büyük bir pozitif büyüklüktür. Diğer terimler, zincirdeki amino asit rezidülerinin doğasına bağlıdır. Apolar gruplar birbirlerinden ziyade su ile daha iyi etkileşime geçebilirler; bu yüzden ΔH_{zincir} değeri katlanmamış duruma oldukça uygundur. Diğer taraftan, su molekülleri apolar açık yan zincirlerden ziyade diğer su molekülleri ile daha iyi etkileşim kurabildikleri için, $\Delta H_{\text{çözücü}}$ katlanmış durum için daha az uygundur. Eşitliğin kritik bileşeni, apolar grupların varlığında büyük ve negatif bir değer olan $-T\Delta S_{\text{çözücü}}$ 'dir. Çünkü apolar gruplar su moleküllerini düzenli olmaya zorlamaktadırlar. Bu hidrofobik etki genellikle apolar rezidülerin kürenin iç kısmında kalmasını sağlar. (Baldwin 2007, Makhatadze ve Privalov 1995).



Şekil 1.7. Protein katlanması enerji diyagramı şematik örneği (Dill ve Chan 1997).

Protein katlanması yapısal uzayda genellikle huni benzeri bir yüzeyde meydana gelmektedir. Huninin şekli ve natif duruma (N) karşılık gelen global minimum, başlangıçtaki katlanmamış veya denatüre olmuş durumdan olması gereken katlanma duruma göre değişiklik gösterir. Huninin duvarları tamamen pürüzsüz değildir ve pürüzler katlanmaya engel teşkil eden yerel minimumlardır (Dill ve Chan 1997).

1.3.2. Protein katlanma kinetiği

Kinetik katlanma ile ilgili olarak yapılan çalışmalar katlanma araçlarının ve enerji bariyerlerinin sayısı ve farklı mekanizmalar arasında fark edilebilen önemli etkileşimlerin tanımı hakkında bilgi sağlayabilir. Örneğin, “inen” (*downhill*) veya “azalan” (*run-down*) katlanma, proteinlerin herhangi bir makroskopik serbest enerji bariyeri ile karşılaşmadığı bir protein katlanma sürecidir. Çoğu protein için daha gerçekçi olanı katlanmamış aşamadan (U) asıl konuma (N) tek bir enerji bariyeri ile katlanarak geçişin olduğu “iki aşamalı” katlanmadır (Jackson ve Fersht 1991):



Proteinlerin asıl konuma (N), geçiş durumlarından (I) ulaşmaları ise (1.3) bağıntısındaki gibi ifade edilebilir (Uversky ve Ptysin 1994).



1.4. Veri Tabanları

Proteinlerin yapısal ve fonksiyonel özelliklerini barındıran bir çok veri tabanı bulunmaktadır. Bunların en başında ve en kapsamlı olanı, protein dizisi ve açıklamalı verileri ihtiva eden Universal Protein Resource, (UniProt 2014) (<http://www.uniprot.org>) veri tabanıdır. UniProt veri tabanı, dizi bilgisi, fonksiyonel açıklamalar ve birçok organizmaya, yapıya, fonksiyona, etkileşime, ontolojiye, bölge (domain) ve diğer özelliklere çapraz referans veren SwissProt ve TrEMBL (Artimo ve ark. 2012) (<http://www.expasy.ch/sprot/>) veri tabanlarının üzerine inşa edilmiştir.

Amerikan Ulusal Biyoteknoloji Bilgi Merkezi'nin (NCBI, <http://www.ncbi.nlm.nih.gov>) Gen Bankası ise en kapsamlı bilgi kaynağıdır. Protein Data Bank (PDB) veri tabanı (Berman ve ark. 2000), yapısal proteinlerin yapılarını atomik koordinatları ile paylaşmaktadır. PDB veri tabanında X-ışını kristalografisi ve nükleer manyetik rezonans (NMR) görüntüleme teknikleri ile belirlenmiş 117882'den fazla yapısal bilgi bulunmaktadır (<http://www.pdb.org>). UniProt ve PDB'deki dizi ve yapı bilgisinin etki açısından yorumlanması gerekir. Diğer veri tabanları domainlerin yapısal tanımlarını ve

otonom katlanmalarını açıklamaktadır. Örneğin CATH veri tabanı (Orengo ve ark. 1997) dört farklı seviyede çalışılmış olan protein domain yapılarının hiyerarşik sınıflandırmasını bulundurmaktadır (<http://www.cathdb.info>). SCOP (Murzin ve ark. 1995) ise yapısal özellikler üzerine kurulmuş evrimsel bir sınıflandırmaya dayalı bir veri tabanıdır (<http://scop.mrc-lmb.cam.ac.uk/scop>).

1.4.1. DISPROT veri tabanı

DisProt (Sickmeier ve ark. 2006), deneysel sonuçlardan elde edilmiş 600'den fazla özünde düzensiz yapılı protein hakkında geniş kapsamlı verileri barındıran bir veri tabanıdır (<http://www.disprot.org>). Bu veri tabanında anahtar kelime veya benzer sekans girişi ile arama yapılabilmektedir. Sorgu sonuçlarıyla birçok veri tabanına (UniProt, SwissProt, NCBI, PDB, PubMed) ilişkisel olarak bağlantılı ve yapısal açıklamalı bilgilerle çapraz referans verilmektedir.

1.4.2. CATH protein sınıflandırma veri tabanı

CATH veri tabanı, Protein Bilgi Bankası'ndaki (PDB) protein yapılarının hiyerarşik domain sınıflandırmasını içerir. Protein yapıları otomatik ve manüel prosedürlerin bileşimi yöntemiyle sınıflandırılır. Bu hiyerarşide başlıca dört seviye vardır.

- Sınıf (*Class*): yapılar ikincil yapı kompozisyonları ile sınıflandırılır (çoğunlukla alfa, çoğunlukla beta, alfa/beta karışık ve az ikincil yapı).
- Mimari (*Architecture*): ikincil yapıların birbirlerine göre üç boyutta oryantasyonlarına göre sınıflandırma (ama yapılar arasındaki bağlantılar ihmal edilir).
- Topoloji (=Fold ailesi) (*Topology*): yapılar bu seviyede ikincil yapıların hem genel şekline hem de bağlantı şekillerine göre fold gruplarına ayrılır.
- Homolog üstaile (*Homologous superfamily*): bu seviyede ortak atadan geldiği düşünülen protein domainleri yani homolog domainler gruplandırılır.

CATH ağaç şeklinde bir hiyerarşik sınıflandırmadır. Sınıflandırma ilk C (*Class*) (=Sınıf) seviyesi ile başlar, yani ağaç dallarını sınıflar oluşturur (örneğin alfa veya beta). Daha sonra bu dallar bir sonraki seviyede birbirlerinden farklı gruplara ayrılırlar ve her ayrımda grubun içindeki domainlerin birbirlerine benzerlik oranı artar.

Çizelge 1.3. CATH hiyerarşik sınıflandırma kodları ve kümeleme kriterleri

Derinlik	Harf	İsim	Kümeleme kriteri
1	C	Sınıf (<i>Class</i>)	İkincil yapı içeriği
2	A	Mimari (<i>Architecture</i>)	İkincil yapıların üç boyutta yerleşimleri
3	T	Topoloji	İkincil yapıların üç boyutta yerleşimleri ve aralarındaki bağlantılar (<i>fold</i>)
4	H	Homolog üstaile (<i>Homologous Superfamily</i>)	Manüel inceleme ile ortaya çıkmış evrimsel ilgi
5	S	Dizi ailesi (<i>Sequence Family</i>) (S35)	\geq % 35 dizi benzerliği
6	O	Ortolog Aile (<i>Orthologous Family</i>) (S60)	\geq % 60 dizi benzerliği
7	L	“Benzer” domain (S95)	\geq % 95 dizi benzerliği
8	I	Aynı domain (S100)	% 100 dizi benzerliği
9	D	Domain counter	Benzersiz domainler

1.4.3. RCSB-PDB: protein bilgi bankası

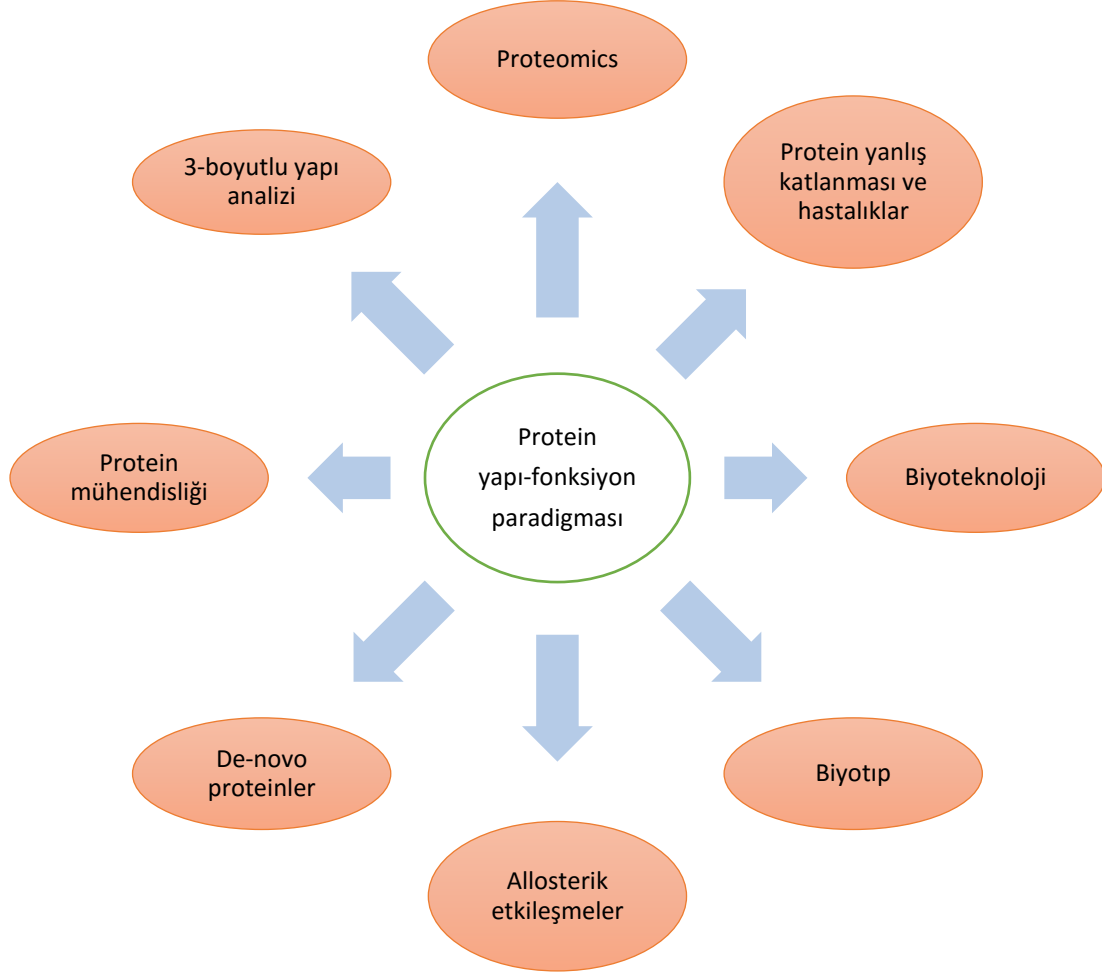
Protein bilgi bankası (Protein Data Bank), farklı organizmalardan elde edilmiş protein ve nükleik asit gibi büyük biyolojik 3-boyutlu yapıların, en yaygın kullanımı olan görüntüleme teknikleri olan X-ışını kristalografisi veya NMR spektroskopisi yöntemleriyle elde edilerek atomların konformasyonel uzaydaki pozisyonlarını temsil eden koordinat verilerini içeren dosyalar bulundurur.

1.5. Özünde Düzensiz Yapılı Proteinler (ÖDYP)

Özünde düzensiz yapılı protein (ÖDYP), sabit veya düzenli 3-boyutlu yapısı olmayan proteindir (Dunker ve ark. 2001, Dyson ve Wright 2005, Dunker ve ark. 2008). ÖDYP'ler kısmen veya tamamen düzensiz olabilirler (Şekil 1.8). Protein yapı-fonksiyon paradigmasının ortaya atılması birçok modern protein bilim dallarını beraberinde getirmiştir. Bunlardan bazıları Şekil 1.9 ile gösterilmektedir.



Şekil 1.8. Özünde düzensiz yapılı protein örneği, C2H2-tipi çinko finger solusyon yapısı proteini (PDB kodu: 1ZR9)



Şekil 1.9. Protein yapı-fonksiyon paradigması sonrası oluşmuş bazı modern bilim dalları

“Anahtar ve kilit” paradigması sayesinde orijinal protein bilimi evreni çoğunlukla iyi tanımlanmış bir yapısı bulunan düzenli proteinler ile ilgilenmiştir. Ancak, en kararlı haldeki proteinler bile katı kristal benzeri oluşumlar olmadığından bu yapılar farklı esneklik uyum dereceleri bulunan dinamik sistemler olarak kabul edilmelidir (Uversky 2013). Aslında protein yapısını stabilize eden kuvvetler zayıftır ve termal dalgalanmalar sebebiyle oda sıcaklıklarında bile kırılabilirler (Uversky ve Dunker 2010, Uversky 2013). Sonuç olarak

proteinlerin X-ışını kristalografisi ile belirlenen 3 boyutlu protein yapıları ve diğer pek çok fiziksel teknik ortalama resimler sunmaktadır (Petsko ve Ringe 2004).

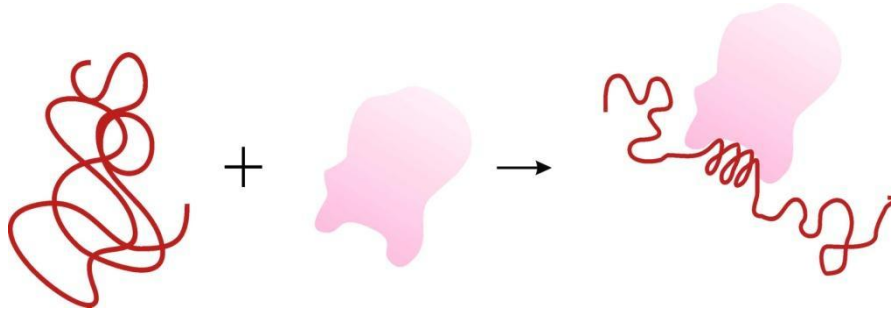
Canlılarda fizyolojik koşullar altında kısmen veya tamamen düzenli bir yapısı olmayan bazı proteinler de bulunmaktadır (Uversky ve Dunker 2010, Dunker ve ark. 1998, Wright ve Dyson 1999, Uversky ve ark. 2000, Dunker ve ark. 2001, Tompa 2002, Daughdrill ve ark. 2005, Dunker ve ark. 2013). Bu proteinlerin ve protein bölgelerinin (ÖDYP) iyi tanımlanmış tek bir denge yapısı bulunmamaktadır. Heterojen konformer topluluklar olan bu proteinlerin veya protein bölgelerinin konformasyonel özelliklerini tanımlamak için Ramachandran açıları tek başına yeterli değildir (Dunker ve ark. 2000, Dunker ve ark. 2013, Ward ve ark. 2004, Uversky ve ark. 2010, Xue ve ark. 2012).

Bu proteinler uzun bir süre içerisinde birbirinden bağımsız olarak tek tek keşfedilmişlerdir ve bu nedenle genel kuralın dışında nadir istisnalar olarak kabul edilmektedirler (Uversky 2013). Sabit yapıların olmadığı biyolojik işlevsellik fenomeni pek çok kez tekrar tekrar keşfedildiğinden ve yazarlar kendi protein ilgi alanlarını yeni terimler ile tanımladığından bu keşiflerin geniş kitleler tarafından fark edilmesi uzun zaman almıştır (Dunker ve ark. 2013). Literatürde bu proteinleri tanımlamak için pek çok terim bulunmaktadır. Bunlardan bazıları yumuşak, bükülen *rheomorphic* (Holt ve ark. 1993), esnek (Pullen ve ark. 1975), mobil (Cary ve ark. 1978), kısmen katlanan (Linderström-Lang ve Schellman 1959), doğası gereği doğal olmayan (Schweers ve ark. 1994), doğası gereği katlanmayan (Uversky ve ark. 2000, Weinreb ve ark. 1996), doğası gereği düzensiz (Daughdrill ve ark. 2005), özünde yapısız (Wright ve Dyson 1999, Tompa 2002), özünde doğal olmayan (Schweers ve ark. 1994), özünde katlanmayan (Weinreb ve ark. 1996), aslen düzensiz (Dunker ve ark. 2001), narin (Chen ve ark. 2008), bukalemun (Uversky 2003), uysal (Fuxreiter ve ark. 2008), 4D (Tsvetkov ve ark. 2008), protein bulutu (Dunker ve Uversky 2010), dans eden proteinler (Livesay 2010), eş bekleyen proteinler (Janin ve Sternberg 2013). Ayrıca “aslen/doğal olarak/kalıtımsal olarak/tabiatı gereği” ifadeleri ile “açık/yapısız/düzensiz/doğal olmayan” ifadelerinin kullanıldığı pek çok farklı kombinasyon bulunmaktadır. Bu nedenle literatürde kullanılan isimlerin çoğu “açık/yapısız/düzensiz/doğal olmayan” durumun, proteinlere ait “asıl/doğal/kalıtımsal/içten gelen” özellikler olduğunu belirtmektedir (Dunker ve ark. 2013). Şekil 1.10 ile ÖDYP’lerin keşfi sonrası ortaya çıkan bazı araştırma konuları gösterilmektedir.



Şekil 1.10. Özünde düzensiz yapılı proteinlerin keşfi ile ortaya çıkan bazı araştırma konuları

ÖDYP'lerin ortaya çıkışı ile, o zamana kadar proteinlerin yapısı ile ilgili söylenen “Proteinin fonksiyonu sabit bir 3-boyutlu yapıya bağlıdır.” paradigması değişti. Yapısal biyoloji alanında yapılan çalışmalar da bu duruma paralel sonuçlarla bir proteinin görevini yapabilmesi için uzayda sabit bir konformasyona sahip olması gerektiği algısını değiştirdi (Wright ve Dyson 1999). ÖDYP'ler sabit 3-boyutlu konformasyonu olmayan fakat fonksiyonel olarak aktif proteinlerin sınıfıdır (Uversky ve ark. 2008). Bazı durumlarda, ÖDYP'ler başka makromoleküllere bağlandıktan sonra (Şekil 1.11) sabit 3-boyutlu yapıya adapte olabilmektedir (Ward ve ark. 2004).



Şekil 1.11. Makromoleküle bağlanmış ÖDYP'nin mevcut 3 boyutlu yapıya adapte olması

1.5.1. Özünde düzensiz yapılı proteinlerin doğal çokluğu

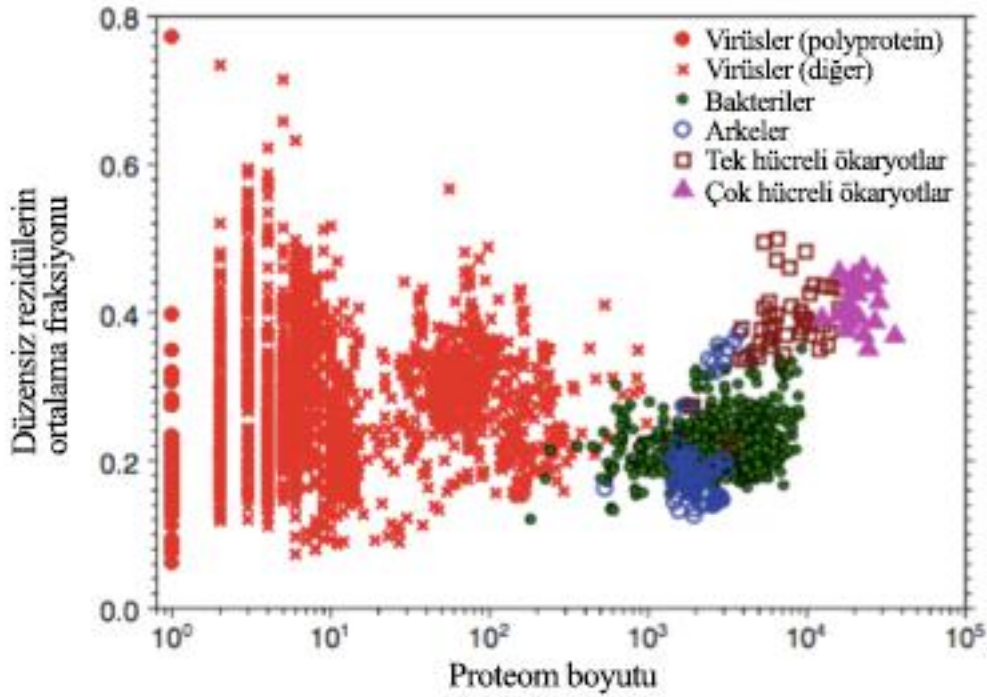
İlk sistematik analizlere göre, proteinlerin özünde düzensizliğinin evrensel bir fenomen olduğu kabul edilirken, 2002 yılından bu yana deneysel çalışmalarla keşfedilen özünde düzensiz yapılı proteinler ve 30-50 rezidürlük kısa polipeptitlerle beraber sayılarının katlanarak artması (Uversky 2002) dikkatleri üzerine çekmiş ve bu yapısız ancak biyolojik olarak aktif proteinlerin yeni bir kural olduğunu farketmişlerdir (Uversky 2000, Uversky 2002).

Sekansa bağlı özünde düzensizlik tayini yapmak için geliştirilen PONDR (Dunker ve ark. 1998, Dunker ve Obradovic 2001) gibi hesaplamalı araçlar kullanılarak (Ferron ve ark. 2006, Bourhis ve ark. 2007, Dosztanyi ve ark. 2007, Dosztanyi ve Tompa 2008, He ve ark. 2009).

ÖDYP ile hibrit proteinlerin dağılımının bulunması için kullanılabilir. Ökaryotların *Caenorhabditis elegans*, *Arabidopsis thaliana*, *Saccharomyces cerevisiae*, ve *Drosophila melanogaster* ile prokaryotlar veya arkedenden daha fazla düzensizlik sergilediği daha önce Uversky tarafından bildirilmiştir (Uversky ve ark. 2014). Ökaryotların proteinlerinin %52-67'sinin ÖDYP olduğu tahmin edilirken; bakteri ve arke proteomlarında bu oranın %16-45 ve 26-51 olduğu tahmin edilmektedir (Dunker ve ark. 2000, Ward ve ark. 2004, Feng ve ark. 2006, Tompa ve ark. 2006, Galea ve ark. 2008, Xue ve ark. 2010, Burra ve ark. 2010). ÖDYP'lerin işlevsel özellikleri arasında genelde sinyal verme, tanımlama ve düzenleme bulunmaktadır. Bunlar ökaryotik ve özellikle çok hücreli ökaryotik organizmalardaki kompleks ve gelişmiş regülasyon ağlarıdır (Uversky ve Dunker 2010, Uversky 2010, Dunker ve ark. 2005, Uversky ve ark. 2005).

Ökaryotlardaki düzensizliğin yüksek olması çalışmalarda sürekli olarak gösterilmektedir. Bu çalışmalar ile tamamlanan proteom analizlerinin sayısı artmakta ve sonuç olarak virüslere ve üç aleme ait yaklaşık 3500 proteomda düzensizlik dağılımı kapsamlı bir şekilde araştırılmıştır (Xue ve ark. 2012). Şekil 1.12'de virüsler, arke, bakteri ve ökaryotlara ait 3,484 türden özünde düzensizlik frekansları ve proteomlar arasındaki korelasyonu gösteren analiz sonuçları temsil edilmektedir. Şaşırtıcı bir şekilde Şekil 1.12, proteom boyutundaki düzensiz rezidürlük fraksiyonunda prokaryotlar ve ökaryotlar arasında iyi

tanımlanan bir boşluk bulunduğunu göstermektedir. Tüm ökaryotların % 32'si veya fazlasında düzensiz rezidüler bulunurken prokaryotik türlerin çoğunda % 27 veya daha az düzensiz rezidü bulunmaktadır (Xue ve ark. 2012). Şekil 1.12'ye göre çok hücreli ökaryotlardaki yapısal düzensizliğin, tek hücreli ökaryotlarda daha az değişken olduğu görülmektedir.



Şekil 1.12. Çeşitli proteomlardaki özünde düzensizlik dağılımları

1.5.2. Özünde düzensizlik tayini

Proteinlerdeki özünde düzensizliği karakterize etmek amacıyla birkaç yöntem kullanılmaktadır. Her bir yöntemin kendisine has güçlü ve zayıf yönleri bulunmaktadır.

1.5.2.1. X-ışını kristalografisi

X-ışını kristalografisi bir kristalin atomik ve moleküler yapısını incelemek için kullanılan ve kristal haline getirilmiş atomların bir X-ışını demetindeki çeşitli yönlerde kırınımı olayına dayanan bir yöntemdir (Tompa 2010). Kırınımına uğrayan bu demetlerin açıları ölçülerek kristaldeki elektronların yoğunluğunun üç boyutlu bir görüntüsü elde edilir. Bu elektron yoğunluğundan kristaldeki atomların kimyasal bağları, kristal yapıdaki düzensizlikler ve bazı başka bilgilerle birlikte ortalama konumları tespit edilebilir (Wikipedia 2015).

1.5.2.2. Nükleer manyetik rezonans spektroskopisi

Nükleer manyetik rezonans (NMR) spektroskopisi (Wüthrich 1986) protein NMR spektroskopisinin alanındaki yenilik kadar, yapısal biyoloji için de önemli bir teknik alanı olmuştur. Bu araştırmaların ortak hedefi X-ışını kristalografisi ile proteinin yüksek çözünürlüklü 3 boyutlu yapısını elde etmektir. X-ışını kristalografisi aksine daha büyük yapılar çözülmüş olsa da, NMR spektroskopisi, genellikle, 35 kDa'lık daha küçük proteinler ile sınırlıdır. NMR spektroskopisi, özünde düzensiz yapılı proteinlerin çok boyutlu yapılarını oluşturmak için kullanılır (Chatterjee ve ark. 2005, Dyson ve Wright 2002, Dyson ve Wright 2004).

1.6. Küresellik Kavramı

Küresellik (*sphericity*), bir objenin ne kadar küresel olduğunun hesaplanmasıdır (Wadell 1935). Böylece bir şeklin tıkkılık (*compactness*) ölçümü yapılmaktadır. Bir objenin küreselliği (Ψ), obje ile aynı hacime (V_p) sahip bir kürenin yüzey alanı ile objenin yüzey alanının (A_p) oranına eşittir. Objenin küreselliği arttıkça bu oran 1'e yaklaşır.

Obje ile aynı hacime (V_p) sahip kürenin yüzey alanı A_s ise:

$$A_s^3 = (4\pi r^2)^3 = 4^3 \pi^3 r^6 = 4\pi(4^2 \pi^2 r^6) = 4\pi 3^2 \left(\frac{4^2 \pi^2}{3^2} r^6\right) = 36\pi \left(\frac{4\pi}{3} r^3\right)^2 = 36\pi V_p^2 \quad (1.4)$$


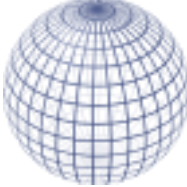
böylece yüzey alanı,

$$A_s = (36\pi V_p^2)^{\frac{1}{3}} = 36^{\frac{1}{3}} \pi^{\frac{1}{3}} V_p^{\frac{2}{3}} = 6^{\frac{2}{3}} \pi^{\frac{1}{3}} V_p^{\frac{2}{3}} = \pi^{\frac{1}{3}} (6V_p)^{\frac{2}{3}} \quad (1.5)$$

sonuç olarak küresellik,

$$\Psi = \frac{A_s}{A_p} = \frac{\pi^{\frac{1}{3}} (6V_p)^{\frac{2}{3}}}{A_p} \quad (1.6)$$

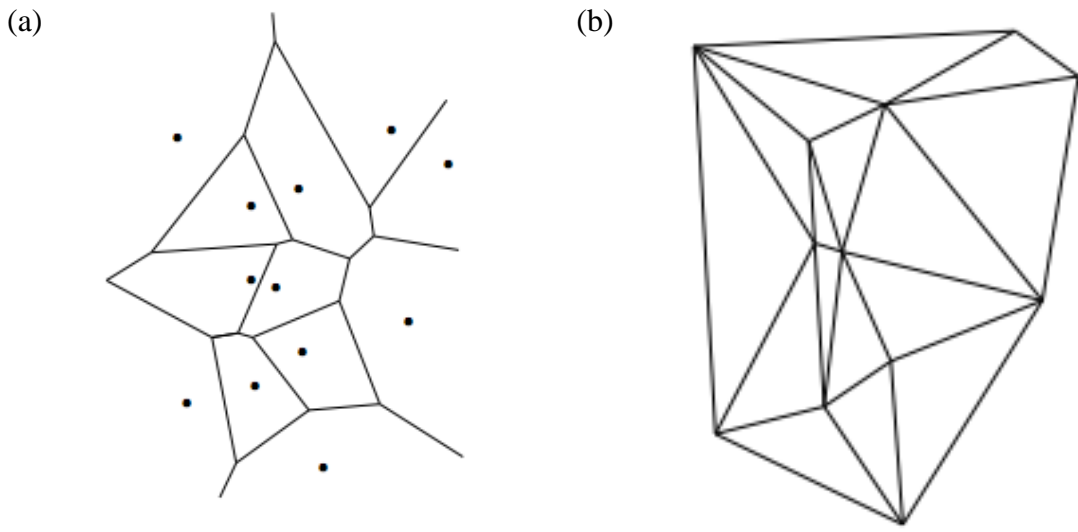
Çizelge 1.4. Bazı 3-boyutlu geometrik şekiller ve küresellik değerleri

Adı	Şekil	Hacim	Yüzey Alanı	Küresellik
tetrahedron (dört yüzlü)		$\frac{\sqrt{2}}{12} s^3$	$\sqrt{3} s^2$	$\left(\frac{\pi}{6\sqrt{3}}\right)^{\frac{1}{3}} \approx 0,671$
sphere (küre)		$\frac{4}{3} \pi r^3$	$4\pi r^2$	1

1.7. Delaunay Teselasyonu

Delaunay teselasyonu hesaplamalı geometride oldukça önemli bir yer tutmaktadır. Bu üçgenlemenin önemini anlayabilmek için geometrik olarak eşleniği sayılabilecek Voronoi diyagramının tanımlanması gerekir. Voronoi diyagramı literatürde Dirichlet, Thiessen veya Wigner-Seitz diyagramı olarak da anılmaktadır (Sukumar ve ark. 2001, Watson ve Philip 1984). Düzlemde yer alan sonlu nokta kümesine ait herhangi bir noktaya, kümedeki diğer noktalardan daha yakın konumda bulunan düzlem noktalarının geometrik yerine o noktanın Voronoi Çokgeni (poligonu) denilmektedir. Kümedeki tüm noktaların Voronoi çokgenlerinin birleşimi, o kümenin Voronoi diyagramını oluşturur.

Şekil 1.13'te bir veri kümesi ve ona ait Voronoi diyagramı görülmektedir. Bu diyagram, en yakın nokta problemleri için kullanılan kesin bir yapıdır. Bir noktanın Voronoi çokgeni herhangi bir noktayı, kendisine en yakın konumdaki komşu noktalardan ayırmaktadır. Çokgenin kenarları, nokta ile komşu noktaları birleştiren doğru parçalarının kenar orta dikmelerinden oluşmakta, her nokta kendisine ait komşu noktalar ile birleştirildiğinde Delaunay Teselasyonu elde edilmektedir. Şekil 1.13'te Voronoi diyagramı verilmiş olan kümenin Delaunay Teselasyonu görülmektedir.



Şekil 1.13. Voronoi Diyagramı (a) ve Delaunay Teselasyonu (b) örneği

1.8. Kolmogorov-Smirnov Testi

χ^2 uygunluk testlerinin alternatifi olan Kolmogorov-Smirnov (K-S) testi, Kolmogorov tarafından 1933 yılında önerilmiştir. Kolmogorov, tek örnek için uyum iyiliği testini önermiştir. 1939 yılında ise Rus matematikçisi olan Smirnov tarafından iki bağımsız örnek için uyum iyiliği testi geliştirilmiştir. Kolmogorov ve Smirnov testi benzerlik nedeniyle, uygulamada Kolmogorov-Smirnov uyum iyiliği testleri olarak bilinirler. χ^2 testinin uygulanabilmesi için beklenen frekansların 5'den büyük olması istenir (Sheskin 2000). Kolmogorov-Smirnov testi böyle bir şarta dayanmadığı için kolayca uygulanabilmektedir (Kartal 1998, Kirkman 1996).

Bilimsel deneylerde, kontrol grubundan elde edilen veriler ile deney grubundan elde edilen veriler karşılaştırılır. Deney grubundaki değerler ile kontrol grubundaki değerler aynı olursa, deney grubunun etkisiz olduğu çıkarımı yapılır. Bu iki grubun sonuçlarının aynı olduğu durumlarda, ne kadar farklı sonuçlar olmalı sorusu ortaya çıkar. Kolmogorov-Smirnov testinde sonuçlara numara tayin edilir. Frekanslar arasındaki anlamlı farklılık P değeri ile gösterilir. Eğer P değeri küçükse sıfır hipotezi reddedilir (Kirkman 1996).

2. MATERYAL ve YÖNTEM

Bu tez çalışması hesaplamalı yapısal biyoloji alanı dahilindedir ve MathWorks MATLAB (Matrix Laboratory) yazılımı kullanılarak hesaplamalar yapılmıştır. Bu tez çalışmasında ÖDYP'lerin PDB'tan elde edilen yapılarının geometrik özellikleri kontrol grubu olarak seçilen alfa ve beta proteinlerin PDB yapılarının geometrik özellikleri ile karşılaştırılmıştır. Bu amaçla her bir protein için PDB yapılarından $C\alpha$ (alfa karbon) atomlarının üç boyuttaki koordinatları Delaunay Teselasyonu ile tetrahedronların uzayda belirlenmesi için kullanılmıştır. Daha sonra bu tetrahedronların küresellik değerleri hesaplanmıştır. Bu tetrahedronların küreselliği regüler tetrahedronda maksimum değer almaktadır (0,671). Tetrahedronların irregülerliği arttıkça bu değer sıfıra yaklaşır. Bir proteindeki tetrahedronların küresellik değerlerinin histogramı o proteine özgüdür. Bu çalışmada kontrol grubu proteinlerinin ve ÖDYP'lerin küresellik değerlerinin dağılımlarının benzerlik ve farklılıklarının incelenmesi ile ÖDYP'lerin kendilerine özgü yapılarına ışık tutulmuştur.

2.1. Kontrol Grubu Proteinleri ve Yapısal Sınıflandırmaları

Kontrol grubu proteinler alfa ve beta proteinlerdir ve daha önce yapılmış bir çalışmada belirlenmiş bir grup protein arasından seçilmişlerdir (da Silveria ve ark. 2009). Bu proteinlerin PDB kodları Çizelge 2.1'de listelenmiştir. Bu proteinlerin CATH sınıflandırmaları da ayrıca belirtilmiştir.

Çizelge 2.1. Kontrol grubu protein yapıları (da Silveria ve ark. 2009)

Grup	Protein Yapıları
Alfa	1LMB, 1B0N, 1M45, 1VRK, 1A7W, 1BGF, 1DK8, 1G33, 1G4I, 1GV2, 1HBK, 1I2T, 1I8O, 1L9L, 1NOG
Beta	1C9O, 1HOE, 1I0C, 1M9Z, 1O5U, 1OH4, 1PMH, 2FCB, 1ROC, 3EZM, 1TUD, 1UMH, 1V05, 1V6P, 1Y0M

Çizelge 2.2. 1LMB kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	1	Mainly Alpha
A	1.10	Orthogonal Bundle
T	1.10.260	434 Repressor (Amino-terminal Domain)
H	1.10.260.40	lambda repressor-like DNA-binding domains

Çizelge 2.3. 1B0N kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	1	Mainly Alpha
A	1.10	Orthogonal Bundle
T	1.10.260	434 Repressor (Amino-terminal Domain)
H	1.10.260.40	lambda repressor-like DNA-binding domains

Çizelge 2.4. 1M45 kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	1	Mainly Alpha
A	1.10	Orthogonal Bundle
T	1.10.238	Recoverin; domain 1
H	1.10.238.10	EF-hand

Çizelge 2.5. 1VRK kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	1	Mainly Alpha
A	1.10	Orthogonal Bundle
T	1.10.238	Recoverin; domain 1
H	1.10.238.10	EF-hand

Çizelge 2.6. 1A7W kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	1	Mainly Alpha
A	1.10	Orthogonal Bundle
T	1.10.20	Histone, subunit A
H	1.10.20.10	Histone, subunit A

Çizelge 2.7. 1BGF kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	1	Mainly Alpha
A	1.10	Orthogonal Bundle
T	1.10.532	Transcription Factor, Stat-4
H	1.10.532.10	Transcription Factor, Stat-4

Çizelge 2.8. 1DK8 kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	1	Mainly Alpha
A	1.10	Orthogonal Bundle
T	1.10.167	Regulator of G-protein Signalling 4; domain 2
H	1.10.167.10	Regulator of G-protein Signalling 4; domain 2

Çizelge 2.9. 1G33 kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	1	Mainly Alpha
A	1.10	Orthogonal Bundle
T	1.10.238	Recoverin; domain 1
H	1.10.238.10	EF-hand

Çizelge 2.10. 1G4I kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	1	Mainly Alpha
A	1.20	Up-down Bundle
T	1.20.90	Phospholipase A2
H	1.20.90.10	Phospholipase A2

Çizelge 2.11. 1GV2 kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	1	Mainly Alpha
A	1.10	Orthogonal Bundle
T	1.10.10	Arc Repressor Mutant, subunit A
H	1.10.10.60	Homeodomain-like

Çizelge 2.12. 1HBK kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	1	Mainly Alpha
A	1.20	Up-down Bundle
T	1.20.80	Acyl-CoA Binding Protein
H	1.20.80.10	-

Çizelge 2.13. 1I2T kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	1	Mainly Alpha
A	1.10	Orthogonal Bundle
T	1.10.1900	c-terminal domain of poly(a) binding protein
H	1.10.1900.10	c-terminal domain of poly(a) binding protein

Çizelge 2.14. 1I8O kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	1	Mainly Alpha
A	1.10	Orthogonal Bundle
T	1.10.760	Cytochrome Bc1 Complex; Chain D, domain 2
H	1.10.760.10	Cytochrome c

Çizelge 2.15. 1L9L kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	1	Mainly Alpha
A	1.10	Orthogonal Bundle
T	1.10.225	NK-Lysin
H	1.10.225.10	Saposin

Çizelge 2.16. 1NOG kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	1	Mainly Alpha
A	1.20	Orthogonal Bundle
T	1.20.1200	Hypothetical Protein Ta1238; Chain: A;
H	1.20.1200.10	-

Çizelge 2.17. 1C9O kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	2	Mainly Beta
A	2.40	Beta Barrel
T	2.40.50	OB fold (Dihydrolipoamide Acetyltransferase, E2P)
H	2.40.50.140	Nucleic acid-binding proteins

Çizelge 2.18. 1HOE kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	2	Mainly Beta
A	2.60	Sandwich
T	2.60.40	Immunoglobulin-like
H	2.60.40.20	alpha-Amylase inhibitor tendamistat

Çizelge 2.19. 1I0C kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	2	Mainly Beta
A	2.30	Roll
T	2.30.30	SH3 type barrels
H	2.30.30.40	SH3 Domains

Çizelge 2.20. 1M9Z kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	2	Mainly Beta
A	2.10	Ribbon
T	2.10.60	CD59
H	2.10.60.10	CD59

Çizelge 2.21. 1O5U kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	2	Mainly Beta
A	2.60	Sandwich
T	2.60.120	Jelly Rolls
H	2.60.120.10	Jelly Rolls

Çizelge 2.22. 1OH4 kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	2	Mainly Beta
A	2.60	Sandwich
T	2.60.120	Jelly Rolls
H	2.60.120.260	Galactose-binding domain-like

Çizelge 2.23. 1PMH kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	2	Mainly Beta
A	2.60	Sandwich
T	2.60.120	Jelly Rolls
H	2.60.120.260	Galactose-binding domain-like

Çizelge 2.24. 2FCB kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	2	Mainly Beta
A	2.60	Sandwich
T	2.60.40	Immunoglobulin-like
H	2.60.40.10	Immunoglobulins

Çizelge 2.25. 1ROC kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	2	Mainly Beta
A	2.60	Sandwich
T	2.60.40	Immunoglobulin-like
H	2.60.40.1490	ASF1-like

Çizelge 2.26. 3EZM kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	2	Mainly Beta
A	2.30	Roll
T	2.30.60	HIV-inactivating Protein, Cyanovirin-n
H	2.30.60.10	HIV-inactivating Protein, Cyanovirin-n

Çizelge 2.27. 1TUD kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	2	Mainly Beta
A	2.30	Roll
T	2.30.30	SH3 type barrels
H	2.30.30.40	SH3 Domains

Çizelge 2.28. 1UMH kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	2	Mainly Beta
A	2.60	Sandwich
T	2.60.120	Jelly Rolls
H	2.60.120.260	Galactose-binding domain-like

Çizelge 2.29. 1V05 kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

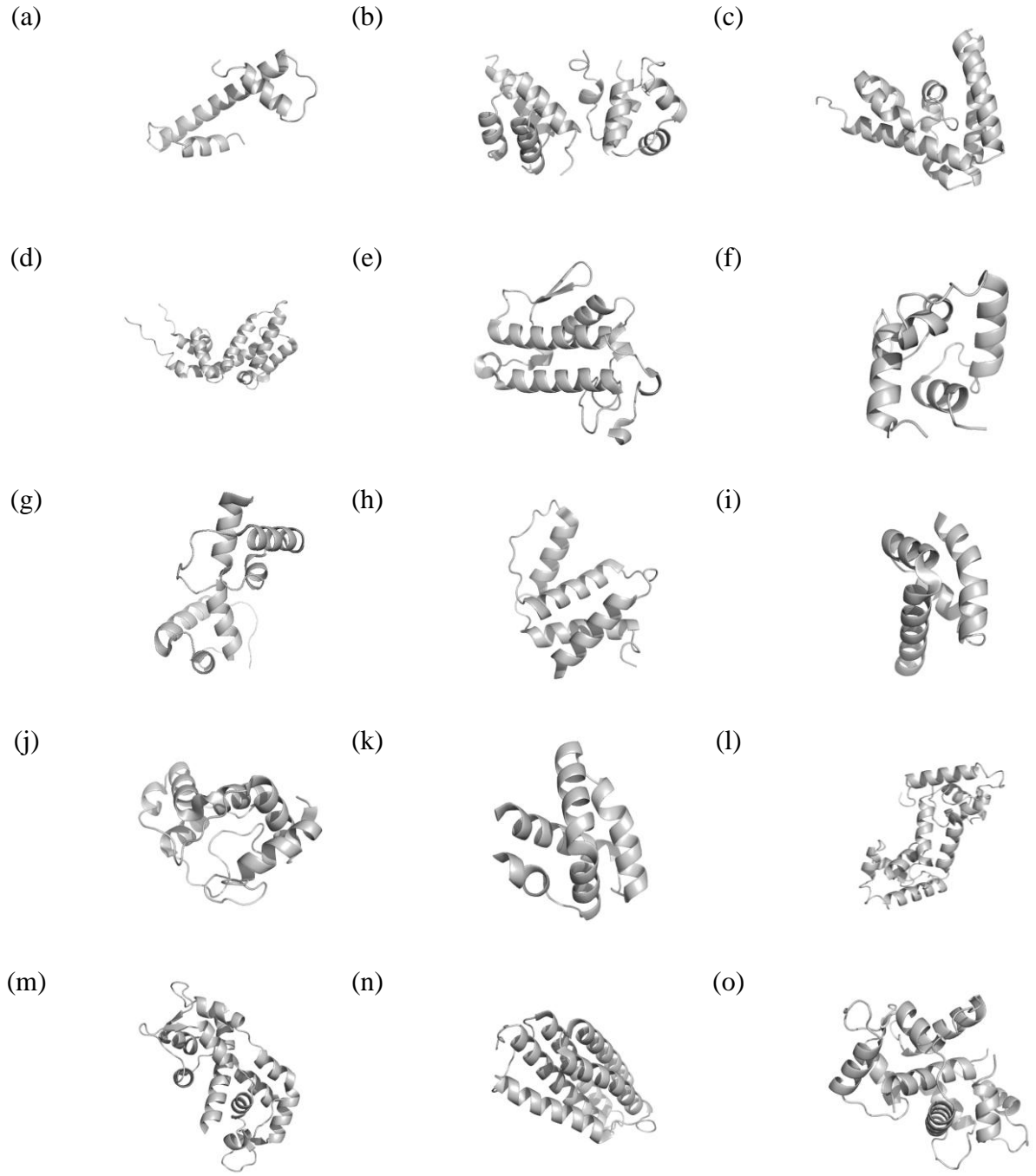
Seviye	CATH Sınıflandırma Kodu	Tanımı
C	2	Mainly Beta
A	2.60	Sandwich
T	2.60.40	Immunoglobulin-like
H	2.60.40.10	Immunoglobulins

Çizelge 2.30. 1V6P kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

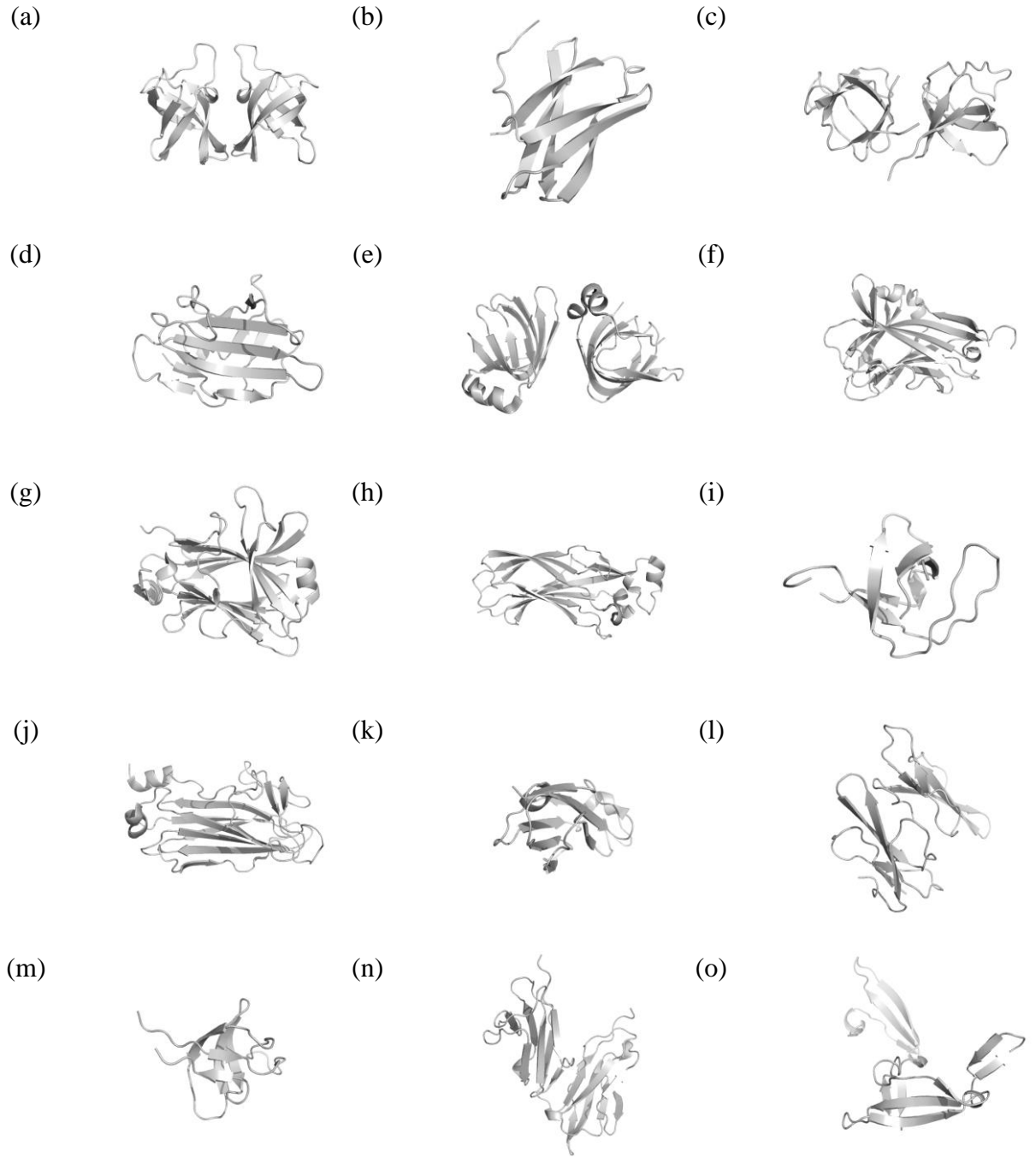
Seviye	CATH Sınıflandırma Kodu	Tanımı
C	2	Mainly Beta
A	2.10	Ribbon
T	2.10.60	CD59
H	2.10.60.10	CD59

Çizelge 2.31. 1Y0M kodlu proteinin CATH sınıflandırma seviyeleri ve tanımları

Seviye	CATH Sınıflandırma Kodu	Tanımı
C	2	Mainly Beta
A	2.30	Roll
T	2.30.30	SH3 type barrels
H	2.30.30.40	SH3 Domains



Şekil 2.1. Alfa proteinlerin 3-boyutlu yapıları. (a) 1A7W, (b) 1B0N, (c) 1BGF, (d) 1DK8, (e) 1G4I, (f) 1G33, (g) 1GV2, (h) 1HBK, (i) 1I2T, (j) 1I8O, (k) 1L9L, (l) 1LMB, (m) 1M45, (n) 1NOG, (o) 1VRK



Şekil 2.2. Beta proteinlerin 3-boyutlu yapıları. (a) 1C9O, (b) 1HOE, (c) 1I0C, (d) 1M9Z, (e) 1O5U, (f) 1OH4, (g) 1PMH, (h) 1ROC, (i) 1TUD, (j) 1UMH, (k) 1V05, (l) 1V6P, (m) 1Y0M, (n) 2FCB, (o) 3EZM

2.2. Özünde Düzensiz Yapılı Proteinlerin Listesinin Hazırlanması

Çevrimiçi açık erişimli DisProt veri tabanından seçilen özünde düzensiz yapılı proteinlerin PDB (Protein Data Bank) veri tabanından elde edilen üç boyutlu yapıları bu araştırmanın materyalini oluşturmaktadır. DisProt veri tabanında 694 ÖDYP listelenmiştir (Mayıs 2013 sürümü). Bunlar arasında PDB yapısı bulunan ve % 100 düzensiz yapılı olan beş protein, % 51 düzensiz yapılı olan bir başka protein bu tez çalışmasında incelenmek için materyal olarak belirlenmiştir. Her birinin Protein Data Bank'ta yapısı araştırılmış, nihai liste Çizelge 2.32 ile verilmiştir .

Çizelge 2.32. İncelenecek olan özünde düzensiz yapılı proteinlerin listesi

DisProt Kodu	PDB Kodu/Zincir	Protein Adı	Düzensizlik	Uzunluk
DP00143	4V4Q/B1	50S ribozomal protein L33	% 100	55
DP00145	4YBB/BL	30S ribozomal protein S12	% 100	124
DP00347	2JU4/A	Retinal rod rodopsin-duyarlı cGMP 3',5'-döngülü fosfodiesteraz alt birim gamma	% 100	87
DP00387	1A6F	Ribonükleaz P protein bileşeni	% 100	116
DP00705	2L9Q	12 kDa ısı şoku proteini	% 100	109
DP00748	2EYY/A	Adaptör molekülü crk	% 51	304

Çizelge 2.33. DP00143 kodlu özünde düzensiz yapılı proteinin künyesi

Eş isimler	RL33_ECOLI
Kaynak organizma	<i>Escherichia coli</i>
Dizi uzunluğu	55
Düzensizlik yüzdesi	% 100
UniProt kodu	P0A7N9
SwissProt kodu	RL33_ECOLI
Natif dizi	<pre> 10 20 30 40 50 60 MAKGIREKIK LVSSAGTGHF YTTTKNKRTK PEKLELKKFD PVVRQHVIYK EAKIK</pre>

Çizelge 2.34. DP00145 kodlu özünde düzensiz yapılı proteinin künyesi

Eş isimler	RS12_ECOLI
Kaynak organizma	<i>Escherichia coli</i>
Dizi uzunluğu	124
Düzensizlik yüzdesi	% 100
UniProt kodu	P0A7S3
SwissProt kodu	RS12_ECOLI
Natif dizi	<pre> 10 20 30 40 50 60 MATVNQLVRK PRARKVAKSN VPALEACPQK RGVCIRVYTT TPKKPNALR KVCVRVLTNG - 60 FEVTSYIGGE GHNLQEHSVI LIRGGRVKDL PGVRYHTVRG ALDCSGVKDR KQARSKYGVK - 120 RPKA</pre>

Çizelge 2.35. DP00347 kodlu özünde düzensiz yapılı proteinin künyesi

Eş isimler	CNRG_BOVIN, GMP-PDE gamma, EC 3.1.4.17, PDE6
Kaynak organizma	<i>Bos taurus</i>
Dizi uzunluğu	87
Düzensizlik yüzdesi	% 100
UniProt kodu	P04972
SwissProt kodu	CNRG_BOVIN
Fonksiyon	Görsel sinyal amplifikasyonu ve iletilmesi işlemlerine katılır. Halkalı guanozin monofosfat (cGMP)-fosfodiesterazlar (PDE), omurgalı çubuk (<i>rod</i>) ve konilerinde (<i>cones</i>) G protein aracılı fototransdüksiyonda efektör moleküllerdir.
Natif dizi	<pre> 10 20 30 40 50 60 MNLEPPKAEI RSATRVMGGP VTPRKGPPKF KQRQTRQFKS KPPKKGVOGF GDDIPGMEGL - 60 GTDITVICPW EAFNHLELHE LAQYGII</pre>

Çizelge 2.36. DP00387 kodlu özünde düzensiz yapılı proteinin künyesi

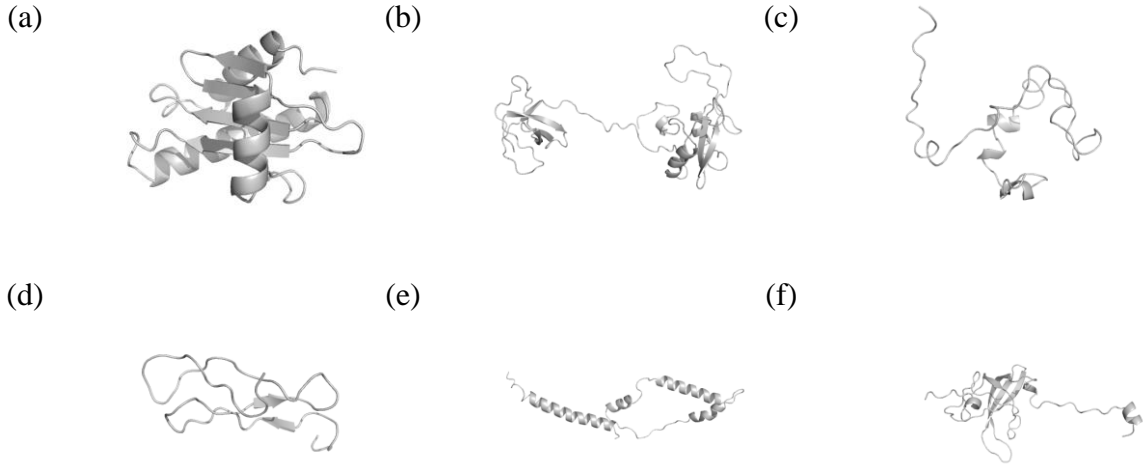
Eş isimler	RNPA_BACSU, RNaseP protein, RNase P protein, EC 3.1.26.5, Protein C5
Kaynak organizma	<i>Bacillus subtilis</i>
Dizi uzunluğu	116
Düzensizlik yüzdesi	% 100
UniProt kodu	P25814
SwissProt kodu	RNPA_BACSU
Fonksiyon	Ribonükleaz P (RNazP) tRNA transkriptlerin prekürsör 5' matürasyonundan sorumlu endoribonükleazdır. Bakterilerde, RNazP katalizör RNA altbirimden oluşur ve holoenzimin substrat spesifikliğı arttıran bir bağlantılı protein altbirimidir. RNazP ön tRNA'dan matür 5' ucunu üretmek için 5'-lider sekansın uzaklaştırılmasını katalize eder. Ayrıca 4.5S RNA gibi diğer RNA substratları da ayırır. Protein bileşeni in vivo ortamda 5'-lider sekansa bağlanması ve ribozimin substrat spesifikliğini genişletilmesi bakımından yardımcı ve önemli bir rol oynar.
Natif dizi	<pre> 10 20 30 40 50 60 MKKRNRLKKN EDFQKVFKHG TSVANRQFVL YTLDPENDE LRVGLSVSKK IGNAVMRNRI - 60 KRLIRQAFLE EKERLKEKDY IIIARKPASQ LTYEETKSL QHLFRKSSLY KKSSSK</pre>

Çizelge 2.37. DP00705 kodlu özünde düzensiz yapılı proteinin künyesi

Eş isimler	HSP12_YEAST, Glucose and lipid-regulated protein
Kaynak organizma	<i>Saccharomyces cerevisiae</i>
Dizi uzunluğu	109
Düzensizlik yüzdesi	% 100
UniProt kodu	P22943
SwissProt kodu	HSP12_YEAST
Fonksiyon	Karbonhidrat kullanım metabolizması ile yağ asidi kullanım metabolizmalarında anahtar bir rol oynar.
Natif dizi	<pre> 10 20 30 40 50 60 MSDAGRKGFG EKASEALKPD SQKSYAEQ GK EYITDKADKV AGKVQPEDNK GVFQGVHDSA - 60 EKGKDNAEQG GESLADQARD YMGAASKLN DAVEYVSGRV HGEEDPTKK</pre>

Çizelge 2.38. DP00748 kodlu özünde düzensiz yapılı proteinin künyesi

Eş isimler	CRK_HUMAN, Proto-oncogene c-Crk, p38, CrkII or CRKII, Crk-II
Kaynak organizma	<i>Homo sapiens</i>
Dizi uzunluğu	304
Düzensizlik yüzdesi	% 51
UniProt kodu	P46108
SwissProt kodu	CRK_HUMAN
Fonksiyon	Crk-I ve Crk-II formlarının biyolojik etkileri farklıdır. Crk-II'nin transform aktivitesi Crk-I'den daha azdır. Crk-II bağlanma ile indüklenen MAPK8 aktivasyonuna, membran fıfırında (<i>ruffling</i>) ve Rac-bağlı hücre hareketine aracılık eder. Apoptotik hücrelerin fagositozunda ve DOCK1 ve DOCK4 ile etkileşimi vasıtasıyla hücre hareketiyle ilişkilidir. EFNA5-EPHA3 sinyalini düzenler.
Natif dizi	<pre> 10 20 30 40 50 60 MAGNFDSEER SSWYWGRLSR QEAVALLQGG RHGVFLVRDS STSPGDYVLS VSENSRVSHY - 60 IINSSGPRPP VFPSPAQPPP GVSPGRLRIG DQEFDSLPLAL LEFYKIHLYD TTTLIEPVSR - 120 SRQSGVILR QEEAEYVRAL FDFNGNDEED LPFKKGDILR IRDKPEEQWW NAEDSEGKRG - 180 MIPVPYVEKY RPASASVSAL IGGNQECSHP QPLGGPEPGP YAQPSVNTPL PNLQNGPIYA - 240 RVIQKRVENA YDKTALALEV GELVKVTKIN VSGQWEGECN GKRGHFPFTH VRLLDQQNPD - 300 EDFS</pre>



Şekil 2.3. Özünde düzensiz yapılı proteinlerin 3-boyutlu yapıları. (a) 1A6F, (b) 2EYY, (c) 2JU4, (d) 4V4Q, (e) 2L9Q, (f) 4YBB

2.3. Irregular Tetrahedron Rezidü Kümelerinin Elde Edilmesi

Sayısal hesaplama platformu olan MathWorks MATLAB yazılımının kendisine özgü programlama dili kullanılarak, her bir protein için PDB yapılarından $C\alpha$ (alfa karbon) atomlarının üç boyuttaki koordinatları, Delaunay Teselasyonu ile tetrahedronların uzayda belirlenmesinde kullanılmıştır.

2.4. ÖDYP'lerin Küresellik Değerlerinin Hesaplanması

Her bir proteine ait irregular tetrahedronların küresellik değerleri hesaplanmıştır ve hesaplanan küresellik değerlerinin histogramları oluşturulmuştur.

3. BULGULAR ve TARTIŞMA

Proteinlerin küresellik hesapları ve ÖDYP'lerin rezidülerinin yerleşim düzenlerine ait hesaplar yapılmıştır.

3.1. Proteinlerin Küresellik Hesapları

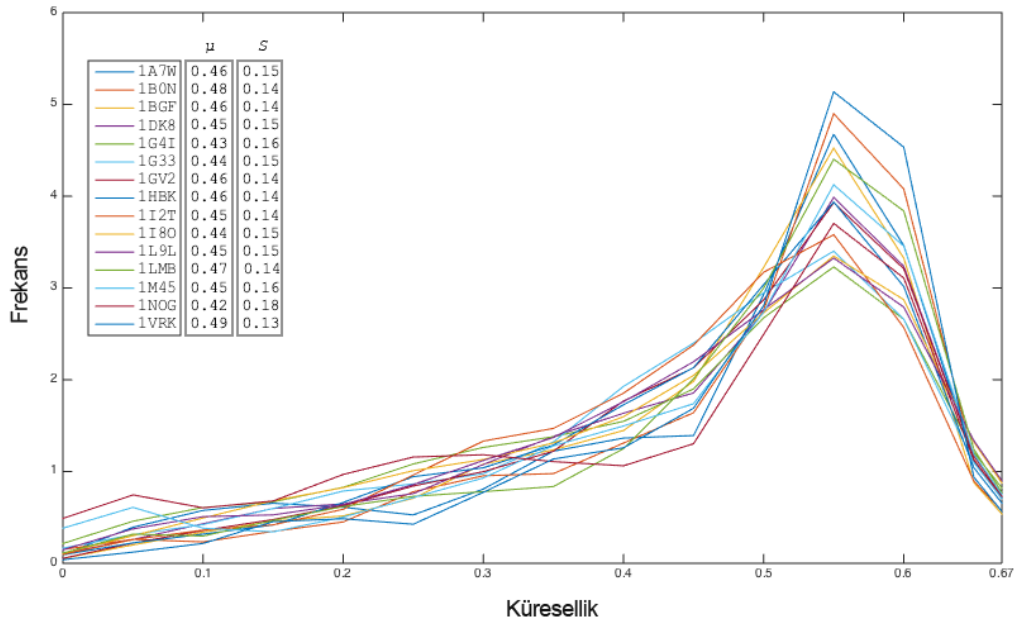
Bu kısımda kontrol grubu ve ÖDYP grubu proteinlerin küresellik hesapları sunulmaktadır.

3.1.1. Kontrol grubu proteinler

Kontrol grubu proteinler alfa ve beta proteinler olarak seçilmiştir. Alfa proteinlerin ve beta proteinlerin küresellik dağılımları hesaplanmış ve Şekil 3.1 ve Şekil 3.2'de gösterilmiştir.

3.1.1.1. Alfa proteinler

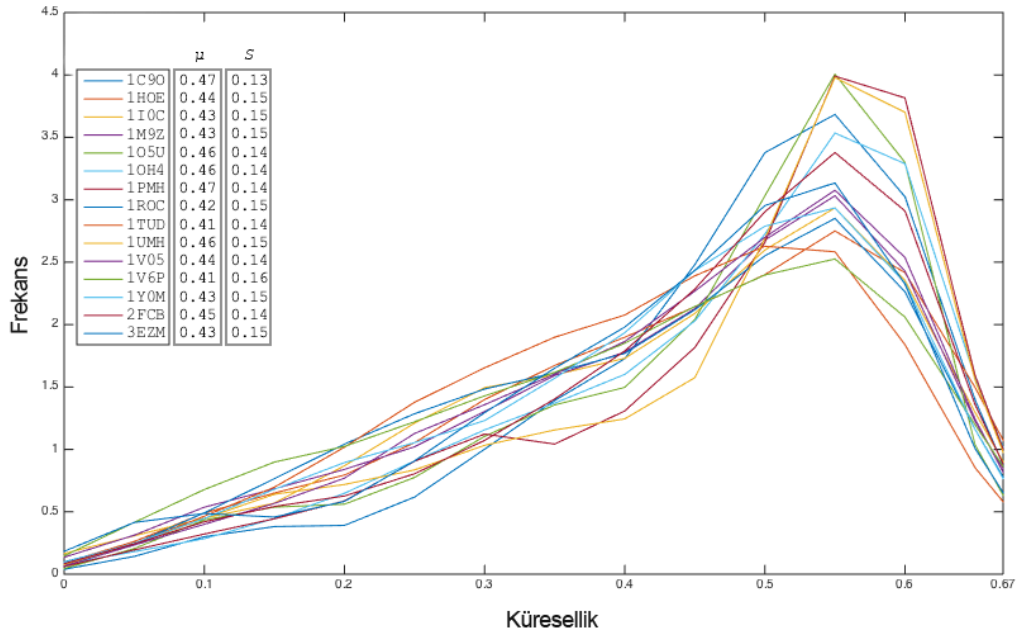
Şekil 3.1 alfa proteinlerdeki tetrahedronların küresellik histogramlarını içermektedir.



Şekil 3.1. Alfa proteinlerdeki tetrahedronların küresellik histogramları

3.1.1.2. Beta proteinler

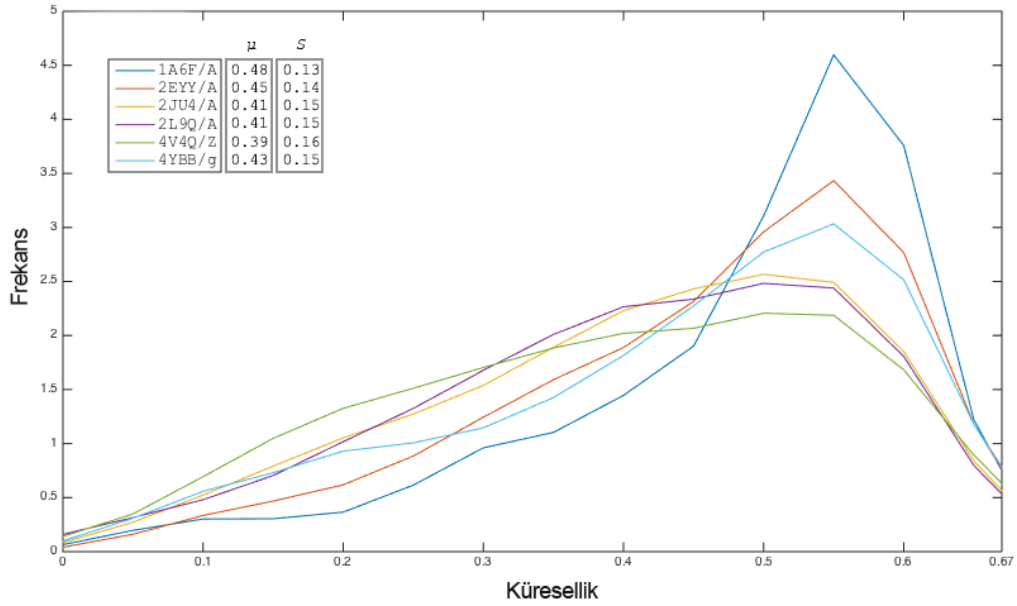
Şekil 3.2 beta proteinlerdeki tetrahedronların küresellik histogramlarını içermektedir. Beta proteinlerin küresellik histogramları ile alfa proteinlerin küresellik histogramları arasında farklılık bulunmaktadır.



Şekil 3.2. Beta proteinlerdeki tetrahedronların küresellik histogramları

3.1.2. Özünde düzensiz yapı proteinler

Şekil 3.3 ÖDYP grubu proteinlerin küresellik histogramlarını içermektedir. Altı ÖDYP'den sadece 1A6F'ye ait histogram alfa grubu proteinlerin histogramlarına benzerlik göstermektedir. Diğerleri kontrol grubu proteinlerden farklılıklar göstermektedir. Bu farklılıkları istatistiksel olarak değerlendirme amacıyla Kolmogorov-Smirnov iki örneklem testine başvurulmuştur. Bu testin sonuçları Çizelge 3.6 ile sunulmuştur.



Şekil 3.3. Özünde düzensiz yapılı proteinlerdeki tetrahedronların küresellik histogramları

3.2. İki Ayrı Proteinde Tetrahedronların Küresellik Dağılımlarının Kolmogorov-Smirnov İki Örneklem Testi İle Karşılaştırılması

Kolmogorov-Smirnov (K-S) iki örneklem testi iki veri setinin istatistiksel olarak anlamlı şekilde farklı olup olmadığını belirlemek için kullanılır. Bu tez çalışmasında K-S iki örneklem testi alfa protein ve beta protein gruplarının kendi içerisinde ne kadar benzer veya farklı olduklarını tespit etmek için ve ÖDYP grubu proteinlerini kontrol grubu ile karşılaştırmak amacıyla kullanılmıştır. Sonuçlar Çizelge 3.1 ile Çizelge 3.6 arasında sunulmuştur. Bu çizelgelerin değerlendirilmesi Çizelge 3.7 ile gösterilmiştir.

3.2.1. Alfa proteinler

K-S iki örneklem testi alfa protein grubu proteinlerin bu çalışmada kullanılan küresellik ölçütü açısından istatistiksel olarak benzer olma sıfır hipotezini test etme amacıyla kullanılmıştır. Çizelge 3.1'de seçilen alfa proteinlerin birbirleri ile benzer olma sıfır

hipotezlerinde kullanılacak olan P deęerleri verilmiřtir. Sıfır hipotezini reddetmek iin kullanılacak P deęeri 0.01 olarak belirlenmiřtir. izelge 3.1’de P ’nin 0.01’den kk olduęu durumlara ait kutular kırmızı renk ile, P ’nin 0.1’den byk olduęu durumlar yeřil ile ve dięer deęerler beyaz ile gsterilmiřtir. Daha sonra verilen izelgelerde de aynı gsterim řekli kullanılmıřtır.

izelge 3.1’de verilen K-S testi sonularının % 26’sında ($r_p < 0.01$, izelge 3.7) P deęerleri 0.01’den kktr. Dolayısıyla P deęeri 0.01 olarak kullanıldıęında alfa proteinler kresellik ltne gre % 74 oranında ($1-r_p < 0.01$, izelge 3.7) birbirlerine benzer bulunmuřtur.

Çizelge 3.1. Alfa proteinlerin küresellik hesaplarının Kolmogorov-Smirnov testi sonucunda elde edilen P değerleri ($\times 10^{-2}$)

		Alfa Proteinler														
		1LMB	1B0N	1M45	1VRK	1A7W	1BGF	1DK8	1G33	1G4I	1GV2	1HBK	1I2T	1I8O	1L9L	1NOG
Alfa Proteinler	1LMB		29.0	4.21	2.16	42.2	26.0	0.09	81.3	0.005	26.8	38.9	2.22	39.5	0.61	1.35
	1B0N			5.46	26.1	14.2	7.02	0.31	43.2	0	83.4	46.2	28.7	26.8	7.10	0.30
	1M45				0.04	62.7	14.8	15.8	58.6	5.01	1.49	36.2	0.56	7.71	0.25	64.7
	1VRK					2.53	0.37	0	10.4	0	12.8	4.20	54.8	7.76	34.9	0.005
	1A7W						74.1	13.0	42.9	3.40	19.8	26.8	6.89	12.5	0.08	38.7
	1BGF							3.77	23.0	0.20	27.2	41.7	1.29	2.86	0.01	2.47
	1DK8								12.8	15.5	0.26	6.63	0.03	0.37	0.02	3.04
	1G33									3.14	37.8	92.6	5.30	50.1	7.91	28.9
	1G4I										0.01	1.74	0.005	0.03	0.003	0.97
	1GV2											11.5	7.37	21.2	14.8	0.38
	1HBK												11.8	35.7	8.44	12.0
	1I2T													13.5	37.5	5.10
	1I8O														13.6	3.30
	1L9L															0.15
	1NOG															
Alfa Proteinlerin CATH Sınıflandırmaları																
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
A	10	10	10	10	10	10	10	10	10	20	10	20	10	10	10	20
T	260	260	238	238	20	532	167	238	90	10	80	1900	710	225	1200	
H	40	40	10	10	10	10	10	10	10	10	60	10	10	10	10	

3.2.2. Beta proteinler

Çizelge 3.2’de verilen K-S testi sonuçlarının % 36’sında ($r_p < 0.01$, Çizelge 3.7) P değerleri 0.01’den küçüktür. Dolayısıyla P değeri 0.01 olarak kullanıldığında beta proteinler küresellik ölçütüne göre % 64 oranında ($1-r_p < 0.01$, Çizelge 3.7) birbirlerine benzer bulunmuştur.

Çizelge 3.2. Beta proteinlerin küresellik hesaplarının Kolmogorov-Smirnov testi sonucunda elde edilen P değerleri ($\times 10^{-2}$)

		Beta Proteinler														
		1C9O	1HOE	1I0C	1M9Z	1O5U	1OH4	1PMH	2FCB	1ROC	3EZM	1TUD	1UMH	1V05	1V6P	1Y0M
Beta Proteinler	1C9O		2.22	3.93	17.7	23.4	3.08	0.07	4.80	6.73	0	57.5	0.45	32.9	0	28.5
	1HOE			1.12	16.7	0.20	33.9	41.4	2.20	1.47	0	7.23	38.3	6.45	0	48.8
	1I0C				2.22	32.9	7.85	0.01	98.2	96.0	0	53.2	0.04	68.5	0	12.1
	1M9Z					31.5	69.6	24.5	4.47	9.50	0	26.9	65.3	18.3	0	65.9
	1O5U						1.89	0.72	30.8	71.7	0	78.8	0.61	35.2	0	28.3
	1OH4							13.3	13.7	14.9	0	48.3	30.8	53.7	0	63.1
	1PMH								0.01	0.06	0	2.33	56.5	0.64	0	17.7
	2FCB									88.2	0	72.0	0.10	88.8	0	15.8
	1ROC										0	54.5	0.37	84.6	0	31.0
	3EZM											0	0	0	1.47	0
	1TUD												6.30	98.7	0	59.9
	1UMH													2.39	0	26.2
	1V05														0	34.1
	1V6P															0
	1Y0M															
Beta Proteinlerin CATH Sınıflandırmaları																
	C	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	A	40	60	30	10	60	60	60	60	60	30	30	60	60	10	30
	T	50	40	30	60	120	120	120	40	40	60	30	120	40	60	30
	H	140	20	40	10	10	260	260	10	1490	10	40	260	10	10	40

3.2.3. Alfa proteinler ve beta proteinler

Bu çalışmada ortaya çıkan bir sonuç kontrol grubunun küresellik ölçütüne göre homojen olmadığıdır. Alfa protein ve beta proteinlerden oluşan kontrol grubundaki proteinler K-S testini benzerlik açısından sırasıyla % 74 ve % 64 oranlarıyla geçmelerine rağmen, alfa protein ve beta protein gruplarına ait proteinler karşılaştırıldığında benzerlik testini geçme oranı % 46'ya ($1-r_p < 0.01$, Çizelge 3.7) düşmüştür.

Çizelge 3.3. Alfa proteinler ile beta proteinlerin küresellik hesaplarının Kolmogorov-Smirnov testi sonucunda elde edilen *P* değerleri ($\times 10^{-2}$)

		Beta Proteinler														
		1C9O	1HOE	1I0C	1M9Z	1O5U	1OH4	1PMH	2FCB	1ROC	3EZM	1TUD	1UMH	1V05	1V6P	1Y0M
Alfa Proteinler	1LMB	0.13	5.88	0.02	28.2	2.08	4.66	33.6	0.04	0.07	0	4.54	4.63	1.49	0	10.6
	1B0N	0.001	2.27	0	7.01	0.005	0.23	22.1	0	0	0	0.11	10.7	0.01	0	1.38
	1M45	0.78	1.48	0.09	30.6	6.50	5.56	4.73	0.25	0.70	0	10.4	9.65	4.49	0	19.9
	1VRK	0	0.49	0	0.13	0	0.001	0.79	0	0	0	0.006	0.06	0.002	0	0.13
	1A7W	0.78	1.31	0.08	43.7	2.94	8.77	9.22	0.26	0.34	0	3.21	14.6	1.68	0	15.9
	1BGF	1.39	0.16	0.07	31.4	6.15	3.43	1.50	0.26	0.23	0	5.99	2.87	2.79	0	33.4
	1DK8	6.55	0.49	8.73	16.4	50.2	7.58	3.42	19.6	29.7	0	59.8	4.58	42.4	0	13.9
	1G33	1.90	13.0	0.48	42.2	11.0	40.9	31.0	1.06	1.40	0	9.70	57.7	5.64	0	19.9
	1G4I	0.02	0.27	27.7	1.97	3.51	3.36	0.38	23.9	13.7	0	5.97	2.65	12.4	0	1.09
	1GV2	0.04	1.43	0.008	13.6	0.05	0.29	10.4	0.001	0.003	0	0.21	3.26	0.02	0	1.99
	1HBK	0.09	4.49	0.008	27.0	0.44	5.02	26.0	0.02	0.04	0	0.91	44.8	0.45	0	7.79
	1I2T	0	1.35	0	0.40	0.001	0.03	2.15	0	0	0	0.01	1.07	0.002	0	0.08
	1I80	0.01	41.2	0.01	3.99	0.14	1.70	32.2	0.005	0.01	0	1.61	18.4	0.43	0	11.2
	1L9L	0	1.58	0	0.36	0.001	0.01	3.75	0	0	0	0.01	1.10	0.001	0	0.27
1NOG	0.03	0.34	0.001	2.81	0.23	0.79	0.37	0.004	0.01	0	1.11	1.83	0.30	0	3.78	

3.2.4. Alfa proteinler ve özünde düzensiz yapılı proteinler

Alfa protein grubundaki proteinlerin K-S testini benzerlik açısından % 74 oranında geçmesine rağmen, alfa protein ve ÖDYP grubundaki proteinler karşılaştırıldığında benzerlik testini geçme oranı % 18'e ($1-r_p < 0.01$, Çizelge 3.7) düşmüştür. Dolayısıyla ÖDYP grubu proteinler, küresellik ölçütü açısından alfa proteinlerden belirgin bir farklılık göstermektedirler.

Çizelge 3.4. Alfa proteinler ile özünde düzensiz yapılı proteinlerin küresellik hesaplarının Kolmogorov-Smirnov testi sonucunda elde edilen P değerleri ($\times 10^{-2}$)

		Özünde Düzensiz Yapılı Proteinler					
		2JU4	2EYY	4V4Q	4YBB	1A6F	2L9Q
Alfa Proteinler	1LMB	0	0	0.002	0	41.0	0
	1B0N	0	0	0	0	43.7	0
	1M45	0	0.002	0.006	0.002	1.79	0
	1VRK	0	0	0	0	4.40	0
	1A7W	0	0.01	0.004	0.004	28.3	0
	1BGF	0	0.001	0.004	0.001	40.7	0
	1DK8	0	1.26	0.41	0.68	0.57	0
	1G33	0	0.05	0.06	0.03	17.0	0
	1G4I	0	2.06	2.07	7.19	0.02	0
	1GV2	0	0	0	0	83.1	0
	1HBK	0	0	0.001	0.001	23.1	0
	1I2T	0	0	0	0	1.19	0
	1I80	0	0	0.001	0	14.9	0
	1L9L	0	0	0	0	5.34	0
	1NOG	0	0	0	0	0.08	0

3.2.5. Beta proteinler ve özünde düzensiz yapılı proteinler

Beta protein grubundaki proteinlerin K-S testini benzerlik açısından % 64 oranında geçmesine rağmen, alfa protein ve ÖDYP grubundaki proteinler karşılaştırıldığında benzerlik testini geçme oranı % 35'e ($1-r_p < 0.01$, Çizelge 3.7) düşmüştür. Dolayısıyla ÖDYP grubu proteinler, küresellik ölçütü açısından beta proteinlerden de belirgin bir farklılık göstermektedirler.

Çizelge 3.5. Beta proteinler ile özünde düzensiz yapılı proteinlerin küresellik hesaplarının Kolmogorov-Smirnov testi sonucunda elde edilen P değerleri ($\times 10^{-2}$)

		Özünde Düzensiz Yapılı Proteinler					
		2JU4	2EYY	4V4Q	4YBB	1A6F	2L9Q
Beta Proteinler	1C9O	0	0.84	0.45	0.86	0.02	0
	1HOE	0	0.01	0.25	0.09	5.41	0
	1I0C	0	70.5	14.9	32.3	0.001	0
	1M9Z	0	0.31	0.08	0.03	17.9	0
	1O5U	0	5.72	1.34	1.62	0.15	0
	1OH4	0	0.09	0.27	0.45	1.86	0
	1PMH	0	0	0.001	0	31.6	0
	2FCB	0	50.9	7.78	29.2	0.01	0
	1ROC	0	39.5	8.36	16.1	0.01	0
	3EZM	38.9	0	3.64	0.001	0	2.10
	1TUD	0.001	66.7	9.49	30.8	0.50	0
	1UMH	0	0.001	0.004	0.001	3.27	0
	1V05	0	17.8	9.82	17.2	0.19	0
	1V6P	1.24	0	3.04	0	0	0.06
1Y0M	0	2.84	0.66	4.22	6.28	0	

3.2.6. Özünde düzensiz yapılı proteinler

Özünde düzensiz yapılı proteinlerin birbirlerine olan K-S benzerlik testlerini geçmeleri %27 ($1-r_p < 0.01$, Çizelge 3.7) oranında olmuştur. Dolayısıyla özünde düzensiz yapılı proteinler hem kontrol grubundaki alfa grubundaki ve beta grubundaki proteinlere, hem de kendi grubundaki diğer proteinlere benzerlik göstermemiştir.

Çizelge 3.6. Özünde düzensiz yapılı proteinlerin küresellik hesaplarının Kolmogorov-Smirnov testi sonucunda elde edilen P değerleri ($\times 10^{-2}$)

		Özünde Düzensiz Yapılı Proteinler					
		2JU4	2EYY	4V4Q	4YBB	1A6F	2L9Q
ÖDYP'ler	2JU4		40.8	2.53	0	0.01	23.9
	2EYY			0.003	0	0	87.3
	4V4Q				0	0.02	0
	4YBB					0	0
	1A6F						0
	2L9Q						

3.2.7. Kolmogorov-Smirnov testi sonuçlarının değerlendirilmesi

Çizelge 3.1-3.6'daki sonuçlar Çizelge 3.7'de özetlenmiştir. Çizelge 3.7 aşağıda verilen denkleme (3.1) göre hazırlanmıştır.

$$r_i = n_i / (n_{p > 0.1} + n_{p < 0.01} + n_{0.01 \leq p \leq 0.1}), i: p > 0.1, p < 0.01, 0.01 \leq p \leq 0.1 \quad (3.1)$$

Çizelge 3.7. Kolmogorov-Smirnov testi sonuçlarındaki oranlar

Karşılaştırma	$r_{p > 0.1}$	$r_{p < 0.01}$	$r_{0.01 \leq p \leq 0.1}$	$r_{p > 0.1} / r_{p < 0.01}$
Alfa-Alfa	0.46	0.26	0.27	1.81
Alfa-Beta	0.17	0.54	0.28	0.31
Beta-Beta	0.45	0.36	0.18	1.23
Beta-ÖDYP	0.15	0.65	0.18	0.23
Alfa-ÖDYP	0.08	0.82	0.08	0.10
ÖDYP-ÖDYP	0.2	0.73	0.06	0.27

4. SONUÇ ve ÖNERİLER

Bu çalışmada özünde düzensiz yapılı proteinlerin rezidü yerleşme düzenleri hesaplamalı yöntemlerle incelenmiştir. Kullanılan ölçüt, proteinlerin alfa karbonlarının koordinatlarından Delaunay Teselasyonu ile elde edilen tetrahedronların küresellik değerleridir. Bu küresellik değerleri tetrahedronları regüler ve irregüler olarak sınıflandırmıştır çünkü en yüksek olası değer regüler tetrahedrona (düzgün dörtyüzlü) aittir. Kontrol grubu olarak seçilen proteinlerin (alfa proteinler, beta proteinler) rezidü yerleşme düzenlerini incelemek için oluşturulan tetrahedronların küresellik değerlerinin dağılımlarının homojen olmadıkları saptanmıştır. Kontrol grubu dolayısı ile alfa proteinler ve beta proteinler olarak iki ayrı gruptan oluşmaktadır.

Kontrol grubu proteinlerin ÖDYP grubu proteinler ile K-S testi ile karşılaştırılmaları ile ÖDYP grubu proteinlerin hem alfa hem de beta proteinlerden küresellik ölçütü kullanıldığında farklılık gösterdiği bulunmuştur. Özünde düzensiz yapılı proteinlerin ise birbirleri ile bir grup oluşturmak yerine kendine özgü bir yapı sergiledikleri sonucuna Çizelge 3.7'nin incelenmesi ile varılabilir.

İleride yapılacak çalışmaların daha fazla sayıda kontrol grubu ve ÖDYP içermesi yerinde olacaktır. Ayrıca alfa+beta ve alfa/beta proteinler de kontrol grubu proteinler arasında olmalıdır.

5. KAYNAKLAR

- Anfinsen CB (1973). Principles that govern the folding of protein chains. *Science*, 181(4096): 223–230.
- Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, Duvaud S, Flegel V, Fortier A, Gasteiger E, Grosdidier A, Hernandez C, Ioannidis V, Kuznetsov D, Liechti R, Moretti S, Mostaguir K, Redaschi N, Rossier G, Xenarios I, Stockinger H (2012). ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res*, 40(W1): W597-W603.
- Baldwin RL (2007). Energetics of protein folding. *J. Mol. Biol*, 371: 283–301.
- Barlow DJ, Thornton JM (1988). Helix geometry in proteins. *J. Mol. Biol*, 201: 601-619.
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000). The Protein Data Bank. *Nucleic Acids Research*, 28: 235-242.
- Bourhis JM, Canard B, Longhi S (2007). Predicting protein disorder and induced folding: from theoretical principles to practical applications. *Current Protein and Peptide Science*, 8(2): 135-149.
- Burra PV, Kalmar L, Tompa P (2010). Reduction in structural disorder and functional complexity in the thermal adaptation of prokaryotes. *PLoS One*, 5(8), e12069.
- Cary PD, Moss T, Bradbury EM (1978). High-resolution proton-magnetic-resonance studies of chromatin core particles. *European Journal of Biochemistry*, 89(2): 475-482.
- Chatterjee A, Kumar A, Chugh J, Srivastava S, Bhavesh NS, Hosur RV (2005). NMR of unfolded proteins. *J. Chem. Sci.*, 117: 3–21.
- Chen J, Liang H, Fernández A (2008). Protein structure protection commits gene expression patterns. *Genome Biol.*, 9(7): R107.
- Chiang YS, Gelfand TI, Kister AE, Gelfand IM (2007). New classification of supersecondary structures of sandwich-like proteins uncovers strict patterns of strand assemblage. *Proteins*, 68 (4): 915–921.
- Chothia C (1973). Conformation of twisted beta-pleated sheets in proteins. *J. Mol. Biol.*, 75: 295-302.
- da Silveira CH, Pires DEV, Minardi RC, Ribeiro C, Veloso CJM, Lopes CJD, Meira W, Neshich G, Ramos CHI, Habesch R, Santoro MM (2009). Protein cutoff scanning: A comparative analysis of cutoff dependent and cutoff free methods for prospecting contacts in proteins. *Proteins*, 74: 727–743.
- Daughdrill GW, Pielak GJ, Uversky VN, Cortese MS, Dunker AK (2005). Natively disordered proteins. *Protein folding handbook*, 275-357.
- Dill KA, Chan HS (1997). From Levinthal to pathways to funnels. *Nat. Struct. Biol.*, 4: 10-19.
- Dosztányi Z, Sándor M, Tompa P, Simon I (2007). Prediction of protein disorder at the domain level. *Current Protein and Peptide Science*, 8(2): 161-171.
- Dosztányi Z, Tompa P (2008). Prediction of protein disorder. *Structural Proteomics: High-Throughput Methods*, 103-115.
- Dunker AK, Garner E, Guillot S, Romero P, Albrecht K, Hart J, Obradovic Z (1998). Protein disorder and the evolution of molecular recognition: theory, predictions and observations. *Pac Symp Biocomputing*, 3: 473–84.
- Dunker AK, Romero P, Obradovic Z, Garner EC, Brown CJ (2000). Intrinsic protein disorder in complete genomes. *Genome Informatics*, 11: 161-171.

- Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CM, Hipps KW, Ausio J, Nissen MS, Reeves R, Kang C, Kissinger CR, Bailey RW, Griswold MD, Chiu W, Garner EC, Obradovic Z. (2001). Intrinsically disordered protein. *Journal of molecular graphics & modelling*, 19 (1): 26-59.
- Dunker AK, Obradovic Z (2001). The protein trinity—linking function and disorder. *Nature biotechnology*, 19(9): 805-806.
- Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN (2005). Flexible nets. *Febs Journal*, 272(20): 5129-5148.
- Dunker AK, Silman I, Uversky VN, Sussman JL (2008). Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.*, 18 (6): 756-64.
- Dunker AK, Uversky VN (2010). Drugs for ‘protein clouds’: targeting intrinsically disordered transcription factors. *Current opinion in pharmacology*, 10(6): 782-788.
- Dunker AK, Babu MM, Barbar E, Blackledge M, Bondos SE, Dosztányi Z, Han KH (2013). What’s in a name? Why these proteins are intrinsically disordered: Why these proteins are intrinsically disordered. *Intrinsically Disordered Proteins*, 1(1), e24157.
- Dyson HJ, Wright PE (2002). Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol.*, 12: 54-60.
- Dyson HJ, Wright PE (2004). Unfolded proteins and protein folding studied by NMR. *Chem. Rev.*, 104: 3607-22.
- Dyson HJ, Wright PE (2005). Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.*, 6(3): 197-208.
- Feng ZP, Zhang X, Han P, Arora N, Anders RF, Norton RS (2006). Abundance of intrinsically unstructured proteins in *P. falciparum* and other apicomplexan parasite proteomes. *Molecular and biochemical parasitology*, 150(2): 256-267.
- Ferron F, Longhi S, Canard B, Karlin D (2006). A practical overview of protein disorder prediction methods. *Proteins: Structure, Function, and Bioinformatics*, 65(1): 1-14.
- Fuxreiter M, Tompa P, Simon I, Uversky VN, Hansen JC, Asturias FJ (2008). Malleable machines take shape in eukaryotic transcriptional regulation. *Nature chemical biology*, 4(12): 728-737.
- Galea CA, High AA, Obenauer JC, Mishra A, Park CG, Punta M, Kriwacki RW (2008). Large-scale analysis of thermostable, mammalian proteins provides insights into the intrinsically disordered proteome. *Journal of proteome research*, 8(1): 211-226.
- Garrett R, Grisham CM (2007). *Biochemistry*. Thomson Brooks/Cole, Belmont CA, USA.
- He B, Wang K, Liu Y, Xue B, Uversky VN, Dunker AK (2009). Predicting intrinsic disorder in proteins: an overview. *Cell research*, 19(8): 929-949.
- Holt C, Sawyer L (1993). Caseins as rheomorphic proteins: interpretation of primary and secondary structures of the α S1-, β - and κ -caseins. *Journal of the Chemical Society, Faraday Transactions*, 89(15): 2683-2692.
- IUPAC-IUB Joint Commission on Biochemical Nomenclature (1984). *Eur. J. Biochem.*, 138: 9-37.
- Jackson SE, Fersht AR (1991). Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry*, 30: 10428–10435.
- Janin J, Sternberg MJ (2013). Protein flexibility, not disorder, is intrinsic to molecular recognition. *F1000 Biol Rep*, 5(2).
- Kartal M (1998). Hipotez Testleri. Şafak Yayınevi, Erzurum, Türkiye.

- Kirkman TW (1996). Statistics to Use. <http://www.physics.csbsju.edu/stats/> (erişim tarihi, 19.04.2016).
- Levinthal C (1969). How to Fold Graciously. *Mossbauer Spectroscopy in Biological Systems: Proceedings of a meeting held at Allerton House, Monticello, Illinois*: 22–24.
- Lin MT, Beal MF (2006). Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases. *Nature*, 443: 787-795.
- Linderström-Lang K, Schellman JA (1959). In *The Enzymes*; 2nd ed.; Boyer PD, Lardy H, Myrback K (Eds.) Academic Press, New York, USA.
- Livesay DR (2010). Protein dynamics. *Current opinion in pharmacology*, 10(6): 706-708.
- Makhatadze GI, Privalov PL (1995). Energetics of protein structure. *Adv. Protein Chem.*, 47: 307-425.
- Murzin AG, Brenner SE, Hubbard T, Chothia C (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.*, 247: 536-540.
- Murzin AG, Finkelstein AV (1988). General architecture of the alpha-helical globule. *J. Mol. Biol.*, 204: 749-769.
- Orengo CA, Michie AD, Jones S, Jones DT, Swindells MT, Thornton JM (1997). CATH—a hierarchic classification of protein domain structures. *Structure*, 5: 1093-1108.
- Pauling L, Corey RB, Branson HR (1951). The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci.*, 37(4): 205-211.
- Petsko GA, Ringe D (2004). *Primers in Biology. Protein Structure and Function*. New Science Press Ltd., Sinauer Associates, Inc. Publishers, Blackwell Publishing, London, UK.
- Petsko GA, Ringe D (2008). *Protein Structure and Function*. University of Oxford Press. London, UK.
- Pullen RA, Jenkins JA, Tickle IJ, Wood SP, Blundell TL (1975). The relation of polypeptide hormone structure and flexibility to receptor binding: the relevance of X-ray studies on insulins, glucagon and human placental lactogen. *Molecular and cellular biochemistry*, 8(1): 5-20.
- Ramachandran GN, Ramakrishnan C, Sasisekharan V (1963). Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.*, 7: 95-99.
- Richmond TJ, Richards FM (1978). Packing of alpha-helices: geometrical constraints and contact areas. *J. Mol. Biol.*, 119: 537-555.
- Rose GD, Fleming PJ, Banavar JR, Maritan A (2006). A backbone-based theory of protein folding. *Proc. Natl. Acad. Sci.*, 103: 16623–16633.
- Salemme FR, Weatherford DW (1981). Conformational and geometrical properties of beta-sheets in proteins. I. Parallel beta-sheets. *J. Mol. Biol.*, 146:101-117.
- Schweers O, Schönbrunn-Hanebeck E, Marx A, Mandelkow E (1994). Structural studies of tau protein and Alzheimer paired helical filaments show no evidence for beta-structure. *Journal of Biological Chemistry*, 269(39): 24290-24297.
- Sickmeier M, Hamilton JA, LeGall T, Vacic V, Cortese MS, Tantos A, Szabo B, Tompa P, Chen J, Uversky VN, Obradovic Z, Dunker AK (2006). DisProt: the Database of Disordered Proteins. *Nucleic Acids Res.*, 35: 786-793.
- Sheskin DJ (2000). *Handbook of Parametric and Nonparametric Statistical Procedures Second Edition*, CHAPMAN&HALL/CRC, America.

- Stothard P (2000). The Sequence Manipulation Suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques*, 28: 1102-1104.
- Sukumar N, Moran B, Semenov A, Belikov VV, (2001). Natural Neighbour Galerkin Methods. *International Journal for Numerical Methods in Engineering*, 50: 1-27.
- The UniProt Consortium Activities at the Universal Protein Resource (UniProt) (2014). *Nucleic Acids Res.* 42: D191-D198.
- Tompa P (2002). Intrinsically unstructured proteins. *Trends in biochemical sciences*, 27(10): 527-533.
- Tompa P, Dosztanyi Z, Simon I (2006). Prevalent Structural Disorder in *E. coli* and *S. cerevisiae* Proteomes. *Journal of proteome research*, 5(8): 1996-2000.
- Tompa P (2010). *Structure and Function of Intrinsically Disordered Proteins*. Chapman and Hall/CRC, America.
- Tsvetkov P, Asher G, Paz A, Reuven N, Sussman JL, Silman I, Shaul Y (2008). Operational definition of intrinsically unstructured protein sequences based on susceptibility to the 20S proteasome. *Proteins: Structure, Function, and Bioinformatics*, 70(4): 1357-1366.
- Uversky VN, Ptysin OB (1994). "Partly folded" state, a new equilibrium state of protein molecules: four-state guanidinium chloride-induced unfolding of beta-lactamase at low temperature. *Biochemistry*, 33: 2782–2791.
- Uversky VN, Gillespie JR, Fink AL (2000). Why are "natively unfolded" proteins unstructured under physiologic conditions?. *Proteins: Structure, Function, and Bioinformatics*, 41(3): 415-427.
- Uversky VN (2002). Natively unfolded proteins: a point where biology waits for physics. *Protein Sci.*, 11: 739-756.
- Uversky VN (2003). A protein-chameleon: conformational plasticity of α -synuclein, a disordered protein involved in neurodegenerative disorders. *Journal of Biomolecular Structure and Dynamics*, 21(2): 211-234.
- Uversky VN, Oldfield CJ, Dunker AK (2005). Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. *Journal of Molecular Recognition*, 18(5): 343-384.
- Uversky VN, Oldfield CJ, Dunker AK (2008). Intrinsically Disordered Proteins in Human Diseases: Introducing the D2Concept. *Annual Review of Biophysics*, 37: 215–246.
- Uversky VN, Dunker AK (2010). Understanding protein non-folding. *Biochim Biophys Acta.*, 1804(6): 1231-1264.
- Uversky VN (2013). A decade and a half of protein intrinsic disorder: biology still waits for physics. *Protein Sci.*, 22(6): 693-724.
- Uversky VN, Davé V, Iakoucheva LM, Malaney P, Metallo SJ, Pathak RR, Joerger AC (2014). Pathological Unfoldomics of Uncontrolled Chaos: Intrinsically Disordered Proteins and Human Diseases. *Chem. Rev.*, 114(13): 6844-6879.
- Xue B, Williams RW, Oldfield CJ, Dunker AK, Uversky VN (2010). Archaic chaos: intrinsically disordered proteins in Archaea. *BMC Systems Biology*, 4(Suppl 1), S1.
- Xue B, Dunker AK, Uversky VN (2012). Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life. *Journal of Biomolecular Structure and Dynamics*, 30(2): 137-149.
- Wadell H (1935). Volume, Shape and Roundness of Quartz Particles". *Journal of Geology* 43(3): 250–280, doi:10.1086/624298.

- Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT (2004). Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *Journal of Molecular Biology*, 337(3): 635–645.
- Watson DF, Philip GM (1984). Triangle Based Interpolation, *Mathematical Geology*, 16(8): 779-795.
- Weinreb PH, Zhen W, Poon AW, Conway KA, Lansbury PT (1996). NACP, a protein implicated in Alzheimer's disease and learning, is natively unfolded. *Biochemistry*, 35(43): 13709-13715.
- Wikipedia: The free encyclopedia. (2015, March 13). FL: Wikipedia Foundation, Inc. Retrieved March 13, 2015, from <http://www.wikipedia.org>.
- Wüthrich K (1986). *NMR of Proteins and Nucleic Acids*, Wiley, New York, USA.
- Wright PE, Dyson HJ (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.*, 293: 321-331.

EKLER

ÖZGEÇMİŞ

12.06.1988 tarihinde İzmir’de doğdu. Lise öğrenimini Samsun Milli Piyango Anadolu Lisesi’nde tamamladı. Lisans eğitimini 2011 yılında İzzet Baysal Üniversitesi, Fen-Edebiyat Fakültesi Biyoloji Bölümü’nde tamamladıktan sonra 2013 yılında Namık Kemal Üniversitesi, Fen Bilimleri Enstitüsü, Biyoloji Anabilim Dalında, 2014 yılında ise Marmara Üniversitesi, Fen Bilimleri Enstitüsü, Biyomühendislik Anabilim Dalında yüksek lisans eğitimine başladı.

Halen İstanbul’da yaşayan Rasim Murat Aydınkal, Ali Nihat Gökyiğit Vakfı tarafından finanse edilen Nezahat Gökyiğit Botanik Bahçesi’nde biyolojik çeşitlilik veri tabanları konusunda çalışmalar yapmaktadır.