GRADUATE SCHOOL
ISTANBUL MEDENIYET UNIVERSITY
DEPARTMENT OF BIOENGINEERING

Molecular and Dynamic Investigation of the Interactions of RAC1, a Target in Cancer and Immune Diseases, with TIPE2 and PAK1 Proteins

Master's Thesis

Hacı Aslan Onur İşcil

June 2025

GRADUATE SCHOOL
ISTANBUL MEDENIYET UNIVERSITY
DEPARTMENT OF BIOENGINEERING

Molecular and Dynamic Investigation of the Interactions of RAC1, a Target in Cancer and Immune Diseases, with TIPE2 and PAK1 Proteins

Master's Thesis

Hacı Aslan Onur İşcil

Supervisor

Asst. Prof. Saliha Ece Acuner Zorluuysal

June 2025

THESIS JURY APPROVAL

This Master thesis titled" Molecular and Dynamic Investigation of the Interactions of RAC1, a Target in Cancer and Immune Diseases, with TIPE2 and PAK1 Proteins" written by Hacı Aslan Onur İşcil at the Department of Bioengineering was accepted by our jury.

JURY MEMBERS **SIGNATURE**

Supervisor:

Asst. Prof. Saliha Ece Acuner Zorluuysal

İstanbul Medeniyet University

Other Jury Members:

Prof. Dr. Pemra Özbek Sarıca

Marmara University

Asst. Prof. Semran İpek

İstanbul Medeniyet University

Date of Defence: 17/06/2025

STATEMENTS

Style and Reference Manual Statement

Having reviewed this thesis written under my supervision, I confirm that it has been written in accordance with APA Manual of Style and used its footnote reference format consistently throughout the entire text.

Signature

Asst. Prof. Saliha Ece Acuner Zorluuysal

Declaration of Originality

I hereby declare that all information in this dissertation has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conducts, I have fully cited and referenced all material and results that are not original to this work.

Signature

Hacı Aslan Onur İşcil

GENİŞ ÖZET

Kanser ve Bağışıklık Hastalıklarında Hedef Olan RAC1'in TIPE2 ve PAK1 Proteinleriyle Olan Etkileşimlerinin Moleküler Seviyede ve Dinamik Olarak İncelenmesi

İşcil, Hacı Aslan Onur

Yüksek Lisans Tezi, Biyomühendislik Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Saliha Ece Acuner Zorluuysal

Haziran 2025

Bu tez çalışması, hücre içi sinyal iletiminde merkezi rol oynayan Rac1 proteini ile TIPE2 ve PAK1 gibi kritik düzenleyici proteinler arasındaki etkileşimleri moleküler düzeyde detaylı bir şekilde incelemeyi amaçlamaktadır. Organizmalarda homeostazın sürdürülebilmesi, hücre büyümesi, farklılaşması, hareketi ve ölümü gibi temel biyolojik olayların doğru ve koordineli bir şekilde işlemesi, büyük ölçüde protein-protein etkileşimlerinin sorunsuz işlemesine bağlıdır. Bu etkileşimlerin temel düzenleyicilerinden biri olan küçük GTPaz proteinleri, hücre içi sinyal iletim yollarının başlıca düğüm noktalarını oluşturur. Rac1 (Ras-related C3 botulinum toxin substrate 1), Ras süper ailesine ait Rho protein ailesinde yer alan bir GTPaz enzimi olup, GTP'ye bağlandığında aktif, GDP'ye bağlandığında ise inaktif konumda bulunur. Aktif Rac1, çeşitli efektör proteinlerle etkileşime girerek hücre iskeletinin yeniden düzenlenmesi, hücre polaritesi, göçü, proliferasyonu ve apoptozu gibi çok sayıda süreci düzenler. Yapısal olarak, Rac1 proteini beş korunumlu G motifini (G1–G5), Switch I ve Switch II bölgelerini, Insert bölgesini ve hiper degişken bir 9 C-terminal bölgeyi içerir. Özellikle Switch bölgeleri, GTP/GDP bağlanmasına göre konformasyonel değişiklikler göstererek proteinin partnerlerine bağlanma yeteneklerini belirler. Bu bölgelerde meydana gelen mutasyonlar, Rac1'in

yapısal kararlılığını ve fonksiyonel etkileşimlerini ciddi biçimde etkileyebilir. Tez kapsamında incelediğimiz mutasyonlar ve kısa bir özeti su şekildedir;

- P29S: Diğer tipik Ras onkojenik mutasyonlarının aksine, bu mutasyon GTPaz aktivitesini korur. GTP hidrolizini yapabilmesine rağmen, GDP'den GTP'ye dönüş hızını önemli ölçüde artırır ve böylece sürekli aktif hale gelir.
- F28L: Özellikle sürekli aktif olma özelliğiyle bilinir. Ana nedeni, switch bölgelerindeki esneklik değişikliği sayesinde GTP/GDP değişiminin hızlanmasıdır (hızlı döngü).
- Q61L: Sürekli aktif bir mutasyondur ve Rac1'in GTP hidrolizini engelleyerek onu sürekli "açık" konumda tutar. Ancak bu aşırı stabil aktif durum, normal Rac1 sinyal dinamiklerini taklit edemediği için işlevsel bozukluğa yol acar.
- Q61R: Kazanım mutasyonudur ve P29S mutasyonuna benzer şekilde Rac1'in GTP bağlama yeteneğini artırarak aktivasyonunu kolaylaştırır. Bu da Rac1'in efektör proteinlerle (örneğin PAK1) etkileşim olasılığını yükseltir ve yapısal değişikliklerle sinyal davranışını etkiler.
- T17N: Dominant-negatif mutasyondur. P-loop bölgesinde yer alır ve GTP/GDP bağlanmasını engeller, Rac1'in aktif hale geçmesini engeller. Ayrıca GEF proteinlerine daha güçlü bağlanarak diğer Rac1'lerin aktivasyonunu rekabetçi olarak engeller ve böylece sinyal iletimini baskılar.

TIPE2 (Tumor necrosis factor alpha-induced protein 8-like protein 2), TIPE ailesine ait immün dengeyi sağlayan önemli bir adaptör proteindir. Altı alfa-heliksten oluşan kompakt bir yapı gösteren TIPE2, merkezinde lipid moleküllerle etkileşim kurabilecek geniş bir hidrofobik boşluk taşır. TIPE2, T hücre reseptörleri (TCR) ve Toll-like reseptörler (TLR) gibi immün sinyal yollarını negatif şekilde düzenleyerek, özellikle aşırı immün yanıtların baskılanmasını sağlar. Bu yönüyle otoimmün hastalıkların ve inflamatuar süreçlerin kontrol altına alınmasında önemli bir rol oynar. Literatürde

TIPE2'nin ekspresyonunun birçok inflamatuar hastalıkta azaldığı; özellikle lökosit infiltrasyonu, sitokin üretimi ve NF-κB aktivasyonu ile ilişkili olduğu belirtilmektedir. Son yıllarda yapılan çalışmalar, TIPE2'nin çeşitli kanser türlerinde baskılanmış olduğunu ve bu baskılanmanın tümör progresyonuyla ilişkili olabileceğini göstermiştir. Bu nedenle, TIPE2'nin bağışıklık sistemiyle olan ilişkisi kadar tümör mikro çevresi üzerindeki etkileri de araştırılmaktadır. Ayrıca TIPE2'nin apoptotik 10 süreçlerde rol oynayabileceği, hücresel enerji homeostazına katkı sağladığına dair bulgular da rapor edilmiştir. Ek olarak, TIPE2'nin Rac1 ile olan etkileşimi STRING veri tabanı incelendiğinde, ko-immünopresipitasyon temelli yöntemlerle etkileşimde bulunabileceğine dair veriler elde edilmiştir. Ancak bu etkileşimin detaylı analizi veya etkileşim sonrasındaki yolak regülasyonları konusunda detaylı bir çalışma bulunmamaktadır.

PAK1 (p21-activated kinase 1) proteini ise, Rho GTPaz ailesindeki Rac1 ve Cdc42 gibi küçük GTPaz'ların doğrudan efektörüdür. Kinaz aktivitesine sahip olan PAK1, inaktif formda homodimer bir yapıda bulunur ve oto inhibisyon mekanizması ile kendi aktivitesini baskılar. GTP-bağlı aktif formdaki Rac1'in PAK1'in p21 bağlanma bölgesine (PBD) bağlanması, oto inhibisyon bölgesinin yapısal olarak devre dışı bırakılmasına ve kinaz aktivitesinin başlamasına yol açar. PAK1'in aktive olması, hücre iskeletinin yeniden düzenlenmesinden hücre göçüne, polarite oluşumundan hücre döngüsü ilerlemesine kadar birçok biyolojik olayda kritik rol oynar. Ayrıca PAK1'in MAPK sinyal yolu, PI3K/Akt yolu ve Wnt sinyal yolağı gibi çok sayıda önemli hücresel yolağı da etkileyebildiği bilinmektedir. PAK1'in anormal aktivasyon- onu ise metastatik tümörlerde artış göstermekte ve bu nedenle PAK1 kanser biyolojisinde hem bir belirteç hem de potansiyel hedef protein olarak değerlendirilmektedir.

Bu bağlamda, Rac1, TIPE2 ve PAK1 üçlüsünün etkileşimleri yalnızca moleküler düzeyde sınırlı kalmayıp, hücre düzeyinde çok yönlü bir etki ağı

olusturmaktadır. Örneğin, Rac1'in immün sistem üzerindeki etkileri TIPE2 üzerinden negatif geri besleme ile düzenlenirken, hücre iskeletini yeniden yapılandıran PAK1 ile olan ilişkisi hücre hareketliliği ve metastazda belirleyici olabilir. Rac1'in PAK1 ile etkileşiminin bozulması, kanser hücrelerinin yönlendirilmiş göçünde aksamalara veya alternatif invazyon yollarının aktifleşmesine neden olabilir. Aynı zamanda bu etkileşimlerin bozulması immün sistemin kontrolsüz tepkisine yol açarak otoimmün hastalıkların şiddetlenmesine neden olabilir. Bu nedenle Rac1'in, mutasyonlara bağlı olarak bu iki proteinle olan etkileşimlerini kaybetmesi çok çeşitli hastalıkların temelinde yer alabilecek bir mekanizmayı temsil eder. Yapılan bu çalışma, hem TIPE2'nin hem de PAK1'in Rac1 aracılığıyla ortak bir düzenleme ekseninde buluştuğunu ortaya koyarak, bu eksenin patofizyolojik önemi konusunda yeni sorular doğurmuştur. Bu sorular, ilerleyen donemde deneysel ve klinik araştırmalarla yanıtlanması gereken önemli araştırma başlıklarını oluşturabilir.

Bu moleküler ilişki ağının çözümlenmesi, özellikle hedefe yönelik tedavi strateji- lirinin geliştirilmesinde büyük bir potansiyel taşımaktadır. Mevcut bulgular doğrultusunda, Rac1'in mutant formlarına karsı geliştirilecek küçük moleküller ya da peptidomimetik inhibitörler, Rac1-TIPE2-PAK1 eksenini yeniden düzenleyerek hastalıkların seyrini 11 değiştirebilir. Bununla birlikte, TIPE2 ve PAK1'in ekspresyon düzeylerinin hastalıkların tanı ve prognozlunda biyo-belirteç olarak kullanılıp kullanılamayacağı da araştırılmaya değerdir. Bu tur moleküler hedeflerin belirlenmesi, özellikle kişiselleştirilmiş tıbbın gelişimi acısından önem taşımaktadır. Sonuç olarak, bu çalışmada elde edilen yapısal ve dinamik veriler, temel biyolojiyi aydınlatmanın yanı sıra translasyonel araştırmalar için sağlam bir temel oluşturma potansiyeline sahiptir.

Bu tez çalışmasında, moleküler dinamik simülasyonları ile Rac1, TIPE2 ve PAK1 arasındaki etkileşimler ve bu etkileşimlerin Rac1 mutasyonlarına

bağlı olarak nasıl değiştiği ayrıntılı biçimde analiz edilmiştir. Bu analiz için özellikle Rac1 TIPE2 ve PAK1'in deneysel kristal yapıları Protein Veri Bankası (PDB)'ndan elde edilmiş olup yapıların üzerindeki protein olmayan hetero atom ve benzeri istenmeyen moleküllerin çıkartılması için PyMOL ve manuel yöntemler tercih edilmiştir. Özellikle Rac1'ın tek başına yapısının olmadığı mutant tipler için wild tip (doğal yapı) üzerinden PyMOL ile in silico mutasyonlar gerçekleştirilmiştir.

Yapılar hazırlandıktan sonra eksik yan zincir veya numaralandırmadaki hatalar giderilip moleküler kenetlenme aşamasına geçilmiştir. PAK1-Rac1 ve TIPE2- Rac1 kompleks tahminleri için birbirinden ayrı adımlar izlenmiştir. PAK1 ile olan etkileşimin ortaya çıkartılmasında ilk olarak PAK1'in Rac1 ile aynı ailedeki başka proteinlerle olan deneysel yapıları incelenmiş ve etkileşimlerden bağlanma bölgesi bilgisi elde edilmiştir. Bu elde edilen bilgi kullanılan moleküler kenetlenme programı HADDOCK'a girdi olarak verilmiş ve bu bağlanma arayüzü çerçevesinde kenetlenmiş pozlar elde edilmiştir. PAK1 ile kenetlenme Rac1'in wild tipi başta olmak üzere F28L, P29S, Q61L, Q61R ve T17N mutant yapıları için tekrarlanmıştır. Elde edilen kenetlenme pozları ise HADDOCK sunucusundan gelen skorlara göre sıralanmış ve en iyi skoru veren yapılar, çalışmanın devamındaki moleküler dinamik simülasyon analizlerinde kullanılmak üzere seçilmiştir. TIPE2 ve Rac1 kenetlenmesi içinse Rosetta paketi içerisindeki kenetlenme yazılımı tercih edilmiştir. Bu tercihimizin ana sebebi kor docking çalışması gerçekleştirecek olmamızdır. Rosetta ˙ Docking paketi ile iki protein birbirleri etrafında bizim verdiğimiz eşik değerleri içerisinde kalmak koşulu ile rotasyon ve konum değişiklikleri gerçekleştirir ve en her bir konum için skor tahminleri gerçekleştirir. HADDOCK üzerinde ise bu işlem çok daha kompleks olmakta ve kullanıcı tarafından en bilgiler istenmektedir. Bu yüzden tercihimiz tamamen kor docking gerçekleştirebilen Rosetta tarafında olmuştur.

Bu aşamada TIPE2 ve Rac1'in yapılarının alternatif konumları Rosetta Relax kullanılarak üretilmiş olup bu alternatif yapılar ile Rac1'in GTP ve GDP'li kompleksleri çapraz olarak kenetlenme işlemine tabi tutulmuş ve her bir kenetlenme işlemi için10.000'er model elde edilmiştir. Bu aşamada 3 alternatif yapı içi kendi içlerinde 12 kombinasyonları elde edilmiş ve bu kombinasyonlarının docking gerçekleştirilmiştir Kombinasyonundan 9 farklı set elde edilmiş ve her bir set için10.000 yapılı docking işlemi gerçekleştirilmiştir. Toplamda 90.000 yapı kombinasyonlardan gelirken ek bir docking seti ise Wild yapıların docking 'den gelmiş ve toplamda 100.000 yapı elde edilmiştir. Bu işlem GTP'li yapılar ve GDP'li yapılar için ayrı ayrı gerçekleştirilmiş olup bütün docking aşamalarının sonunda 200.000 yapı elde edilmiştir.

Bu elde edilen yapılar ile doğrudan simülasyona gidilemeyeceği için aralarında en uygun yapının tercih edilmesi aşaması başlamıştır. Bu aşamada Hawdock sunucusu kullanılarak her bir set içerisinde seçilen en iyi bağlanma arayüzü ve toplam skorlarına sahip yapıların MM/GBSA enerjileri analiz edilmiş ve bu verilere dayanılarak Rac1 ve TIPE2 için MD simülasyonlarında kullanılacak başlangıç pozisyonları elde edilmiştir.

Yapı hazırlama ve kenetlenme çalışmaları sonrası Rac1'ın mutasyonlu ve doğal yapıları ile modellenen PAK1 ve TIPE2 komplekslerinin MD simülasyonu aşamasına geçilmiştir. Bu aşamada Amber simülasyon motoru tercih edilmiş olup, simülasyona başlamadan her bir yapı çeşitli kademelerden oluşan minimizasyon ve eşitleme aşamalarından geçirilmiştir. Daha sonrasında ise 3 tekrarlı 1000 nano saniyelik simülasyonların gerçekleştirilebilmesi için hidrojen kütle yeniden dağıtımı denilen bir teknik uygulanmış ve protein üzerindeki hidrojenlerin kütleleri değiştirilmiştir. Bu sayede 2 femtosaniyelik adım boyutu 2 katına çıkartılmış ve 1000 nano saniyelik simülasyonların hesaplama maliyeti düşürülmüştür. Bu hazırlık aşamalarından sonra simülasyonlar koşturulmaya başlanmıştır.

Moleküler Dinamik (MD) simülasyonları, proteinlerin kompleks yapısındaki konformasyonel değişimlerini, bağlanma bölgelerini ve dinamik hareketlerini yüksek zaman çözünürlüğünde ortaya koymuştur. Rac1'in Switch I ve II bölgelerindeki mutasyonların hem proteinin yapısal kararlılığını hem de TIPE2 ve PAK1 ile olan bağlanma afinitesini nasıl etkilendiği sayısal olarak ortaya konmuştur. Ayrıca, GDP/GTP ve Mg2+ gibi kofaktörlerin Rac1'in etkileşimlerindeki rolü ve bu bağlamda işlevselliğinin nasıl değiştiği de detaylı şekilde incelenmiştir. RMSD, RMSF gibi yapısal dinamik analiz metrikleri, mutasyonların protein hareketliliği ve esnekliği üzerindeki etkilerini nicel olarak göstermiştir.

Ek olarak tez kapsamında Ek olarak tez kapsamında gerçekleştirilen simülasyonların genel bir özeti tablo 1'de gösterilmektedir.

Elde edilen önemli bulgular ise Rac1-PAK1 ve Rac1-TIPE2 olarak iki başlıkta incelenebilir.

Table 1: Simülasyonların genel özeti

Sistem	**İkincil Analizler**	**Süre**	**Tekrar**	**Toplam Süre**
Rac1 (Tek başına)				
Wild Type Rac1 (GTP)	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
Wild Type Rac1 (GDP)	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
F28L Rac1 (GTP)	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
F28L Rac1 (GDP)	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
P29S Rac1 (GTP)	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
P29S Rac1 (GDP)	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
Q61R Rac1 (GTP)	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
Q61R Rac1 (GDP)	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
Q61L Rac1 (GTP)	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
Q61L Rac1 (GDP)	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
T17N Rac1 (GTP)	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
T17N Rac1 (GDP)	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
Rac1 – PAK1 Kompleksi				
Wild Type Rac1 (GTP) – PAK1	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
Wild Type Rac1 (GDP) – PAK1	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
F28L Rac1 (GTP) – PAK1	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
F28L Rac1 (GDP) – PAK1	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
P29S Rac1 (GTP) – PAK1	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
P29S Rac1 (GDP) – PAK1	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
Q61R Rac1 (GTP) – PAK1	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
Q61R Rac1 (GDP) – PAK1	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
Q61L Rac1 (GTP) – PAK1	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
Q61L Rac1 (GDP) – PAK1	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
T17N Rac1 (GTP) – PAK1	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
T17N Rac1 (GDP) – PAK1	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
Rac1 – TIPE2 Kompleksi				
Wild Type Rac1 (GTP) – TIPE2	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
Wild Type Rac1 (GDP) – TIPE2	MMGBSA + 2ns Entropi	1000 ns	3	3000 ns
			Toplam	**78 μs**

Rac1-PAK1 etkileşimi:

- F28L Mutasyonu: GDP-bağlı durumda olan mutant Rac1'in PAK1'e olan bağlanma enerjisinde doğal yapısına göre artış gözlemlenmiştir. Bu durum, Switch I bölgesindeki artan esneklikle birlikte değerlendirilmiştir. Genellikle inaktif form olarak bilinen GDP-bağlı Rac1'in, F28L mutasyonu sonucunda PAK1'e bağlanma eğiliminin artabileceği düşünülmektedir.
- Q61L Mutasyonu: GDP-bağlı formdaki mutant Rac1'in PAK1 ile etkileşiminin güçlendiği görülmüştür. Literatürde "aktive edici" mutasyon olarak tanımlanan Q61L, bağlanma enerjisindeki artış sayesinde PAK1 etkileşimini desteklemekte; bu da hücre içi sinyal iletimini sürdürebilir ve hücresel proliferasyon ya da hareketliliğin artmasına neden olabilir.
- Q61R Mutasyonu: Hem GDP- hem de GTP-bağlı formlardaki mutant Rac1'in bağlanma enerjisinde düşüş kaydedilmiştir. Bu düşüş, PAK1 ile zayıflamış bir etkileşime işaret etmektedir. Ayrıca Switch I bölgesinde artan esneklik, bağlanma yüzeyindeki yapısal çeşitliliği artırarak kararlı bir etkileşimi engelleyebilir. Bu onkojenik mutasyonun Rac1-PAK1 etkileşimini zayıflatarak hücresel sinyal yollarında bozulmalara yol açabileceği öne sürülmektedir.
- Switch I Bölgesi Esnekliği: Switch I bölgesindeki esneklik, PAK1'e bağlanmada kritik bir rol oynamaktadır. Yüksek esneklik genellikle bağlanma afinitesinin azalmasıyla ilişkilenirken, düşük esneklik ve daha stabil konformasyonlar daha güçlü bağlanmayı desteklemektedir (Özellikle Q61R mutantı için bu geçerli iken, istisnai olarak F28L de artan esneklikle birlikte PAK1 bağlantıda iyileşme görülmüştür).
- Nükleotid Bağlanma Durumu: GTP veya GDP'ye bağlı olma durumu, mutasyonların etkisini büyük ölçüde belirlemektedir. Örneğin, F28L mutasyonu özellikle GDP bağlıyken etkileşimi artırırken, GTP bağlı halde bu etkileşim azalmaktadır. Q61L mutasyonu ise özellikle GDP bağlı durumda etkileşimde artış gösterirken, Q61R mutasyonu nükleotid

bağlılığına bakılmaksızın etkileşimi zayıflatmakta ve bu durum hücresel sinyal iletiminde farklı sonuçlara yol açmaktadır.

Rac1-TIPE2 etkileşimi:

Yapılan her 3 tekrar simülasyonun (3 run Rac1(GTP)-TIPE2 ve 3 run Rac1(GDP)-TIPE2 olmak üzere) 2 tanesinde iki protein arasında kopma gözlemlenmiş ancak bu kopmalardan birisinde alternatif bir bağlanma konumu ortaya konmuştur. Ancak entropi katkıları da hesaba katıldığında, altı simülasyonun tamamı pozitif bağlanma enerjisi göstermiştir; bu da termodinamik açıdan bağlanmanın elverişsiz olduğunu ortaya koymaktadır. Bu yüzdenden bağlantı konusunda tam bir yorumda bulunmanın zor olacağı görülmüştür. Ancak TIPE2'nin ortasında yer alan büyük cep bölgesi değerlendirildiğinde buraya bağlanabilecek bir küçük molekül ya da peptit parçasının TIPE2'nin stabilitesini değiştirerek Rac1 ile olan etkileşimi düzenleyebileceği yorumunda bulunulmuştur.

Sonuç olarak ise Rac1'in üzerinde gerçekleşen mutasyonların PAK1 ile olan etkileşimini ciddi bir şekilde etkileme potansiyeline sahip olduğu gözlemlenmiş ve değişimler moleküler seviyede incelenmiştir. Bu sonuçların, mutasyonlu Rac1 yapılarını hedef alan ilaç moleküllerinin geliştirilmesine katkısı olacağı düşünülmektedir. Rac1 ile TIPE2 arasındaki etkileşimin ortaya konabilmesi için daha fazla araştırmaya ihtiyaç olduğu tespit edilmiş; gelecek çalışmalarda özellikle TIPE2 veya Rac1 üzerinde gerçekleştirilecek modifikasyonlar ya da ikincil/üçüncül proteinlerin bağlanması yoluyla bu etkileşimin daha kararlı bir şekilde kurulup kurulamayacağının detaylı biçimde incelenmesi önerilmektedir.

Anahtar Kelimeler: Rac1, PAK1, TIPE2, Moleküler Dinamik Simülasyon, protein- protein etkileşimi, mutasyon etkileri

ABSTRACT

Molecular and Dynamic Investigation of the Interactions of RAC1, a Target in Cancer and Immune Diseases, with TIPE2 and PAK1 Proteins

İşcil, Hacı Aslan Onur

Master's Thesis, Department of Bioengineering

Supervisor: Asst. Prof. Saliha Ece Acuner Zorluuysal

June 2025

Intracellular signal transduction and protein-protein interactions are crucial for the proper functioning of biological systems, regulating key processes such as metabolism, environmental response, and cell cycle control. Disruptions or mutations in these proteins or their interactions can lead to various pathological conditions, including cancer and immune dysfunctions. In this context, Rac1 is a small GTPase involved in regulating the cytoskeleton; Pak1 is a serine/threonine kinase activated by Rac1 that controls cell motility, polarity, and proliferation; TIPE2 is a regulatory protein that suppresses the immune response by balancing cell signaling pathways to keep inflammation under control. Notably, TIPE2 acts as an inhibitor of both inflammation and cancer. In this study, the interactions between Rac1 (Ras-related C3 botulinum toxin substrate 1), TIPE2 (Tumor necrosis factor alpha-induced protein 8-like protein 2), and PAK1 (p21-activated kinase 1) were investigated, along with the effects of Rac1 mutations on these interactions. Molecular dynamics simulations were employed to analyze the impacts of mutations on protein structural dynamics and binding mechanisms. Metrics such as RMSD and RMSF provided insight into protein flexibility and conformational changes. Binding energies and structural analyses revealed that mutations in Rac1 differentially modulate its interactions with PAK1 and TIPE2 depending on the nucleotide-bound state. Specifically, the F28L mutation increased the binding affinity to PAK1 in the GDP-bound form, likely due to increased flexibility in the Switch I region. In contrast, the Q61R mutation

weakened Rac1-PAK1 interactions in both GDP- and GTP-bound forms, accompanied by increased Switch I flexibility, potentially disrupting cellular signaling. Simulations involving Rac1 and TIPE2 showed structural displacements under certain conditions, with one GTP-bound run indicating promising conformations. However, entropy calculations suggested thermodynamically unfavorable binding across all parallel runs. Overall, the study highlights the critical role of Switch I flexibility and the bound nucleotide in regulating Rac1 interactions, providing insights into the functional consequences of Rac1 mutations in cellular signaling pathways.

Keywords: Rac1, PAK1, TIPE2, molecular dynamics simulations, protein-protein interactions, mutation effect

ACKNOWLEDGEMENTS

Felix qui potuit rerum cognoscere causas

Atque metus omnes, et inexorabile fatum

Subjecit pedibus, strepitumque Acherontis avari.

Virgil, Georgics, Book II, lines 490–492

(29 BC).

First and foremost, I would like to express my deepest gratitude to my mother Konca Çevik, my grandmother Zehra Çevik, and my late grandfather Bektaş Çevik for raising me with endless love and support. I know that I would not be where I am today without them. Whenever I had a dream or an idea, they were always there to support and encourage me.

However, my concept of family extends beyond biological ties. I would especially like to thank my advisor, Asst. Prof. Dr. Saliha Ece Acuner Zorluuysal. She has not only been my academic mentor but also a mother, a sister, and a dear friend to me. I have always felt her unwavering support. Being by her side and trying to reciprocate her guidance (even though it may be impossible) will remain a lifelong moral responsibility for me.

I also want to express my heartfelt thanks to my dear friends Ecem Bulut and Damla Nur Çamlı, who have become like sisters to me. From the very first day I arrived in Istanbul and at Medeniyet University, you embraced me without hesitation and treated me as one of your own. We shared our sorrows and found comfort together. You will always remain my sisters, and I am so grateful to have met you.

Lastly, I would like to thank Hacer Nur Eksi, my companion since our under- graduate years. We laughed together, had fun, and collected many

memories, both joyful and challenging. We stood strong against life's difficulties. I can never thank you enough for everything you've brought into my life.

CONTENTS

1.ABSTRACT XII
2.INTRODUCTION 1
3.LITERATURE REVIEW 2
3.1 Rac1 2
3.1.1 Rac1 Structural Properties 2
3.1.2 Rac1 Mutations 5
3.2 TIPE2 8
3.2.1 TIPE2 Topology 9
3.3 PAK1 12
3.4 Rac1, PAK1 and TIPE2 Interactions 14
4. MATERIALS AND METHODS 18
4.1 Structure Preparation for Molecular Docking and MD Simulations 19
4.2 Modeling PAK1 and TIPE2 complexes with Rac1 19
4.2.1 Modeling Rac1-TIPE2 Complex 19
4.2.2 Modeling Rac1-PAK1 Complex 21
4.3 Molecular Dynamics Simulations 24
4.3.1 Force Field And Parameters..... 24
4.3.2 Hydrogen Mass Repartationing (Hmr) 24
4.3.3 Steps of Molecular Dynamics Simulation Runs 26
4.3.4 Analyses Performed After MD Simulations..... 29
4.3.5 Summary Of MD Simulations And Analyses..... 32
4.4 Visualization of Structures and Cavity Detection 33
5. RESULTS 34
5.1 Modeling Protein-Protein Complexes 34
5.1.1 Binding Residues 34
5.1.2 Docking with Rosetta 42
5.2 Analysis of MD Simulations 44
5.2.1 RMSD and RMSF 44
5.2.2 Binding Dynamics Analysis with MM/GBSA and Binding Free Energy ... 52
6. DISCUSSION AND CONCLUSION 75
7. FUTURE WORK 78
BIBLIOGRAPHY 79

List of Figures

Figure 3.1 3D Structure and Functional Domains of Rac1.
A - Key motifs from G1 to G5 and the positions of GTP/GDP and Mg.
B - Hypervariable region 4

Figure 3.2 Topological and Secondary Structures of Rac1.
(A) Secondary structures and the positions of key motifs on these structures.
(B) Topological representation showing the connections and order of alpha helices and beta sheets. 5

Figure 3.3 Mutations and Locations in the Rac1.
……..... 8

Figure 3.4 3D Structure and Topology of TIPE2.
(A) Hydrophilic and hydrophobic regions.
(B) Helices and their positions.
(C) Cavity between helices. 11

Figure 3.5 Domains and Activation Cycle of PAK1.
A - Key motifs and their positions in PAK1.
B - Representation of the inactive and active cycle of PAK1 13

Figure 3.6 Potential Interaction Partners of Rac1, TIPE2, and PAK1.
The thickness of the lines on the graph represents the reliability of protein-protein interactions and the abundance and diversity of information about these interactions.
A - For Rac1, B - For TIPE2, C - For PAK1 17

Figure 4.1 Overall workflow of the study 18

Figure 4.2 Workflow of the Docking and Post-Docking Scoring Processes for Rac1 and TIPE2 20

Figure 4.3 PAK1 - Rac3 Crystal Structure (PDB ID: 2QME)
In the figure, the molecule GCP is visualized in stick representation, the green sphere represents magnesium, the wheat-colored region corresponds to the PAK1 protein, and the light green region indicates the RAC3 protein
..... 23

Figure 4.4 Hydrogen Mass Repartitioning (HMR) Process. The colors and their corresponding atoms are as follows: yellow-carbon, white-hydrogen, gray-repartitioned hydrogen, blue-nitrogen, and red-oxygen 25

Figure 5.1 Interface Residues for Rac3 and PAK1 Proteins
A- Active residues shown in raspberry color on Rac1 protein (shown on the 2QME structure)
B- Active residues shown in raspberry color on PAK1 protein 34

Figure 5.2 Interface Residues for Rac1 and TIPE2 Proteins
A- Active residues shown in raspberry color on Rac1 protein
B- Active residues shown in raspberry color on TIPE2 protein 35

Figure 5.3 Docking Pose Selected to Study the Interaction Between Rac1 and TIPE2 42

Figure 5.4 RMSD plots of Rac1 trajectories across various mutations and wild type (F28L, P29S) 46

Figure 5.5 RMSD plots of Rac1 trajectories across various mutations (Q61R, Q61L, T17N) 47

Figure 5.6 RMSD plots of trajectories for docking structures of Rac1 and PAK1 across wild type and various mutations (F28L, P29S) 48

Figure 5.7 RMSD plots of trajectories for docked Rac1 and PAK1 structures across various mutations (Q61L, Q61R, T17N) 49

Figure 5.8 Panels A to C show the first and last frames of the three runs of Rac1(GTP) and TIPE2, while panels D to F show the first and last frames of the three runs of Rac1(GDP) and TIPE2 51

Figure 5.9 RMSD plots of Rac1(GTP)–TIPE2 and Rac1(GDP)–TIPE2 simulations 52

Figure 5.10 RMSF plots of Rac1 trajectories across various mutations and wild type (F28L, P29S) 54

Figure 5.11 RMSF plots of Rac1 trajectories across various mutations (Q61R, Q61L, T17N) 55

Figure 5.12 RMSF plots of Rac1 trajectories across various mutations and wild type (F28L, P29S) 56

Figure 5.13 RMSF plots of Rac1 trajectories across various mutations (Q61R, Q61L) 57

Figure 5.14 MMPBSA Energy Data for GTP- and GDP-bound Rac1 Wild Type, F28L, and P29S Mutants 60

Figure 5.15 MMPBSA Energy Data for GTP- and GDP-bound Rac1 Q61R, Q61L and T17N Mutants 61

Figure 5.16 Binding free energy graph for Rac1. Binding free energy values (in kcal/mol) of Rac1 in complex with GTP or GDP, calculated from three replicate molecular dynamics simulations. Bars represent the mean values, and error bars denote standard deviations. Variants of Rac1, including wild type and mutants (P29S, F28L, Q61R, Q61L, T17N), are shown for both GTP and GDP bound forms 65

Figure 5.17 MM/GBSA plots of trajectories for docking structures of Rac1 and PAK1 across wild type and various mutations (F28L, P29S) 67

Figure 5.18 MM/GBSA plots of trajectories for docked Rac1 and PAK1 structures across various mutations (Q61L, Q61R, T17N) 68

Figure 5.19 Binding free energy graph for PAK1. Binding free energy values (in kcal/mol) of PAK1 in complex with Rac1-bound GTP or GDP, based on three replicate simulations. Mean binding energies are shown as bars, with standard deviations indicated by error bars. The impact of Rac1 mutations on the binding affinity of PAK1 is visualized across variants 72

Figure 5.20 MM/GBSA Values for Rac1-TIPE2 Complex different nucleotide states 73

List of Tables

Table 4.1 General Summary of Simulations ……..... 33

Table 5.1 Haddock Docking score of Wild type Rac1(GTP) with PAK1 36

Table 5.2 Haddock Docking score of Wild type Rac1(GDP) with PAK1 36

Table 5.3 Haddock Docking score of F28L Rac1(GTP) with PAK1 .....…37

Table 5.4 Haddock Docking score of F28L Rac1(GDP) with PAK1 37

Table 5.5 Haddock Docking score of P29S Rac1(GTP) with PAK1 38

Table 5.6 Haddock Docking score of P29S Rac1(GDP) with PAK1 38

Table 5.7 Haddock Docking score of Q61R Rac1(GTP) with PAK1 39

Table 5.8 Haddock Docking score of Q61R Rac1(GDP) with PAK1 39

Table 5.9 Haddock Docking score of Q61L Rac1(GTP) with PAK1 40

Table 5.10 Haddock Docking score of Q61L Rac1(GDP) with PAK1 40

Table 5.11 Haddock Docking score of T17N Rac1(GTP) with PAK1 41

Table 5.12 Haddock Docking score of T17N Rac1(GDP) with PAK1 41

Table 5.13 Rac1(GTP) vs TIPE2 Docking Poses. Energy values of the best pose in each set based on total score 43

Table 5.14 Rac1(GTP) vs TIPE2 Docking Poses. Energy values of the best pose in each set based on interface score 43

Table 5.15 Rac1(GDP) vs TIPE2 Docking Poses. Energy values of the best pose in each set based on total score 43

Table 5.16 Rac1(GDP) vs TIPE2 Docking Poses. Energy values of the best pose in each set based on interface score 43

Table 5.17 Translational, Rotational, and Vibrational Entropy Contributions and Total−TΔS Values (kcal/mol) for Protein-Ligand (GTP/GDP) Complexes. Values Obtained from Three Replicates for Each Complex. 63

Table 5.18 Entropy Contribution−TΔS Values of the Groups, with Wild Type Mean and Comparative Analysis 64

Table 5.19 Binding free energy values of GTP/GDP obtained from three replicate runs and their averages 65

Table 5.20 Translational, Rotational, and Vibrational Entropy Contributions and Total−TΔS Values (kcal/mol) for Protein-Protein (Rac1/PAK1) Complexes. Values Obtained from Three Replicates for Each Complex. 69

Table 5.21 Entropy Contribution−TΔS Values of the Groups (Rac1/PAK1), with Wild Type Mean and Comparative Analysis 70

Table 5.22 Binding free energy values and averages of PAK1 obtained from three replicate runs 72

Table 5.23 Translational, Rotational, and Vibrational Entropy Contributions and Total−TΔS Values (kcal/mol) for Rac1 TIPE2 Complexes. Values Obtained from Three Replicates for Each Complex. 74

Table 5.24 Binding energy data obtained from MD simulations of Rac1 (GTP and GDP) with TIPE2 74

ABBREVIATIONS

Å: Ångström

Amber: Assisted Model Building with Energy Refinement (simulation software)

AmberTools: Toolkit within AMBER for parameter generation and file processing

fs: femtosecond

μs: microsecond

GCP: Phosphomethylphosphonic Acid Guanylate Ester

GDP: Guanosine diphosphate

GTP: Guanosine triphosphate

HADDOCK: High Ambiguity Driven biomolecular DOCKing

HMR: Hydrogen Mass Repartitioning

K: Kelvin

kcal: kilocalorie

MD: Molecular dynamics

MM/GBSA: Molecular Mechanics, General Born Surface Area

Mg^{2+}: Magnesium ion with a 2+ charge

ns: nanosecond

PDB: Protein Data Bank

PAK: p21-activated kinase, Serine/threonine-protein kinase PAK

PAK1: p21-activated kinase 1, Serine/threonine-protein kinase PAK 1

PAK2: p21-activated kinase 2, Serine/threonine-protein kinase PAK 2

PAK3: p21-activated kinase 3, Serine/threonine-protein kinase PAK 3

PAK4: p21-activated kinase 4, Serine/threonine-protein kinase PAK 4

PAK5: p21-activated kinase 5, Serine/threonine-protein kinase PAK 5

PAK6: p21-activated kinase 6, Serine/threonine-protein kinase PAK 6

Rac1: Ras-related C3 botulinum toxin substrate 1

Rac3: Ras-related C3 botulinum toxin substrate 3

RMSD: Root Mean Square Deviation

RMSF: Root Mean Square Fluctuation

TIPE (TNFAIP8): Tumor necrosis factor alpha-induced protein 8

TIPE1 (TNFAIP8L1): Tumor necrosis factor alpha-induced protein 8-like protein 1

TIPE2 (TNFAIP8L2): Tumor necrosis factor alpha-induced protein 8-like protein 2

TIPE3 (TNFAIP8L3): Tumor necrosis factor alpha-induced protein 8-like protein 3

ΔG: Gibbs Free Energy Change

–TΔS: Temperature-scaled Entropy Change

ParmEd: AMBER-compatible parameter editing tool used in HMR

ff14SB: AMBER force field for proteins

TIP3P: Three-point water model used in MD simulations

PDBsum: Web server for PDB structural analysis

INTRODUCTION

Intracellular signal transduction and protein-protein interactions are critically important for the functioning of biological systems. Processes such as the maintenance of metabolic pathways, cellular responses to changing environmental conditions, and many others are regulated through these mechanisms. Disruptions or mutations affecting the proteins involved in these pathways can lead to loss of function, which may severely impact living systems. In fact, such impairments are commonly observed in diseases including cancer, metabolic disorders, immune system dysfunctions, and various other pathological conditions. In this context, we aimed to investigate the interactions among Rac1 (Ras-related C3 botulinum toxin substrate 1), TIPE2 (Tumor necrosis factor alpha-induced protein 8-like protein 2), and PAK1 (p21-activated kinase 1, Serine/threonine-protein kinase PAK1), with a particular focus on the impact of Rac1 mutations. To achieve this goal, we performed molecular dynamics (MD) simulations. MD simulations are especially effective in understanding how mutations affect protein structures, exploring the nature of protein-protein interactions, and analyzing binding energies. Following the simulations, trajectory data were analyzed using RMSD, RMSF, and similar techniques to gain insight into protein binding dynamics. By integrating our findings, we were able to identify potential interaction pockets between RAC1 and TIPE2 and further investigated how the RAC1–PAK1 interaction is modulated by mutations. Additionally, we conducted a detailed structural analysis of the impact of mutations on Rac1, with particular emphasis on the Switch I and Switch II regions, and their interactions with GDP/GTP molecules and Mg^{+2} ions. The obtained trajectories were further analyzed using methods such as RMSD and RMSF, and binding energy values between Rac1–TIPE2 and Rac1–PAK1 were calculated.

LITERATURE REVIEW

3.1 Rac1

The Ras superfamily is a large group with over 150 known members (according to the study by Rojas et al., 167 members were identified as of 2012) [Rojas et al., 2012]. Among these, 39 belong to the Ras family, 30 to Arf, 22 to Rho, 65 to Rab, and 1 to the Ran family. Also, The Ras superfamily is a group of proteins that play crucial roles in various intracellular processes, ranging from the regulation of cell motility and cell cycle control to vesicle transport, actin cytoskeleton organization, and even human tumorigenesis [Oxford and Theodorescu, 2003].

Additionally, this protein family is represented across nearly all major groups of organisms, from Homo sapiens to Mus musculus, Xenopus tropicalis, Ciona intestinalis, Drosophila melanogaster, Caenorhabditis elegans, Saccharomyces pombe, Arabidopsis thaliana, and Plasmodium falciparum [Rojas et al., 2012].

Although there are sequence-level variations among members, the domain responsible for GTP/GDP binding is highly conserved.

3.1.1 Rac1 Structural Properties

Structurally, Rac1 contains a G-domain that extends from G1 to G5, as well as a hypervariable region (HVR) [Ma et al., 2023].

Post-translational modifications occurring particularly at the highly variable C- terminal region enable membrane attachment. These modifications often include prenylation, palmitoylation, and geranylation.

In Ras family, the transition between GDP and GTP bounded states—commonly referred to as the" off" and" on" conformations—occurs through structural changes in regions called" Switch" domains, which surround the

nucleotide-binding pocket. These changes are crucial in determining whether the protein can bind to its effector.

Five characteristic motifs surrounding this region are identified, labeled from G123 (Figure 3.1). This region is functionally important for maintaining nucleotide to G5 (Figure 3.1 3.2):

- G1 / p-loop / Walker A motif: This motif typically contains a conserved sequence such as GxxxxGKS/T. In Rac1, this sequence is GDGAVGKT [Vetter, 2014] affinity. The lysine residue within the P-loop forms direct interactions with the beta and gamma phosphate oxygens of the nucleotide, while the ser- ine/threonine at the end of the loop interacts with magnesium [Vetter, 2014].
- G2 / Switch II: This region lies approximately between residues 26–37 in Rac1 and contains a highly conserved threonine residue (at position 35 in Rac1) [Vetter, 2014, Tanabe et al., 2000] (Figure3.1).
- G3 (Walker B motif): It contains the conserved DxxG sequence and is located in close proximity to the Switch II region. The aspartate residue (D) typically coordinates with magnesium. In Rac1, this region is represented by the sequence LWDTAGQE [Vetter, 2014, Tanabe et al., 2000] (Figure3.1).
- G4: 115 Located between residues 115–119, this motif includes the sequence N/TKxD. In Rac1, it is found as TKLDL [Vetter, 2014, Tanabe et al., 2000] (Fig-ure3.1).
- G5: This weakly conserved motif is typically composed of a SAK or SAL se-quence. In Rac1, it appears as ECSALT and is involved in interacting with the oxygen atoms of the guanine base [Vetter, 2014, Tanabe et al., 2000] (Figure 3.1).

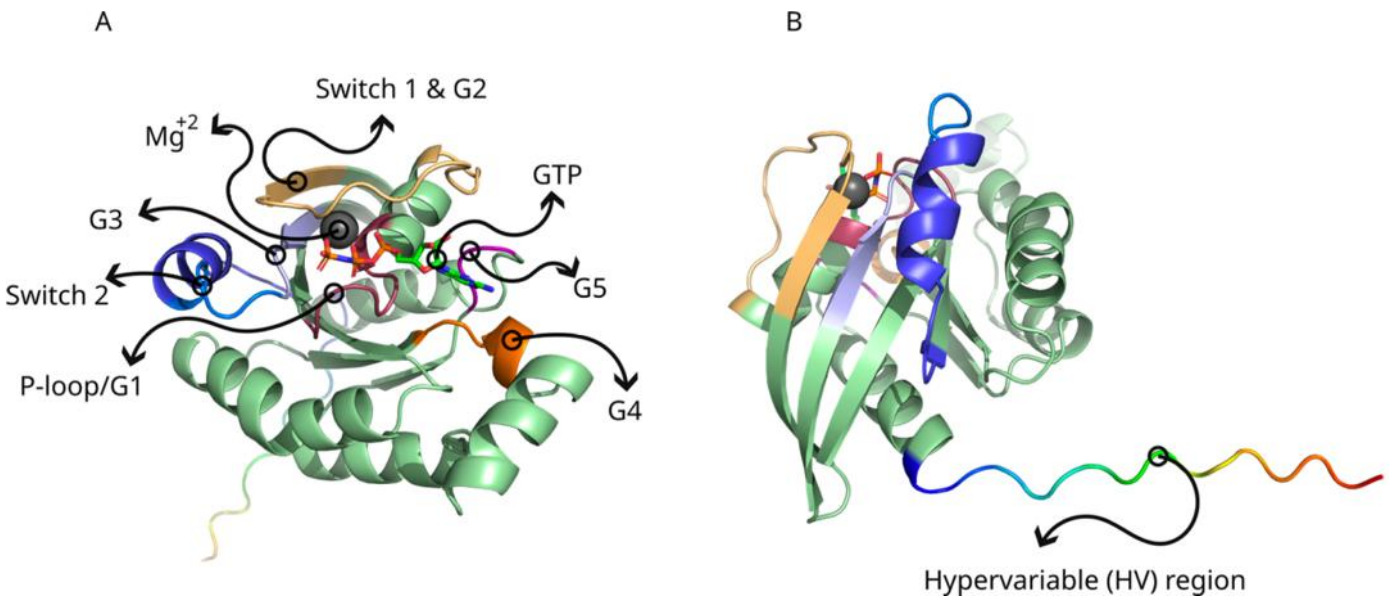


Figure 3.1: 3D Structure and Functional Domains of Rac1. A - Key motifs from G1 to G5 and the positions of GTP/GDP and Mg. B - Hypervariable region

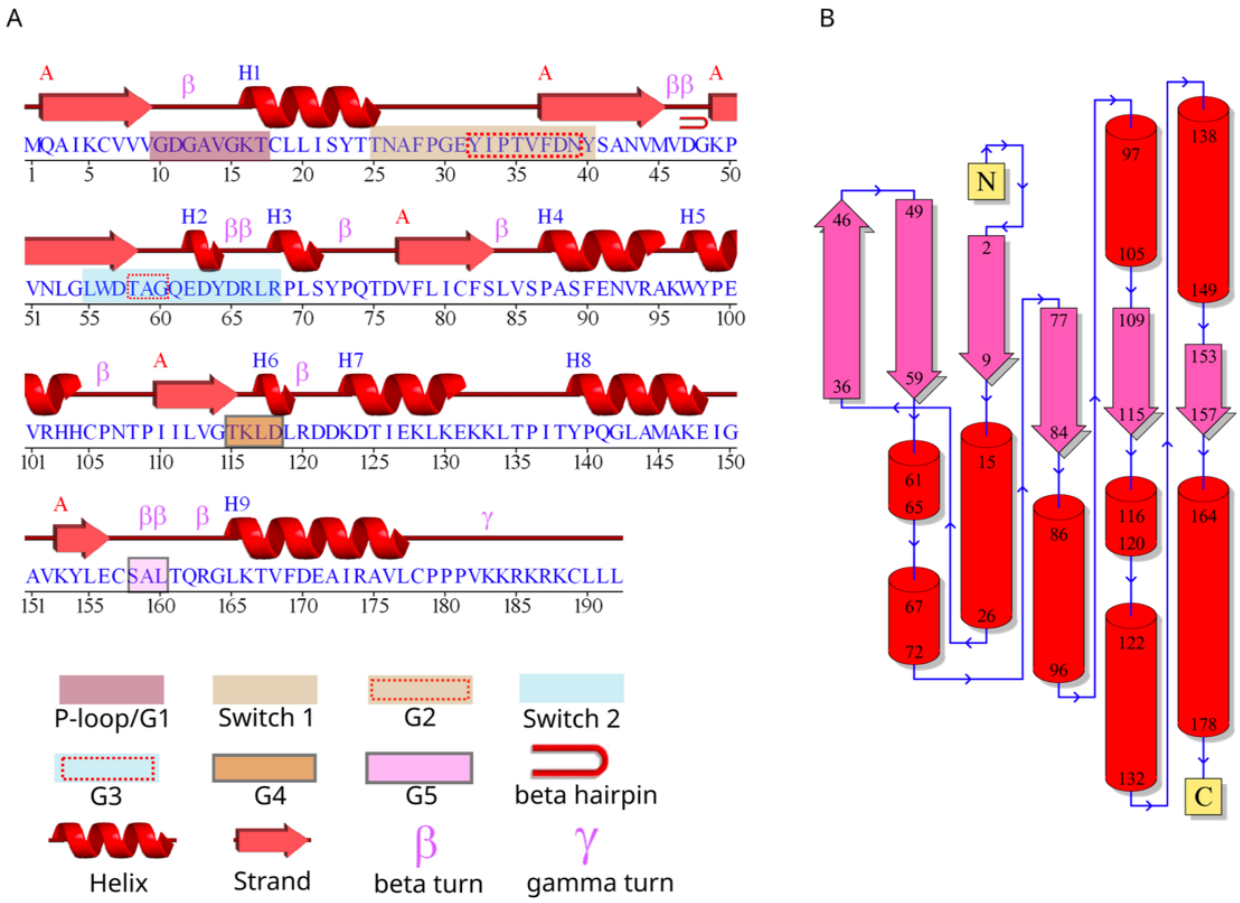


Figure 3.2: Topological and Secondary Structures of Rac1. (A) Secondary structures and the positions of key motifs on these structures. (B) Topological representation showing the connections and order of alpha helices and beta sheets.

3.1.2 Rac1 Mutations

In the scope of this thesis, several mutations in the Rac1 protein, many of which are implicated in diseases such as cancer, have been analyzed structurally and dynamically at the molecular level. The mutations investigated and their associated functional outcomes are summarized below:

- P29S: In contrast to typical oncogenic mutations observed in Ras proteins, this mutant retains its GTPase activity (Figure 1). Although it can

hydrolyze GTP, it exhibits a significantly accelerated GDP-to-GTP exchange rate—referred to as spontaneously active, resulting in its constitutive activation [4GZM].

- F28L: This mutant is particularly known for its constitutive activity, primarily due to altered flexibility in the switch regions, which facilitates rapid GTP/GDP exchange—referred to as 'fast cycling' [Kumar et al., 2013]. A comparable effect has also been observed in HRAS mutants at the corresponding residue position (Figure 1).
- Q61L: A mutation that keeps Rac1 in an active state by blocking the conversion of GTP to GDP, thereby reducing its inherent GTPase function (Figure 1). As a result, Rac1 remains locked in a persistently 'on' (active) conformation. However, this mutation leads to a functional impairment, as the overly stabilized active state fails to recapitulate the dynamic nature of normal Rac1 signaling. The Q61L variant represents a constitutively active yet functionally compromised form, capable of aberrantly modulating intracellular signaling pathways without properly mimicking the physiological signaling cycle [Acuner et al., 2021].
- Q61R: Q61R is a gain-of-function mutation for Rac1, similar in effect to the activating P29S mutation (Figure 1). This substitution promotes the activation of Rac1 by facilitating its binding to GTP. While it does not completely abolish GTPase activity, it enhances Rac1's intrinsic activation potential, allowing it to adopt the active conformation more readily.

As a result, the probability of Rac1 interacting with downstream effector proteins, such as PAK1, increases. Moreover, Q61R induces structural rearrangements by modulating the allosteric communication between key functional regions of Rac1—namely, the Switch I & II regions, the p-loop, and the insert region—thereby altering its conformational dynamics and signaling behavior [Acuner et al., 2021].

- T17N: The T17N mutation is classified as a dominant-negative variant of Rac1. It is located within the phosphate-binding loop (p-loop), a region critical for coordinating both magnesium ions and nucleotide (GDP/GTP) binding (Figure 1).

This mutation severely disrupts Rac1 activation by impairing its ability to bind GTP and weakening its affinity even for GDP, effectively locking the protein in an inactive or nucleotide-free state.

Interestingly, the T17N variant exhibits a stronger binding affinity for guanine nucleotide exchange factors (GEFs) compared to the wild-type. By sequestering GEFs, it competitively inhibits the activation of other Rac1 molecules, thereby exerting a dominant-negative effect on Rac1 signaling.

Functionally, this mutation drastically diminishes the ability of Rac1 to interact with downstream effectors such as PAK1. The loss of nucleotide binding not only reduces Rac1–PAK1 affinity but also leads to increased flexibility around key residues like T17 and T35, disrupting the formation of stabilizing hydrogen bonds with GTP. These structural changes further compromise Rac1's signaling capabilities [Acuner et al., 2021].

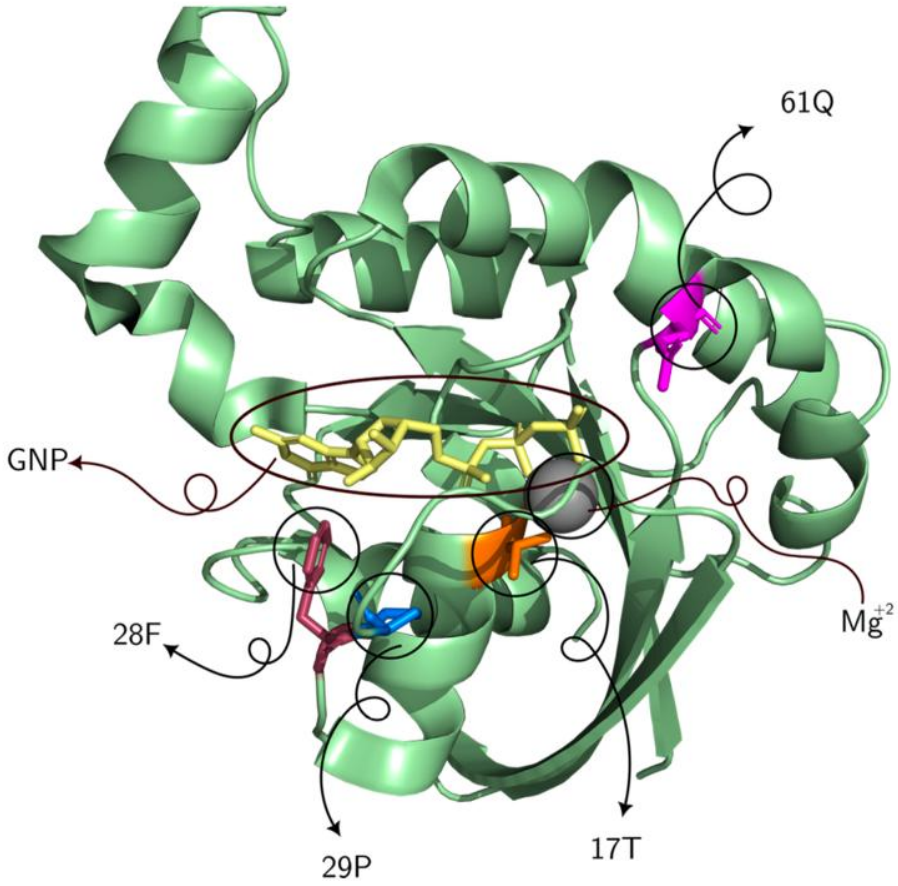


Figure 3.3: Mutations and Locations in the Rac1

3.2 TIPE2

TIPE2 (TNFAIP8-like protein 2) is one of four members of the tumor necrosis factor alpha-induced protein 8-like (TIPE) superfamily, which includes TNFAIP8 (TIPE), TIPE1 (TNFAIP8L1), TIPE2 (TNFAIP8L2), and TIPE3 (TNFAIP8L3). They possess structurally conserved domains and play roles in modulating immune function, programmed cell death, and cancer development [Bordoloi et al., 2018].

Members of the TIPE family are known to modulate inflammatory pathways and maintain immune homeostasis. Notably, TIPE2 functions as a negative regulator of both innate and adaptive immunity by suppressing T cell receptor (TCR) and Toll-like receptor (TLR) signaling, thereby preventing

excessive inflammation and autoimmunity. Dysregulation of TIPE2 has been associated with autoimmune diseases and various malignancies, highlighting its therapeutic potential [Sun et al., 2008; Zhang et al., 2009].

Chronic inflammation is widely recognized as a key contributor to tumor development. Inflammatory stimuli activate immune cells, leading to the secretion of tumor necrosis factor-alpha (TNF-α), which binds to TNFR1 and TNFR2 receptors. This interaction activates the NF-κB signaling pathway, subsequently enhancing the transcription of TIPE family genes, including TIPE2 [Gu et al., 2020].

Interestingly, some TIPE proteins, particularly TIPE and TIPE3, have been implicated in promoting drug resistance in cancer, whereas TIPE1 and TIPE2 have demonstrated anti-tumorigenic properties by promoting apoptosis. Among them, TIPE2 plays a central role in regulating inflammation and preserving immune system balance, which further underlines its importance in cancer biology and immune regulation [Gu et al., 2020; Sun et al., 2008].

3.2.1 TIPE2 Topology

Structurally, the TIPE protein exhibits a conserved helical topology composed of six α-helices, designated as α1 – α6 (Figure 2A) [Zhang et al., 2009].

These helices are arranged in a compact structure that creates a central hydrophobic cavity—one of the key structural features of the TIPE family.

Closer examination of the cavity reveals a markedly hydrophobic inner surface, in contrast to the relatively polar outer surface of the protein (Figure 2B). The volume of this cavity has been calculated to be approximately 1217 $Å^3$, providing adequate space for the binding of small ligands or hydrophobic signaling molecules (Figure 2C) [Zhang et al., 2009].

This structural feature supports the proposed role of TIPE proteins in lipid binding and signaling regulation, similar to TIPE3, which functions as a transfer protein for lipid second messengers [Cao et al., 2013; Xu et al., 2024].

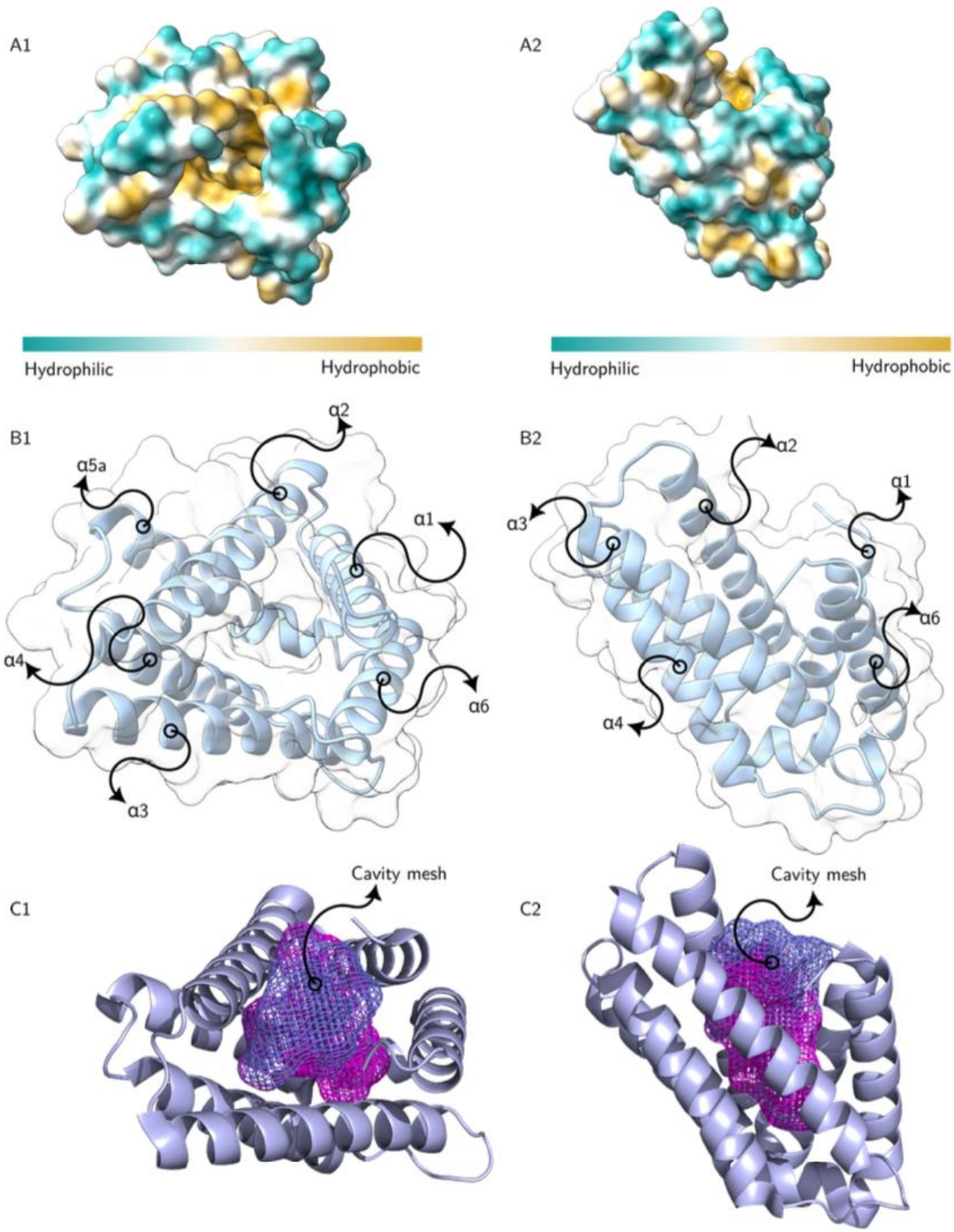


Figure 3.4: 3D Structure and Topology of TIPE2. (A) Hydrophilic and hydrophobic regions. (B) Helices and their positions. (C) Cavity between helices.

3.3 PAK1

p21-activated kinases (PAKs) constitute a family of serine/threonine kinases that serve as critical effectors for the Rho family small GTPases, particularly Cdc42 and Rac1. These kinases are pivotal in regulating various cellular processes, including cytoskeletal dynamics, cell motility, survival, and proliferation. The PAK family is subdivided into two groups based on structural and functional characteristics: Group I (PAK1, PAK2, PAK3) and Group II (PAK4, PAK5, PAK6) [Pak Genel; Ha et al., 2015; Rane et al., 2014; Ching et al., 2003].

Group I PAKs, including PAK1, are characterized by an N-terminal regulatory domain that encompasses a p21-binding domain (PBD) overlapping with an autoinhibitory domain (AID), and a C-terminal kinase domain. In their inactive state, Group I PAKs form homodimers wherein the AID of one monomer inhibits the kinase domain of the other. Activation occurs upon binding of active GTP-bound Cdc42 or Rac1 to the PBD, leading to conformational changes that release the AID-mediated inhibition and activate the kinase domain [Molli et al., 2009].

PAK1, in particular, plays a significant role in various cellular functions such as cytoskeletal reorganization, gene expression, and cell cycle progression. Its activity is tightly regulated through the aforementioned autoinhibitory mechanism, ensuring precise control over its kinase function in response to upstream signals [Kumar et al., 2006].

In the case of PAK1, which is the focus of this thesis, the GTPase-binding region is generally located between residues 75 and 105, the autoinhibitory domain (AID) spans residues 70 to 149, and the kinase domain is found between residues 272 and 523 [Pak Genel].

For activation, PAK1 normally exists as an asymmetrical parallel dimer. At this stage, one of the dimeric units adopts an active conformation while the other remains inactive [Wang et al., 2011; Pak Genel]. The binding interface and relative positions of the dimers are illustrated in Figure 3.

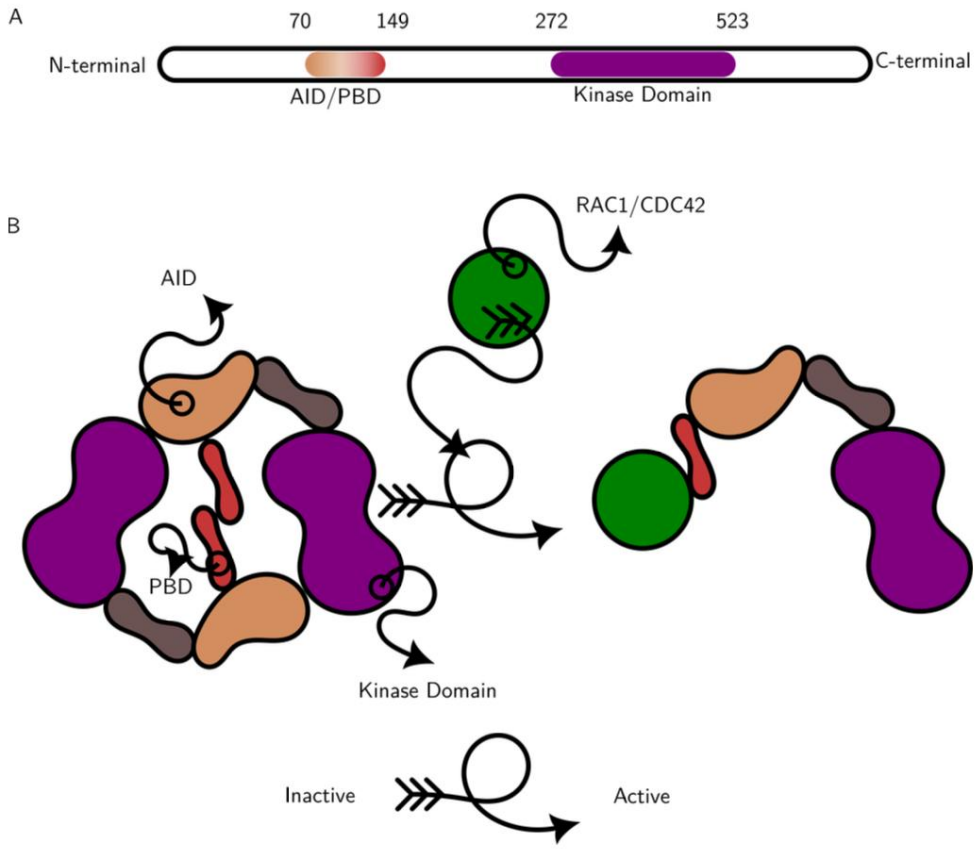


Figure 3.5: Domains and Activation Cycle of PAK1. A - Key motifs and their positions in PAK1, B - Representation of the inactive and active cycle of PAK1

PAK1 and Cancer

PAK1 is frequently overexpressed in almost all cancer types, and its role becomes more pronounced in malignant tumors [Pak Genel].

Specifically, PAK1 has been shown to play an active role in various stages of cancer development. During the initiation phase, it contributes to genomic instability, oncogene activation, and inflammation. In the tumor growth phase, PAK1 is involved in hypoxia response, cellular proliferation, mitosis, resistance to apoptosis, and angiogenesis.

In later stages such as metastasis and cancer cell survival, PAK1 supports invasion, continued resistance to apoptosis, cell migration, chronic inflammation, angiogenesis, and immunosuppression [Pak Genel].

Moreover, PAK1 is implicated in cancer cell metabolic reprogramming and immune evasion. It also plays a significant role in the development of drug resistance, making it a potential therapeutic target in multiple cancer types. Studies have also highlighted its crosstalk with key oncogenic pathways such as MAPK, PI3K/AKT, and NF-κB, further amplifying its impact on tumor progression and resistance to therapy [Liu et al., 2021].

3.4 Rac1, PAK1 and TIPE2 Interactions

The interactions between Rac1–PAK1 and Rac1–TIPE2 have been the subject of studies employing of experimental and computational methods. Below is a summary of the current findings and key gaps in the literature regarding these protein–protein interactions.

In particular, crystallography-based structural studies conducted by Sells et al. demonstrated that the active GTP-bound form of Rac1 binds to the p21-binding domain (PBD) of PAK1, thereby releasing the autoinhibitory segment on PAK1 and initiating its kinase activity [Sells et al., 1997]. This activation directly triggers crucial cellular processes such as actin cytoskeleton reorganization and cell migration. Moreover, the Rac1–PAK1 interaction is not limited to conformational activation; it also facilitates PAK1 autophosphorylation and the activation of downstream signaling pathways, including MAPK and LIMK, thus

opening the door to a broader signaling network. Consequently, this interaction is considered a key mechanism influencing cellular behavior in various pathological conditions, including cancer, inflammation, and neurodegenerative diseases [Bokoch, 2003].

However, several critical gaps in the literature remain. Primarily, the impact of mutations on the Rac1–PAK1 interaction has not been thoroughly investigated. In particular, structural responses of this complex to mutations have not been analyzed in detail using molecular dynamics (MD) simulations or similar approaches. Furthermore, the extent to which GTP and GDP contribute to this interaction through the Switch I and Switch II regions remains unclear.

When examining the interaction between Rac1 and TIPE2, numerous studies can be found in the literature, yet the structural basis of this interaction has not been fully elucidated. For instance, in a study conducted on non-small lisuppress Rac1 activity; however, the underlying structural mechanism of this suppression was not explained [Li et al., 2016]. Similarly, another study reported that TIPE2 expression was significantly reduced in hepatocellular carcinoma (HCC) tissues, while forced expression of TIPE2 inhibited tumor growth, invasion, and metastasis—effects that were attributed to the downregulation of Rac1 activity. Nevertheless, as in the previous case, the molecular-level mechanism of this interaction was not thoroughly investigated [Cao et al., 2013b].

Building on these findings, bioinformatics analyses further highlight the central role of Rac1 as a molecular hub. Various interaction databases, such as STRING, provide evidence that Rac1 interacts or potentially interacts with both TIPE2 and PAK1, supporting the experimental observations presented above. Figure 1 illustrates the interaction map of Rac1, PAK1, and TIPE2 obtained from the STRING database [Szklarczyk et al., 2023].

When a Rac1-centered query is performed in the STRING database, as shown in Figure 1A, Rac1 appears to potentially interact with numerous proteins,

including PAK1. Similarly, in Figure 1B, a TIPE2-based search reveals that TIPE2 also interacts with various proteins, including Rac1. Figure 1C presents the results of a PAK1-centered query, which largely overlaps with the interaction profile observed for Rac1.

Taken together, the comparison of these networks highlights the functional and possibly regulatory interplay between RAC1, TIPE2 and PAK1. Notably, RAC1 appears as a central node linking both TIPE2 and PAK1 networks, suggesting it may serve as a molecular hub in pathways involving cytoskeletal remodeling, immune regulation, and cancer progression.

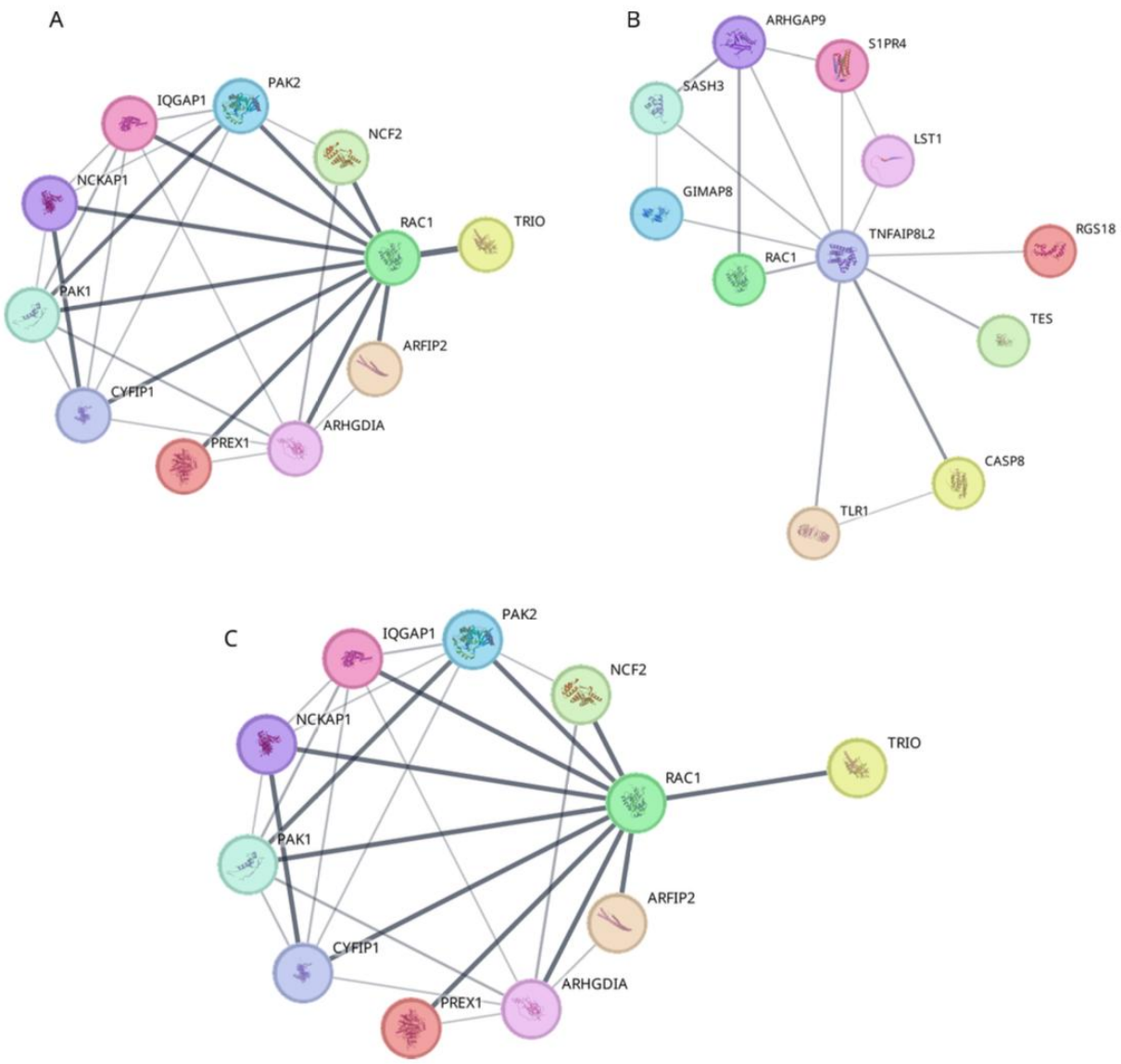


Figure 3.6: Potential Interaction Partners of Rac1, TIPE2, and PAK1. The thickness of the lines on the graph represents the reliability of protein-protein interactions and the abundance and diversity of information about these interactions. A – For Rac1, B - For TIPE2, C - For PAK1

In this thesis, the overall workflow can be broadly divided into four main stages: structure preparation, molecular docking, molecular dynamics (MD) simulations, and post-simulation analyses. The flowchart representing these steps is presented in Figure 4.1. Detailed descriptions of each step are provided below in a step-by-step manner.

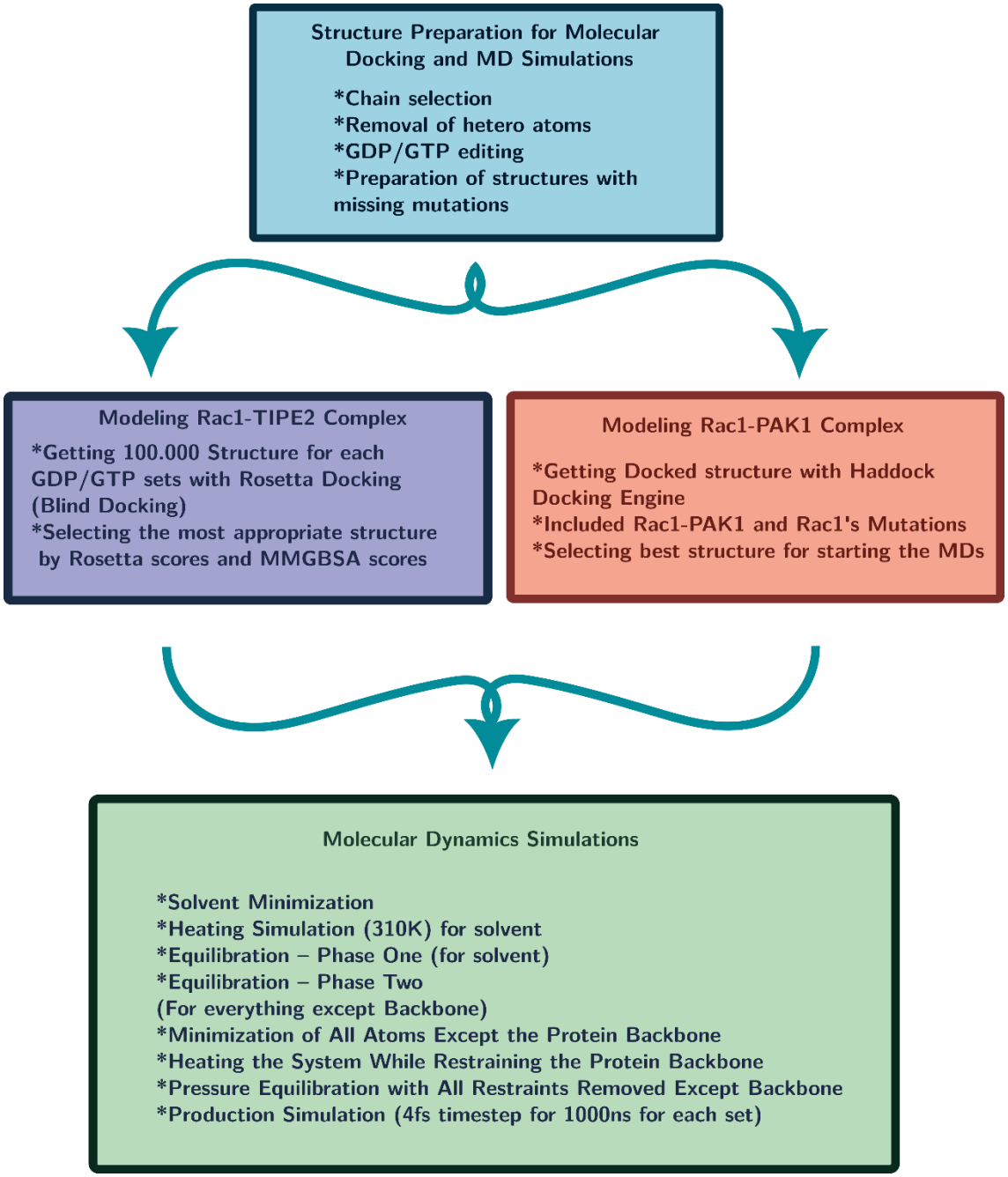


Figure 4.1: Overall workflow of the study

4.1 Structure Preparation for Molecular Docking and MD Simulations

Structure preparation steps for docking and molecular dynamics simulations (such as removal of undesired heteroatoms, chain selection, etc.) were carried out manually using PyMOL or by editing the structure files directly [Schr¨odinger, LLC, 2015b].

4.2 Modeling PAK1 and TIPE2 complexes with Rac1

4.2.1 Modeling Rac1-TIPE2 Complex

To prevent potential sequence inconsistencies in the numbering of atoms within the PDB files from affecting the docking processes, atom and residue numbers were renumbered sequentially. For residues that could potentially adopt multiple positions, a single conformation was selected.

Improvements on the structure were made using the Relax module within Rosetta, which minimizes the structure's energy by allowing limited movements in the protein backbone and changes in side-chain conformations (rotamers). This module was used to generate three distinct structures. To increase structural diversity, each of these three structures was subjected to 10 repeated relaxation cycles [Tyka et al., 2011, Khatib et al., 2011, Maguire et al., 2021, Conway et al., 2014, Conway et al., 2014 Niv´on et al., 2013].

The following steps outline the use of the Rosetta relax protocol on the input structure complex A.pdb:

- Use the input PDB file complex A.pdb as the starting structure.
- Apply the relax protocol with 10 default minimization cycles for each model.
- Generate a total of 3 independent models (nstruct = 3).
- Save the resulting PDB structures to the folder ./tutorial output.
- Save the score (energy) files to the folder ./expected output.
- Append the suffix N relax cycles to all output files for easy identification.

The actual command used:

relax.static.linuxgccrelease \

-s kompleks_renumbere_A.pdb \

-out:suffix _N_relax_cycles \

-relax:default_repeats 10 \

-nstruct 3 \

-out:path:pdb ./structure_output \

-out:path:score ./expected_output

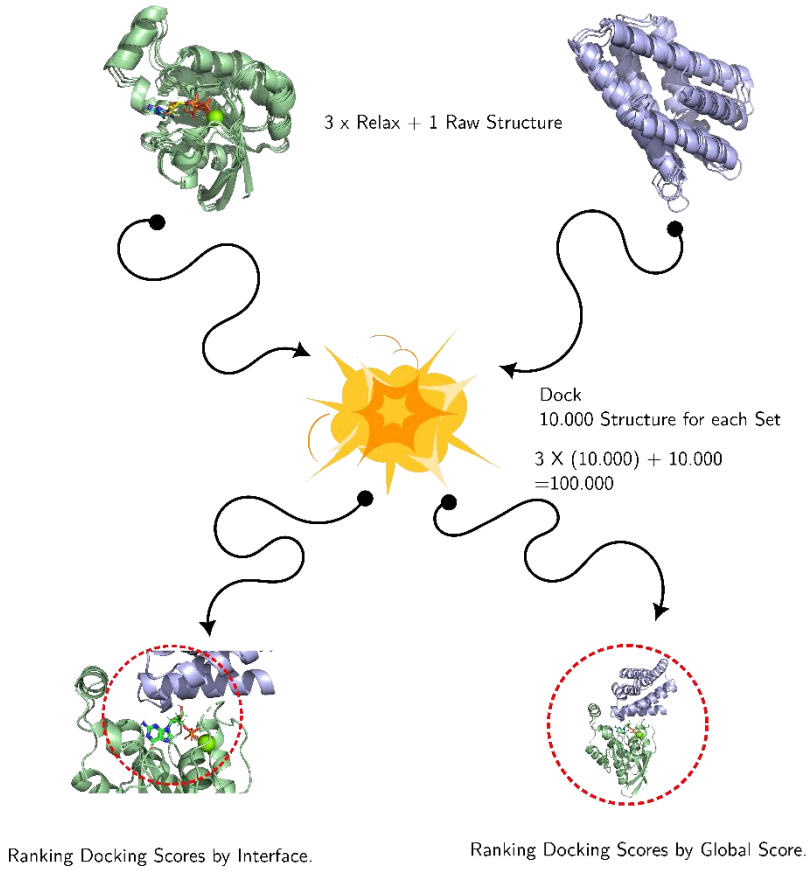


Figure 4.2: Workflow of the Docking and Post-Docking Scoring Processes for Rac1 and TIPE2

Global docking was performed for each docking set using Rosetta Dock [Gray et al., 2003, Marze et al., 2018, Wang et al., 2005, Wang et al., 2007,

Chaudhury and Gray, 2008, Chaudhury and Gray, 2008, Chaudhury et al., 2011, Marze et al., 2017].

The docking parameters we used are as follows

- The command -dock pert 3 8 applies a random translation up to 3 ˚A and rotation up to 8◦ to the structure at the start.
- The commands -randomize1 and -randomize2 randomly position Chain A and Chain B, respectively, in space.
- The command -spin spins the second docking partner around axes from the center of mass of the first partner to the second partner (global).
- The commands -ex1 and -ex2aro add extra side-chain rotamers; -ex2aro adds extra chi1 sub-rotamers specifically for aromatic residues, while -ex1 adds extra chi1 sub-rotamers for all residues that pass.

After docking, the MM-GBSA and VD-MM-GBSA calculation tools available on the HawkDock server were utilized to evaluate the scores of the complexes [Weng et al., 2019]

To identify the best docking structure from the obtained results, threshold values were established for the Total energy and Interface energy from Rosetta docking, as well as the MM/GBSA energies from HawkDock for a single structure.

The determined threshold values are as follows:

- Threshold for HawkDock MMGBSA energy: −20 or lower
- Threshold for HawkDock (VD)MMGBSA energy: −90 or lower
- Threshold for Rosetta Docking Interface score: −30 or lower
- Threshold for Rosetta Docking Total score: −1000 or lower

4.2.2 Modeling Rac1-PAK1 Complex

For the docking process, the structures of the Rac1 and PAK1 proteins were saved individually in PDB file format, and all heteroatoms—except for

GDP/GTP and Mg2+ ions—were removed. PyMOL 3.1 software was used for these preparation steps [Schrödinger, LLC, 2015b].

The structures used and the preparation steps are as follows:

For wild-type Rac1 (GDP-bound), 5N6O was used [Ferrandez et al., 2017]

- For wild-type Rac1 (GTP-bound), 3TH5 was used [Krauthammer et al., 2012]
- For p29s mutant RAC1 (GTP-bound), 3SBD was used (GNP was replaced with GTP using PyMOL) [Krauthammer et al., 2012]
- For p29s mutant RAC1 (GDP-bound), 3SBD was used (GNP was replaced with GDP using PyMOL) [Krauthammer et al., 2012]
- For F28L mutant Rac1 (GDP-bound), 4GZM was used (GSP was replaced with GDP using PyMOL) [Davis et al., 2013]
- For F28L mutant Rac1 (GTP-bound), 4GZM was used (GSP was replaced with GTP using PyMOL) [Davis et al., 2013]
- For Q61L mutant Rac1 (GTP-bound), 3TH5 was used; GNP was replaced with GTP using PyMOL, and the Q61L mutation was introduced in PyMOL accordingly [Krauthammer et al., 2012]
- For Q61L mutant Rac1 (GDP-bound), 5N6O was used and the Q61L mutation was introduced using PyMOL [Ferrandez et al., 2017]
- For F28L mutant Rac1 (GTP-bound), 4GZM was used (GSP was replaced with GTP using PyMOL) [Davis et al., 2013]
- For Q61L mutant Rac1 (GTP-bound), 3TH5 was used; GNP was replaced with GTP using PyMOL, and the Q61L mutation was introduced in PyMOL accordingly [Krauthammer et al., 2012]
- For T17N mutant Rac1 (GDP-bound), 5N6O was used and the T17N mutation was introduced using PyMOL [Ferrandez et al., 2017]
- For T17N mutant Rac1 (GTP-bound), 3TH5 was used and the T17N mutation was introduced using PyMOL [Krauthammer et al., 2012]

Interaction Site Identification

To identify the interface residues, the crystal structure of PAK1 - Rac3 complex (PDB ID: 2QME), Rac1 isoform and another member of the Rho GTPase family, was used [Ugochukwu et al., 2007] (Figure 4.3).

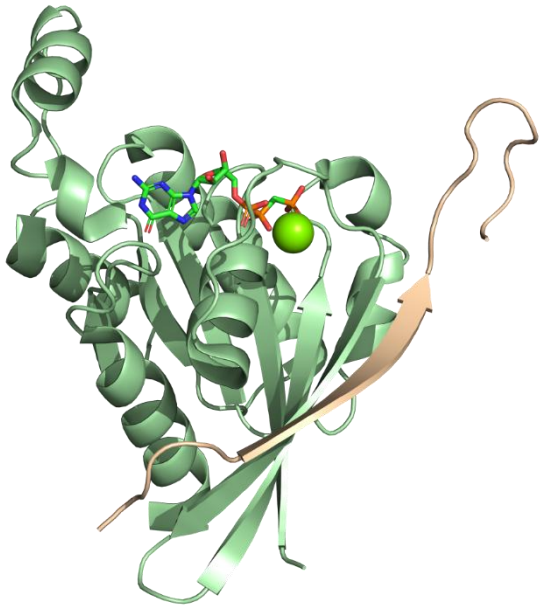


Figure 4.3: PAK1 - Rac3 Crystal Structure (PDB ID: 2QME) In the figure, the molecule GCP is visualized in stick representation, the green sphere represents magnesium, the wheat-colored region corresponds to the PAK1 protein, and the light green region indicates the RAC3 protein

The 2QME structure was analyzed via PDBsum to determine the interface residues of Rac3 and PAK1 [Laskowski, 2007]. Similarly, PDBsum was also utilized to visualize the interface between Rac1 and TIPE2.

Molecular Docking

For the docking of PAK1 and Rac1 proteins, the HADDOCK 2.4 web server was used [Honorato et al., 2024, Honorato et al., 2021]. At this stage, the interface regions identified in the previous step for Rac3 and PAK1 were used as active residues in HADDOCK 2.4 to guide the binding interface, and the docking process was performed using default parameters.

To ensure the docking process proceeded smoothly and that heteroatoms such as GDP/GTP and Magnesium were handled without errors, atom and residue naming was adjusted according to HADDOCK's internal library requirements.

After docking, the resulting docking poses were clustered by the HADDOCK 2.4 server. Among these clusters, a representative model for further analysis was chosen based on multiple criteria, including Docking score, cluster size, restraint violation energy, and RMSDs. From the selected cluster, the structure with the most favorable energy value was chosen to proceed with the molecular dynamics analyses.

4.3 Molecular Dynamics Simulations

4.3.1 Force Field and Parameters

Throughout the simulations, the ff14SB force field was used for proteins, along with the TIP3P water model, which is the most compatible water model for this force field [Maier et al., 2015].

For the Mg2+ ions, the frcmod.ions234lm 1264 tip3p library, which includes 12-6-4 Lennard-Jones parameters for divalent ions developed by Li and Merz, was employed [Li and Merz Jr, 2014, Li and Merz Jr, 2014].

The parameters for GDP and GTP were taken from the files developed by Carlson et al. [Meagher et al., 2003].

4.3.2 Hydrogen Mass Repartitioning - HMR

In molecular dynamics simulations, the time step plays an important role, especially for computational cost. Since hydrogen atoms vibrate at high frequencies, it is common to use a time step of 2 femtoseconds (fs). If we try to make larger time steps, we cannot properly run our simulations. To overcome

this limitation and run simulations more efficiently, the Hydrogen Mass Repartitioning (HMR) technique can be used. HMR can be summarized as transferring mass from heavy atoms bonded to hydrogens onto the hydrogen atoms themselves, thereby both conserving the total mass and increasing the mass of the hydrogens.

In this approach, a small amount of mass is shifted from the heavy atoms to the hydrogens they are bonded to. This lowers the vibrational frequency of hydrogen atoms, making it possible to use a longer time step without compromising stability. In our simulations, this technique allowed us to increase the time step from 2 fs to 4 fs, which helped to reduce the computational cost.

The mass redistribution was carried out using the ParmEd tool, which is part of the Amber software suite. Using the hmassrepartition command, hydrogen atoms were reassigned a mass of 3.024 Daltons (figure 4.4) [Hopkins et al., 2015]. The masses of nearby atoms were scaled accordingly to preserve the overall mass of the system. The process included the following steps:

The HMR procedure was carried out using the following commands:

parmed paremetre_dosyası.parm7

hmassrepartition

outparm ¸cıktı.parm7

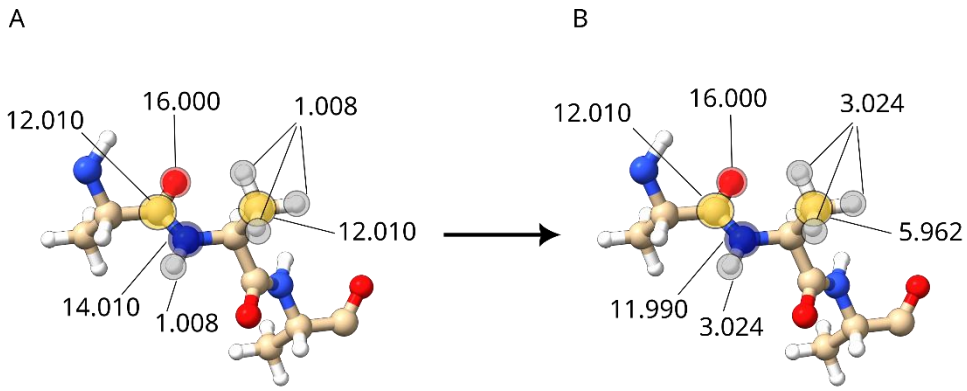


Figure 4.4: Hydrogen Mass Repartitioning (HMR) Process. The colors and their corresponding atoms are as follows: yellow-carbon, white-hydrogen, gray-repartitioned hydrogen, blue-nitrogen, and red-oxygen

4.3.3 Steps of Molecular Dynamics Simulation Runs

Molecular dynamics (MD) simulations of the docked poses were performed using the Amber software suite (AmberTools + pmemd.cuda). The simulation followed an 8-step protocol designed to adapt the system to the simulation environment. These steps include minimization, heating, pressure and temperature equilibration, and production phases. Each step was prepared with specific parameters to ensure system stability and to mimic experimental conditions as closely as possible.

1. Solvent Minimization

In the first step, only the energy of the solvent part of the system—water molecules and ions—was minimized. This helped removal of possible bad contacts, steric clashes, or strain that may have resulted from the initial structure, resulting in a more stable and optimal configuration. The rest of the system, including the protein, GTP/GDP, and the Magnesium ion, was restrained using restraintmask=':1-XXX' and restraint wt=500.

Energy minimization was carried out for a total of 1000 steps. The first 500 steps used the steepest descent algorithm (ncyc=500), and the remaining 500 steps used the conjugate gradient algorithm.

- The steepest descent algorithm moves the system quickly away from high-energy regions by following the energy gradient. It is especially useful in the early stages to resolve severe clashes.

- The conjugate gradient algorithm, on the other hand, is more refined and efficient, giving better results in lower-energy regions.

The parameters used at this stage were: maxcyc=1000, ncyc=500, and cut=10.0 (which defines a 10 ˚A cutoff for short-range interactions).

2. Heating Simulation

After minimization, the temperature of the system was gradually increased to prepare it for molecular dynamics conditions. Using a Langevin thermostat, the temperature was raised gradually from 100 K to 310 K.

During this heating phase, the protein, GTP/GDP, and Mg2+ ion were restrained using the restraintmask, allowing only the water molecules and ions

to move freely. This helps prevent any structural distortions during heating. The heating lasted a total of 50 ps, with a time step of 2 fs.

3. Equilibration – Phase One

In this step, the system was adjusted to constant pressure conditions to allow the volume to stabilize. While keeping the temperature fixed at 310 K, the system was equilibrated at 1 atm pressure. A Monte Carlo barostat (barostat=2) was used to control the pressure.

During this phase, the protein and ligand were restrained, while water molecules and ions were free to move. This gradual approach helps the system adapt to more realistic simulation conditions.

4. Equilibration – Phase Two

In the first pressure equilibration step, the system generally approached equilibrium. In this phase, the goal was to allow the system to become more flexible. To achieve this, the positional restraints on the protein backbone (restraintmask='@C,CA,N,O') were significantly reduced, and the force constant was lowered to 10 kcal/mol·°A2. This value permits atomic movement while preventing excessive deviations.

The simulation lasted 50 ps and was carried out at 310 K and 1 atm pressure using a Monte Carlo barostat (barostat=2) and a Langevin thermostat (ntt=3). During this step, the solvent and ions were completely free to move. The structure was expected to gradually adopt a more natural conformation.

5. Minimization of All Atoms Except the Protein Backbone

At this stage, the system underwent energy minimization again; however, this time only the protein backbone was restrained (restraintmask='@C,CA,N,O', restraint wt=500). Side chains, solvent molecules, and ligands were allowed to move freely. This approach helps to relieve any local stresses that might have developed during the heating and equilibration phases, especially to allow a more natural positioning of water and ions.

The minimization was performed for 1500 steps: the first 750 using the steepest descent algorithm, and the remaining 750 using the conjugate gradient

method. During this process, a 10 ˚A cutoff distance (cut=10.0) was applied for short-range electrostatic interactions to adequately account for these effects.

6. Heating the System While Restraining the Protein Backbone

At this stage, all atoms in the system—especially water molecules, ions, and protein side chains—were allowed to move freely under near-physiological temperature conditions. However, the protein backbone was kept restrained to maintain the overall structural integrity. This setup ensures that the surrounding molecules can equilibrate without disturbing the main framework of the protein.

The temperature was held constant at 310 K using the Langevin thermostat (ntt=3, gamma ln=1.0), and the simulation was run for 50 ps. This step is crucial for allowing the solvent molecules to settle into equilibrium and to create a more realistic environment.

7. Pressure Equilibration with All Restraints Removed Except Backbone

At this stage, all restraints were lifted except for a mild restraint on the protein backbone (restraint wt=5.0). This value provides gentle guidance to the backbone while largely allowing the atoms to move freely. As a result, the system becomes almost fully unconstrained.

The pressure was maintained at 1 atm and the temperature at 310 K, with the simulation running for 100 ps. Under these NPT conditions, both volume and temperature reached equilibrium, preparing the system for the production run. During this phase, parameters such as total potential energy, temperature, and volume are expected to stabilize.

8. Production Simulation

After the system was fully equilibrated, the production simulation was started. In this stage, a time step of 4 fs was used to run three replicate simulations for each system lasting 1000 ns.

4.3.4 Analyses Performed After MD Simulations

RMSD

The Root Mean Square Deviation (RMSD) is a measure of the average distance between the atoms (usually backbone atoms) of superimposed protein structures over time in a molecular dynamics simulation. In this thesis, RMSD was used to evaluate the overall structural stability of the protein and to compare the conformational changes induced by mutations.

The Root Mean Square Deviation (RMSD) calculations were performed using the cpptraj tool included in the Amber software suite [Roe and Cheatham III, 2013]. Throughout the entire simulation, structures were compared to the reference structure (the first frame) to determine their root mean square deviation (RMSD).

The calculation was based on the protein and the protein-bound GDP/GTP and Mg ions, focusing on these components for a more accurate evaluation. RMSD is defined by the following formula (Equation 4.1):

$$RMSD = \sqrt{\frac{\sum_{i=0}^{N} [m_i \cdot (X_i - Y_i)^2]}{M}} \tag{4.1}$$

Here, - N represents the total number of atoms, - mi is the mass of atom i, -Xi are the coordinates of atom i in the target structure, - Yi are the coordinates of atom i in the reference structure, - M denotes the total mass.

RMSF

The Root Mean Square Fluctuation (RMSF) is a measure of the average deviation of an atom or residue from its mean position over the course of a molecular dynamics simulation. In this thesis, it was used to analyze the effects of mutations on the protein, particularly in terms of flexibility. The Root Mean Square Fluctuation (RMSF) calculations were performed using the cpptraj tool included in the Amber software suite [Roe and Cheatham III, 2013].

The Root Mean Square Fluctutation (RMSF) calculations were performed using the equations given below (Equation 4.2):

$$\mathrm{RMSF}_i = \sqrt{\langle (x_i - \langle x_i \rangle)^2 \rangle} \tag{4.2}$$

Here, xi represents the position of the atom over time, and, ⟨xi⟩ is the average position over the entire time interval.

Subsequently, using the values obtained here, a mass-weighted relative RMSF value was calculated. The atom's mass was included in the calculations according to the equation below (Equation 4.3).

$$\langle \mathrm{Fluct} \rangle = \frac{\sum \mathrm{AtomFluct}_i \times \mathrm{Mass}_i}{\sum \mathrm{Mass}_i} \tag{4.3}$$

This provides the average fluctuations weighted by the mass of each atom and takes into account the effect of atomic masses.

This approach is used to calculate fluctuations in a mass-weighted manner, helping us obtain more accurate results. Similarly, RMSF calculations were carried out using the cpptraj tool included in the Amber software suite.

- **AtomFluct**: The fluctuation value of each atom. It indicates how much the position of an atom deviates from its average position over time.
- **Mass**: The mass of each atom. This value depends on the chemical properties of the atom.
- $\sum \mathrm{AtomFluct}_i \times \mathrm{Mass}_i$: Represents the multiplication of each atom's fluctuation by its mass, summed over all atoms. In other words, each atom's fluctuation is calculated weighted by its mass.
- $\sum \mathrm{Mass}_i$: The total mass of all atoms. This sum includes the masses of all atoms in the system.

MM/GBSA and Binding Score Calculations

The second type of the Generalized Born (GB) model (GB/SA model) was selected, and the MM/GBSA energy of the system was calculated over all frames.

The parameter –radii mbondi2 was used for the atomic radii in these calculations. This approach was used to estimate the binding free energy of the complex in an implicit solvent environment.

In the MM/GBSA approach, the binding free energy of a ligand (L) to a receptor (R) forming a complex (RL) is computed as (Equation 4.4):

$$\Delta G_{\text{bind}} = G_{\text{complex}} - (G_{\text{receptor}} + G_{\text{ligand}}) \tag{4.4}$$

The total free energy of each component consists of gas-phase molecular mechanics energies and solvation free energies (Equation 4.5 4.6 4.7):

$$\Delta G_{\text{bind}} = \Delta E_{\text{MM}} + \Delta G_{\text{sol}} \tag{4.5}$$

$$\Delta E_{\text{MM}} = \Delta E_{\text{ele}} + \Delta E_{\text{vdW}} \tag{4.6}$$

$$\Delta G_{\text{sol}} = \Delta G_{\text{polar}} + \Delta G_{\text{nonpolar}} \tag{4.7}$$

Here, ΔE_{ele} and ΔE_{vdW} represent the electrostatic and van der Waals interactions in the gas phase, respectively [Wang et al., 2019]. The polar solvation term ΔG_{polar} is computed using the Generalized Born model, while the non-polar contribution $\Delta G_{\text{nonpolar}}$ is estimated based on the solvent-accessible surface area (SASA) (Equation 4.8):

$$\Delta G_{\text{nonpolar}} = \gamma \cdot \text{SASA} + b \tag{4.8}$$

The last 2 nanoseconds of the simulations were selected and quasi-harmonic entropy estimation was performed using ptraj. This calculation was carried out with the `entropy = 1` parameter, which enables entropy calculation. The purpose of this step was to estimate the entropic contribution to the binding free energy.

The binding free energy (ΔG_{bind}) was then calculated by subtracting the entropic contribution estimated using the quasi-harmonic approximation ($T\Delta S$) from the enthalpic terms obtained via the MM/GBSA method:

$$\Delta G_{\text{bind}} = \Delta G_{\text{MM/GBSA}} - T\Delta S \tag{4.9}$$

Both the entropy and MM/GBSA calculations were performed using the MMPBSA.py module included in the AmberTools package [Miller III et al., 2012].

4.3.5 Summary of MD Simulations and Analyses

A total of 26 different simulations, amounting to 78 microseconds (µs), will be performed within the scope of this thesis. The system details are presented in Table 4.1.

Table 4.1: General Summary of Simulations

System	Secondary Analyses	Duration	Repetition	Total Duration
Rac1 (Alone)				
Wild Type Rac1 (GTP)	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
Wild Type Rac1 (GDP)	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
F28L Rac1 (GTP)	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
F28L Rac1 (GDP)	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
P29S Rac1 (GTP)	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
P29S Rac1 (GDP)	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
Q61R Rac1 (GTP)	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
Q61R Rac1 (GDP)	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
Q61L Rac1 (GTP)	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
Q61L Rac1 (GDP)	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
T17N Rac1 (GTP)	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
T17N Rac1 (GDP)	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
Rac1 – PAK1 Complex				
Wild Type Rac1 (GTP) – PAK1	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
Wild Type Rac1 (GDP) – PAK1	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
F28L Rac1 (GTP) – PAK1	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
F28L Rac1 (GDP) – PAK1	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
P29S Rac1 (GTP) – PAK1	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
P29S Rac1 (GDP) – PAK1	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
Q61R Rac1 (GTP) – PAK1	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
Q61R Rac1 (GDP) – PAK1	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
Q61L Rac1 (GTP) – PAK1	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
Q61L Rac1 (GDP) – PAK1	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
T17N Rac1 (GTP) – PAK1	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
T17N Rac1 (GDP) – PAK1	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
Rac1 – TIPE2 Complex				
Wild Type Rac1 (GTP) – TIPE2	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
Wild Type Rac1 (GDP) – TIPE2	MMGBSA + 2 ns Entropy	1000 ns	3	3000 ns
			Total	**78 μs**

4.4 Visualization of Structures and Cavity Detection

CavitOmiX (v1.0) was used to detect cavities in the protein structures. Hydrophobicity assessment was performed with the VASCo program [Steinkellner et al., 2009], while cavity detection relied on a modified LIGSITE algorithm [Hendlich et al., 1997]."

For protein visualization and trajectory analysis PyMOL and VMD software used [Schrödinger, LLC, 2015a, Schrödinger, LLC, 2015b] .

RESULTS

5.1 Modeling Protein-Protein Complexes

5.1.1 Binding Residues

As a result of PDBsum analysis, the interface regions (active residues) for Rac3 and PAK1 were identified. According to PDBsum, the active residues for Rac3 are residues numbered 23, 24, 25, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 56, 64, 166, 170, 173, and 174, respectively. For PAK1, the active residues are numbered 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 25, and 32. In Figure 5.1, part A shows the active residues of Rac3 highlighted in raspberry color, while part B shows the active residues on PAK1 highlighted in raspberry color.

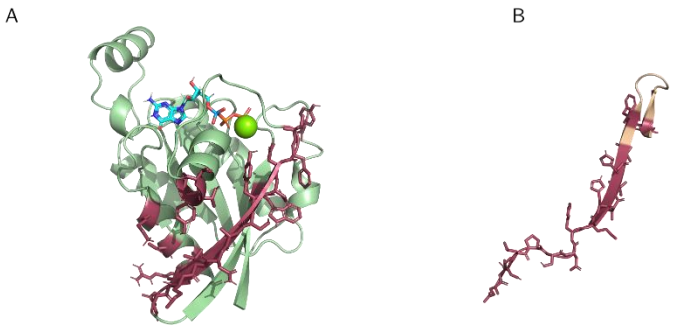


Figure 5.1: Interface Residues for Rac3 and PAK1 Proteins A- Active residues shown in raspberry color on Rac1 protein (shown on the 2QME structure) B- Active residues shown in raspberry color on PAK1 protein

As a result of PDBsum analysis, the interface regions (active residues) for Rac1 and TIPE2 were identified. According to PDBsum, the active residues for Rac1 are numbered 11, 12, 29, 30, 31, 33, 58, 59, 60, 61, 83, 84, 86, 87, 114, 126, 127, 132, 177, and 178, while for TIPE2 they are numbered 7, 10, 11, 13, 14, 17, 18, 21,

23, 27, 31, 133, 136, and 140. In Figure 5.2, part A shows the active residues of Rac1 highlighted in raspberry color, whereas part B shows the active residues of TIPE2 highlighted in raspberry color.

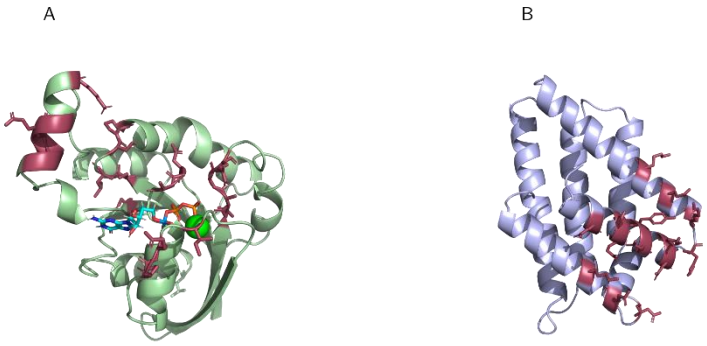


Figure 5.2: Interface Residues for Rac1 and TIPE2 Proteins A- Active residues shown in raspberry color on Rac1 protein B- Active residues shown in Raspberry color on TIPE2 protein

Docking with HADDOCK

Docking studies were performed between Rac1 and PAK1 using predefined active residues as input parameters, resulting in the generation of multiple clusters. From these, the clusters with the best Docking scores were selected. The corresponding Docking scores, cluster sizes, RMSD values within clusters, van der Waals energy, electrostatic energy, desolvation energy, restraints violation energy, buried surface area (BSA), and Z-scores for each set—namely Wild type, F28L, P29S, Q61R, Q61L, and T17N—are presented in Tables 5.1 to 5.12.

Notably, in structures where targeted mutations were introduced via PyMOL rather than using directly modeled mutant structures, Docking scores, cluster sizes, and other energy parameters were found to be similar to those of the initial structures, showing no significant differences. For example, the docking results for wild type Rac1 (GDP-bound) were comparable to those of the Q61R(GDP), Q61L(GDP), and T17N(GDP) mutants. A similar observation was

made for the docking results of the wild-type Rac1 (GTP-bound) with these mutant structures.

Overall, although no substantial differences were observed between wild-type and mutant forms, GTP-bound structures showed a relatively stronger binding affinity based on Docking scores.

Table 5.1: Haddock Docking score of Wild type Rac1(GTP)with PAK1

Parameter	**Value**
Docking score	-123.9 ± 0.8
Number of members in the cluster	126
RMSD	0.4 ± 0.2
VdW energy	-75.8 ± 1.9
Elec. energy	-118.1 ± 17.2
Desolv. energy	-26.3 ± 2.9
Restraints score	19.2 ± 6.6
Interface Area	1774.2 ± 25.2
Z-Score	-1.2

Table 5.2: Haddock Docking score of Wild type Rac1(GDP)with PAK1

Parameter	**Value**
Docking score	-105.1 ± 3.5
Number of members in the cluster	36
RMSD	4.1 ± 0.3
VdW energy	-71.1 ± 1.8
Elec. energy	-54.6 ± 46.7
Desolv. energy	-29.5 ± 2.0
Restraints score	63.9 ± 51.2
Interface Area	1761.0 ± 47.4
Z-Score	-1.7

Table 5.3: Haddock Docking score of F28L Rac1(GTP)with PAK1

Parameter	**Value**
Docking score	-128.6 ± 2.1
Number of members in the cluster	171
RMSD	0.5 ± 0.3
VdW energy	-70.3 ± 2.7
Elec. energy	-134.5 ± 14.8
Desolv. energy	-33.8 ± 1.8
Restraints score	24.7 ± 6.0
Interface Area	1866.8 ± 20.4
Z-Score	-1.0

Table 5.4: Haddock Docking score of F28L Rac1(GDP)with PAK1

Parameter	**Value**
Docking score	-126.9 ± 0.6
Number of members in the cluster	155
RMSD	0.9 ± 0.5
VdW energy	-71.6 ± 2.4
Elec. energy	-145.5 ± 8.7
Desolv. energy	-27.6 ± 3.6
Restraints score	14.7 ± 4.1
Interface Area	1818.6 ± 22.3
Z-Score	-1.0

Table 5.5: Haddock Docking score of P29S Rac1(GTP)with PAK1

Parameter	**Value**
Docking score	-119.1 ± 1.7
Number of members in the cluster	93
RMSD	0.3 ± 0.2
VdW energy	-73.9 ± 1.9
Elec. energy	-117.4 ± 9.4
Desolv. energy	-23.0 ± 2.7
Restraints score	12.7 ± 6.3
Interface Area	1690.2 ± 40.4
Z-Score	-1.9

Table 5.6: Haddock Docking score of P29S Rac1(GDP)with PAK1

Parameter	**Value**
Docking score	-116.5 ± 1.1
Number of members in the cluster	103
RMSD	0.4 ± 0.3
VdW energy	-74.0 ± 3.5
Elec. energy	-99.8 ± 12.0
Desolv. energy	-24.2 ± 3.7
Restraints score	16.6 ± 11.4
Interface Area	1687.8 ± 47.3
Z-Score	-1.5

Table 5.7: Haddock Docking score of Q61R Rac1(GTP)with PAK1

Parameter	**Value**
Docking score	-125.2 ± 2.3
Number of members in the cluster	144
RMSD	0.4 ± 0.2
VdW energy	-76.3 ± 0.6
Elec. energy	-120.1 ± 10.6
Desolv. energy	-26.3 ± 4.4
Restraints score	14.7 ± 8.7
Interface Area	1796.5 ± 40.5
Z-Score	-1.3

Table 5.8: Haddock Docking score of Q61R Rac1(GDP)with PAK1

Parameter	**Value**
Docking score	-104.7 ± 2.3
Number of members in the cluster	21
RMSD	26.5 ± 0.0
VdW energy	-57.7 ± 2.6
Elec. energy	-102.4 ± 4.1
Desolv. energy	-30.8 ± 2.9
Restraints score	42.4 ± 11.0
Interface Area	1723.9 ± 30.2
Z-Score	-1.2

Table 5.9: Haddock Docking score of Q61L Rac1(GTP)with PAK1

Parameter	**Value**
Docking score	-125.4 ± 2.9
Number of members in the cluster	138
RMSD	0.4 ± 0.3
VdW energy	-75.1 ± 3.1
Elec. energy	-114.3 ± 8.7
Desolv. energy	-29.3 ± 1.3
Restraints score	18.6 ± 4.0
Interface Area	1807.8 ± 16.8
Z-Score	-1.4

Table 5.10: Haddock Docking score of Q61L Rac1(GDP)with PAK1

Parameter	**Value**
Docking score	-111.1 ± 9.3
Number of members in the cluster	18
RMSD	38.9 ± 0.1
VdW energy	-62.7 ± 7.7
Elec. energy	-116.1 ± 7.8
Desolv. energy	-30.1 ± 4.9
Restraints score	49.0 ± 23.9
Interface Area	1732.5 ± 85.6
Z-Score	-1.7

Table 5.11: Haddock Docking score of T17N Rac1(GTP)with PAK1

Parameter	**Value**
Docking score	-124.1 ± 0.5
Cluster size	141
RMSD	0.6 ± 0.5
VdW energy	-74.3 ± 3.6
Elec. energy	-121.7 ± 17.8
Desolv. energy	-26.9 ± 0.9
Restraints score	14.6 ± 5.8
Interface Area	1817.1 ± 20.8
Z-Score	-1.3

Table 5.12: Haddock Docking score of T17N Rac1(GDP)with PAK1

Parameter	**Value**
Docking score	-110.3 ± 5.6
Cluster size	58
RMSD	0.7 ± 0.5
VdW energy	-60.0 ± 6.2
Elec. energy	-124.3 ± 11.4
Desolv. energy	-29.3 ± 1.2
Restraints score	39.4 ± 14.6
Interface Area	1794.6 ± 88.7
Z-Score	-1.7

5.1.2 Docking with Rosetta

As part of the docking workflow, 100,000 complex structures modeling the interaction between GTP-/GDP-bound Rac1 and TIPE2 were initially generated using Rosetta. These structures were then evaluated based on energy thresholds defined in the Methods section. After filtering, a subset of structures remained, and their binding affinities were further assessed using MM/GBSA calculations performed via the HawkDock server. Among these, structure 5 was identified as the top candidate model. The scores from Rosetta (”Total Energy” and ”Interface Energy”) and HawkDock (”Total MM/GBSA” and ”Total (VD) MM/GBSA”) are summarized in Tables 5.13 through 5.16, while Table 5.14 and Figure 5.3 highlight the selected models. Additionally, a detailed comparison of MM/GBSA and Rosetta energy scores for various Rac1–TIPE2 combinations is provided in Tables 5.13–5.16.

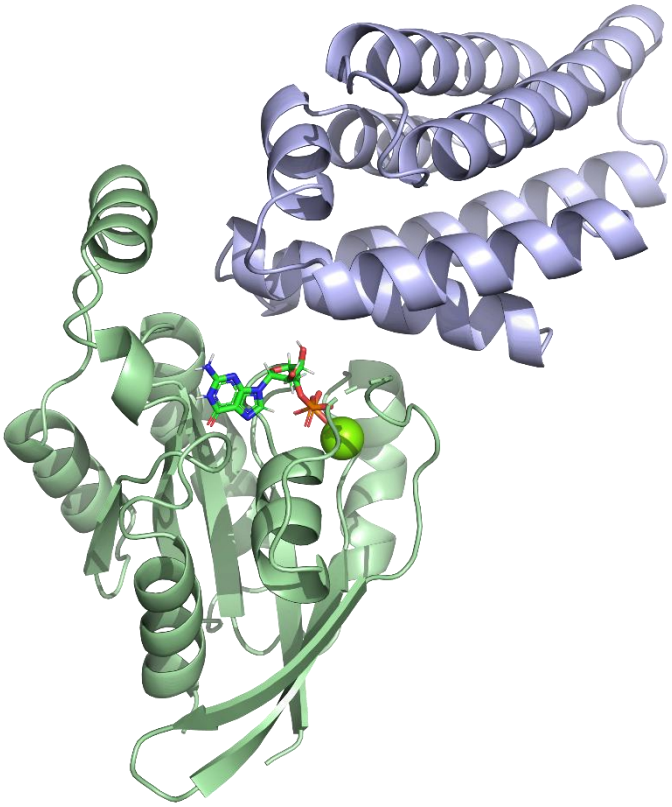


Figure 5.3: Docking Pose Selected to Study the Interaction Between Rac1 and TIPE2

Table 5.13: Rac1(GTP) vs TIPE2 Docking Poses. Energy values of the best pose in each set based on total score

Grup	Total MM/GBSA (kcal/mol)	Total (VD) MM/GBSA (kcal/mol)	Total Energy	Interface Energy
1	-42.79	-110.03	-619.367	-39.24
2	-17.4	-77.54	-1053.43	-21.416
3	-15.46	-104.44	-1054.31	-21.55
4	-20.86	-117.1	-1061.365	-20.402
5	-35.62	-60.19	-1054.614	-20.121
6	-18.51	-72.58	-1059.077	-20.37
7	-47.99	-93.22	-1062.3	-21.617
8	-37.3	-103.73	-1048.726	-24.768
9	-2.93	-82.26	-1056.367	-23.25
10	-2.0	-54.44	-1057.995	-23.837

Table 5.14: Rac1(GTP) vs TIPE2 Docking Poses. Energy values of the best pose in each set based on interface score

Grup	Interface MM/GBSA (kcal/mol)	Interface (VD) MM/GBSA (kcal/mol)	Total Energy	Interface Energy
1	-12.31	-40.79	-680.356	-27.561
2	-18.97	-75.54	-1035.64	-28.405
3	-8.31	-51.28	-1026.153	-27.477
4	-44.84	-148.59	-1021.737	-30.455
5	-54.39	-147.76	-1042.476	-29.736
6	-18.63	-78.22	-960.04	-26.986
7	-8.2	-34.82	-1040.904	-27.547
8	-30.71	-53.88	-1041.428	-31.164
9	-44.45	-109.14	-1003.071	-28.465
10	-9.36	-36.56	-1016.137	-28.266

Table 5.15: Rac1(GDP) vs TIPE2 Docking Poses. Energy values of the best pose in each set based on total score

Grup	Total MM/GBSA (kcal/mol)	Total (VD) MM/GBSA (kcal/mol)	Total Energy	Interface Energy
1	-41.15	-87.62	-670.627	-23.991
2	-0.29	-105.54	-1031.158	-20.558
3	-15.46	-104.44	-1037.013	-21.291
4	-6.65	-22.08	-1039.308	-20.921
5	-8.77	-54.19	-1026.7	-18.597
6	-29.47	-112.88	-1033.913	-23.979
7	-16.4	-55.46	-1036.13	-27.647
8	-34.89	-78.75	-1027.508	-19.341
9	-17.65	-87.87	-1032.976	-18.907
10	-6.06	-76.35	-1034.429	-17.132

Table 5.16: Rac1(GDP) vs TIPE2 Docking Poses. Energy values of the best pose in each set based on interface score

Grup	Interface MM/GBSA (kcal/mol)	Interface (VD) MM/GBSA (kcal/mol)	Total Energy	Interface Energy
1	-56.13	-131.8	-608.481	-29.605
2	-23.84	-120.19	-1024.705	-26.964
3	-8.31	-51.28	-1006.416	-28.117
4	-6.89	-62.08	-989.823	-31.244
5	-26.66	-81.11	-1021.064	-27.133
6	-39.76	-132.87	-1001.019	-28.626
7	-25.08	-82.86	-1028.161	-30.608
8	-2.29	-57.96	-1013.961	-30.936
9	-52.94	-114.89	-1017.347	-31.788
10	-55.45	-139.18	-1025.119	-25.703

5.2 Analysis of MD simulations

5.2.1 RMSD and RMSF

RMSD of Rac1 and Rac1-PAK1

- The results of the MD simulations conducted on Rac1 protein alone are summarized as follows:
- Wild-type Rac1: The GTP-bound form is significantly more stable compared to the GDP-bound form, as evidenced by lower and less fluctuating RMSD values (Figure 5.4 A1, A2).
- F28L mutant: Both GTP- and GDP-bound forms show slightly higher and more fluctuating RMSD profiles than the wild-type (Figure 5.4 B1, B2).
- P29S mutant: Structural differences between GDP and GTP forms were found insignificant; however, GDP form appears somewhat more stable (Figure 5.4 C1, C2).
- Q61L mutant: The GTP-bound form generally exhibits a more stable RMSD trajectory than the GDP-bound form, although it does not significantly deviate from the wild-type behavior (Figure 5.5 A1, BA).
- Q61R mutant: The GTP-bound form shows increased fluctuations relative to the wild-type, while the GDP-bound form is relatively stable except for Run 1, where minor instability is noted(Figure 5.5 B1, B2).
- T17N mutant: The GTP-bound form follows a motion pattern similar to the wild-type in all runs except Run 1. The GDP-bound form displays RMSD behavior comparable to the wild-type (Figure 5.5 C1, C2).

The results of the simulations of Rac1 in complex with PAK1 are summarized as follows:

- Wild-type Rac1: The GTP form is more stable than the GDP form, as indicated by lower RMSD values (Figure 5.6 A1, A2) .
- F28L mutant: Excluding Run 1 of the GTP form, no significant difference is observed between the GTP and GDP states, nor when compared to the wild-type (Figure 5.6 B1, B2)

- P29S mutant: No substantial variation is observed for the GTP and GDP forms (Figure 5.6 C1, C2).
- Q61L mutant: The GDP-bound form exhibits higher RMSD values compared to the wild-type, indicating reduced stability (Figure 5.7 A1, A2).
- Q61R mutant: Both GDP and GTP forms show elevated RMSD values relative to the wild-type, with the GDP form being particularly affected (Figure 5.7 B1, B2).
- T17N mutant: The RMSD profile is similar to the wild-type, indicating comparable stability (Figure 5.7 C1, C2).

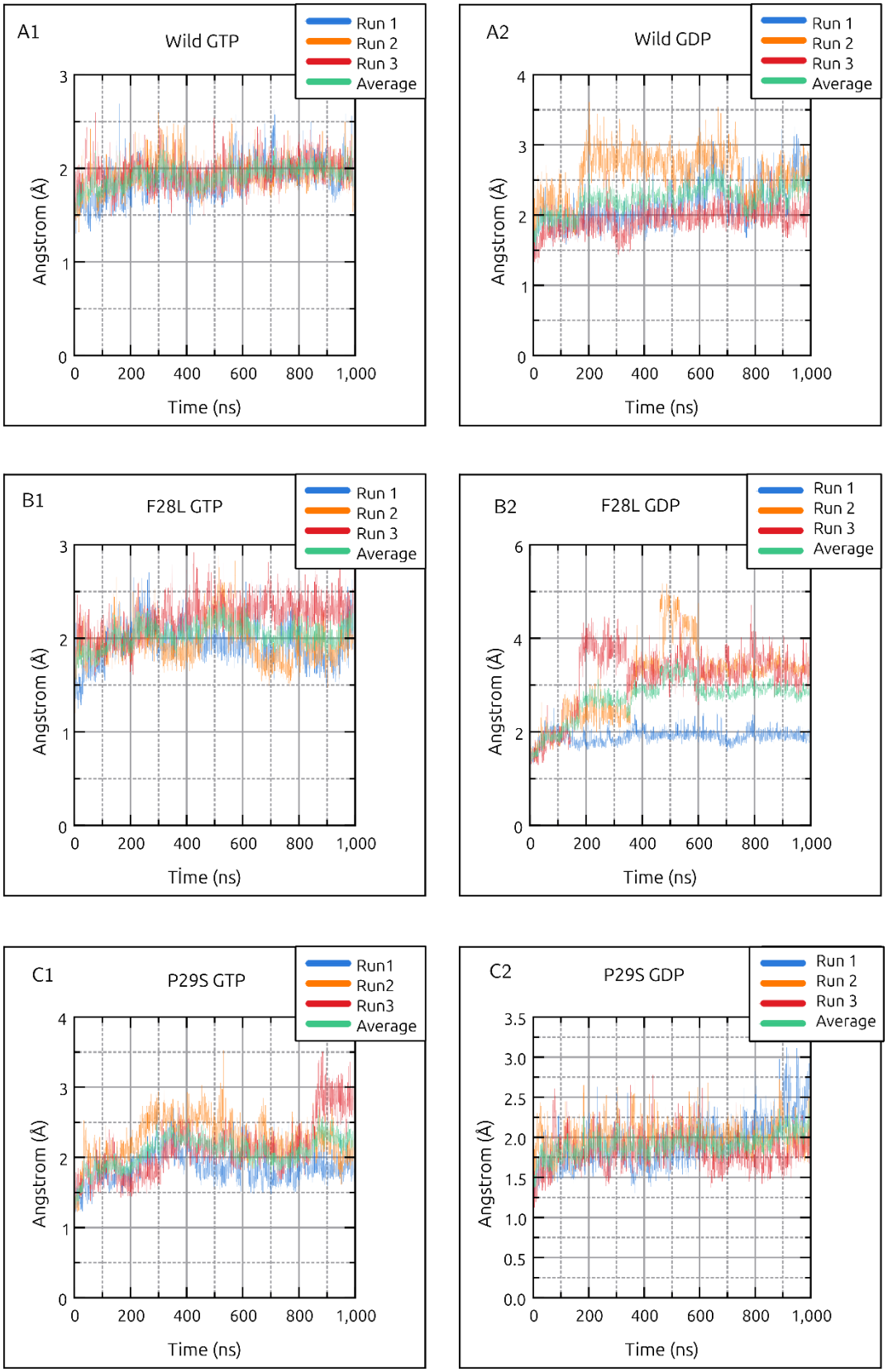


Figure 5.4: RMSD plots of Rac1 trajectories across various mutations and wild type (F28L,P29S)

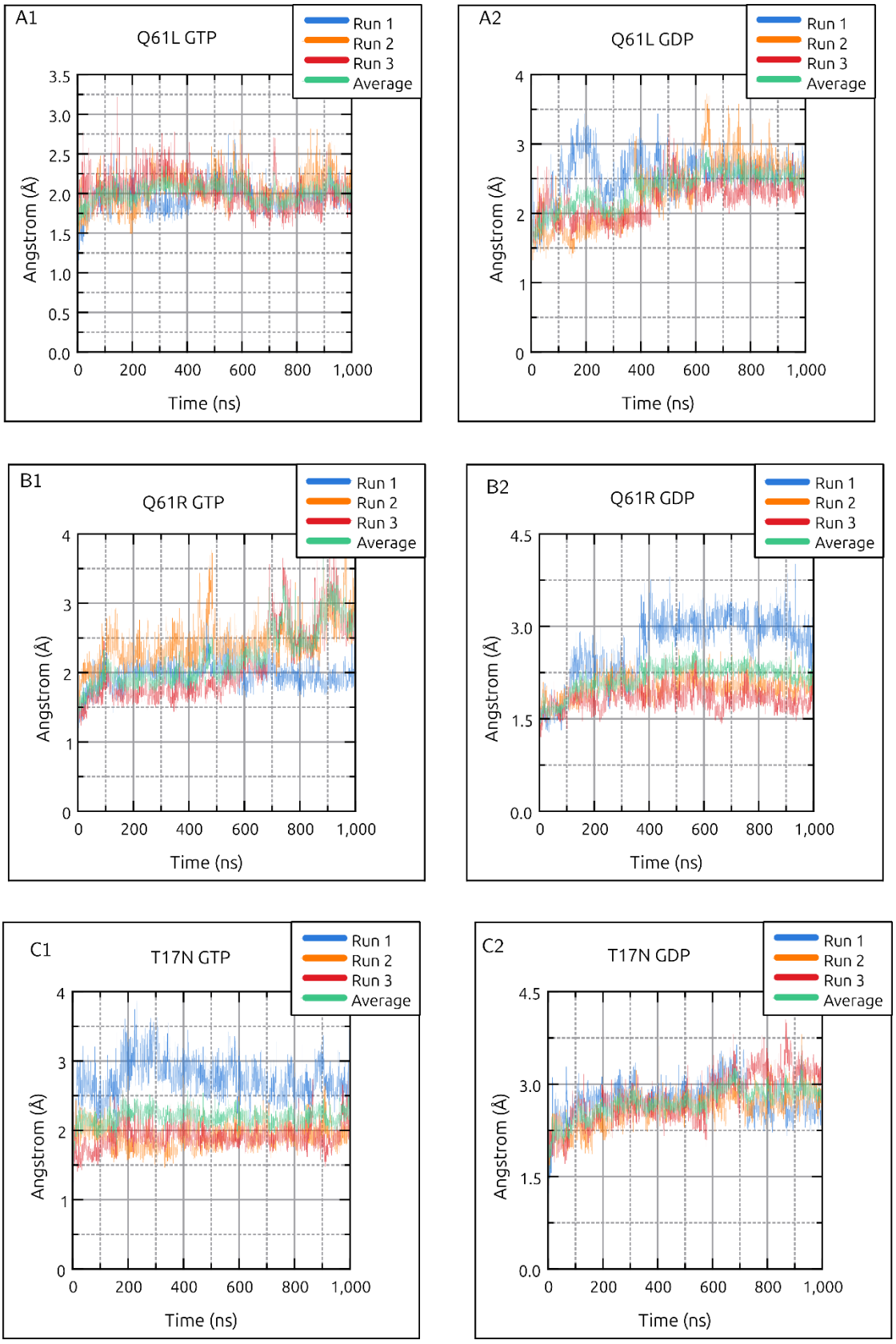


Figure 5.5: RMSD plots of Rac1 trajectories across various mutations (Q61R, Q61L,T17N)

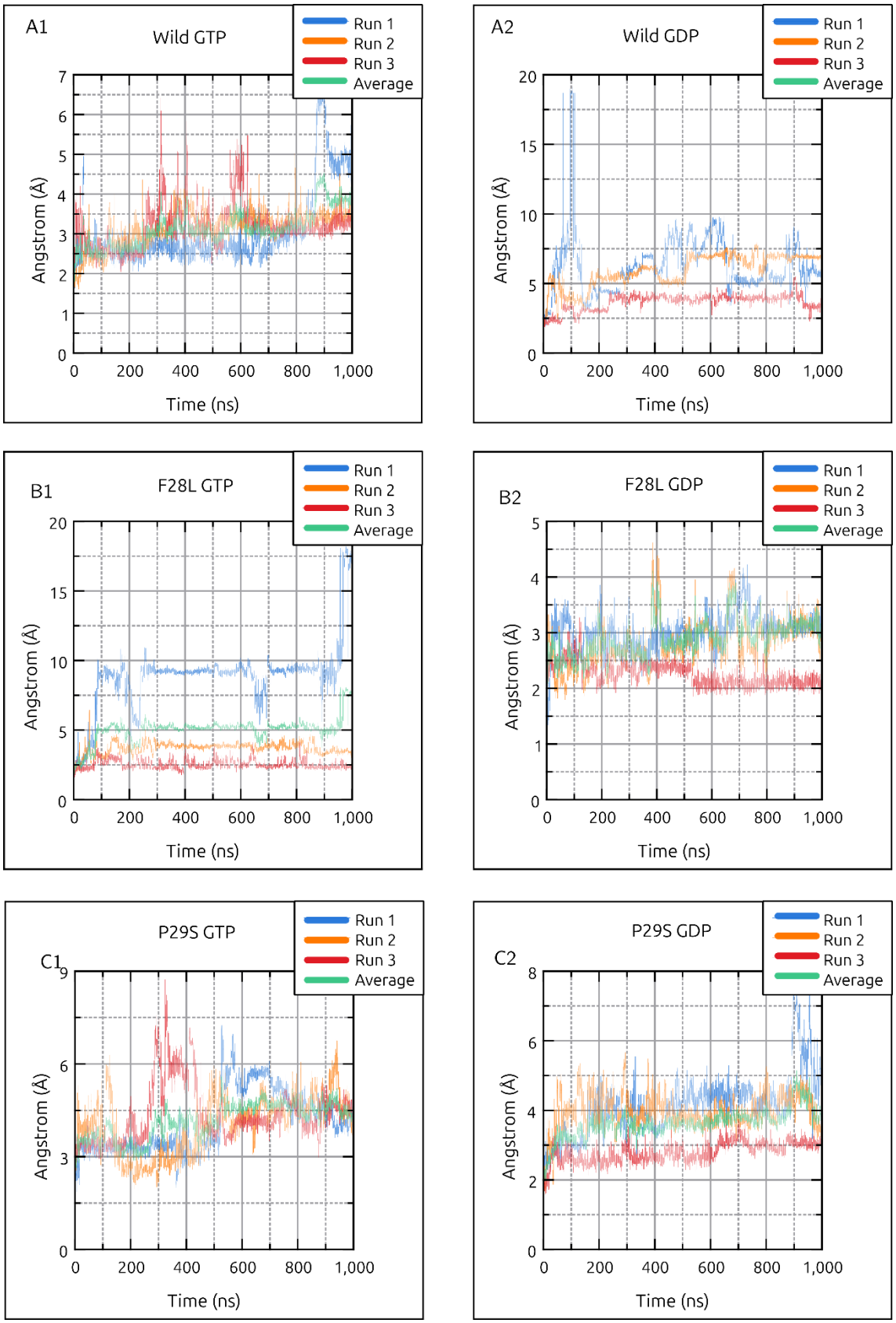


Figure 5.6: RMSD plots of trajectories for docking structures of Rac1 and PAK1 across wild type and various mutations (F28L, P29S)

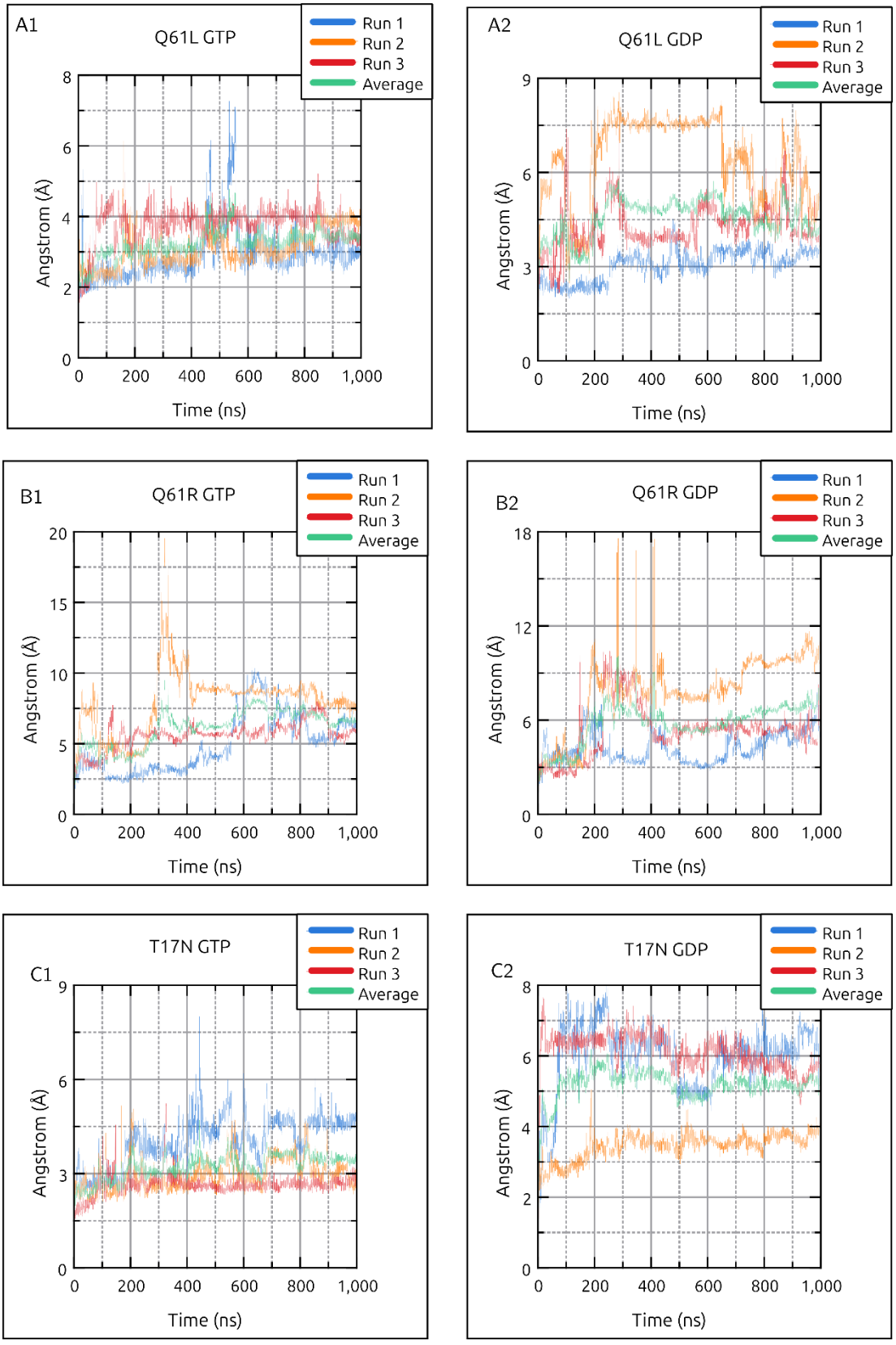


Figure 5.7: RMSD plots of trajectories for docked Rac1 and PAK1 structures across various mutations (Q61L, Q61R, T17N)

RMSD and Displacement Analysis of Rac1 and TIPE2

As a result of the three replicate simulations between Rac1 and TIPE2, the RMSD plots showed a stable plateau for Run 2 of Rac1(GDP) and Run 1 of Rac1(GTP). Upon examination of the simulation files, it was observed that these two simulations remained stable from the beginning to the end of the runs (Figure 5.9). Significant changes were observed in all remaining runs for both the GTP- and GDP-bound versions, with the GTP-bound Run 2 in particular reaching a plateau after 400 nanoseconds and remaining stable until the end of the simulation. Visualization of the simulation revealed that, especially after 400 nanoseconds, TIPE2 and Rac1 bound tightly to each other in a different conformation (Figure 5.9). These changes are also evident in Figure 5.8, which shows a comparative view of the first and last frames of the runs.

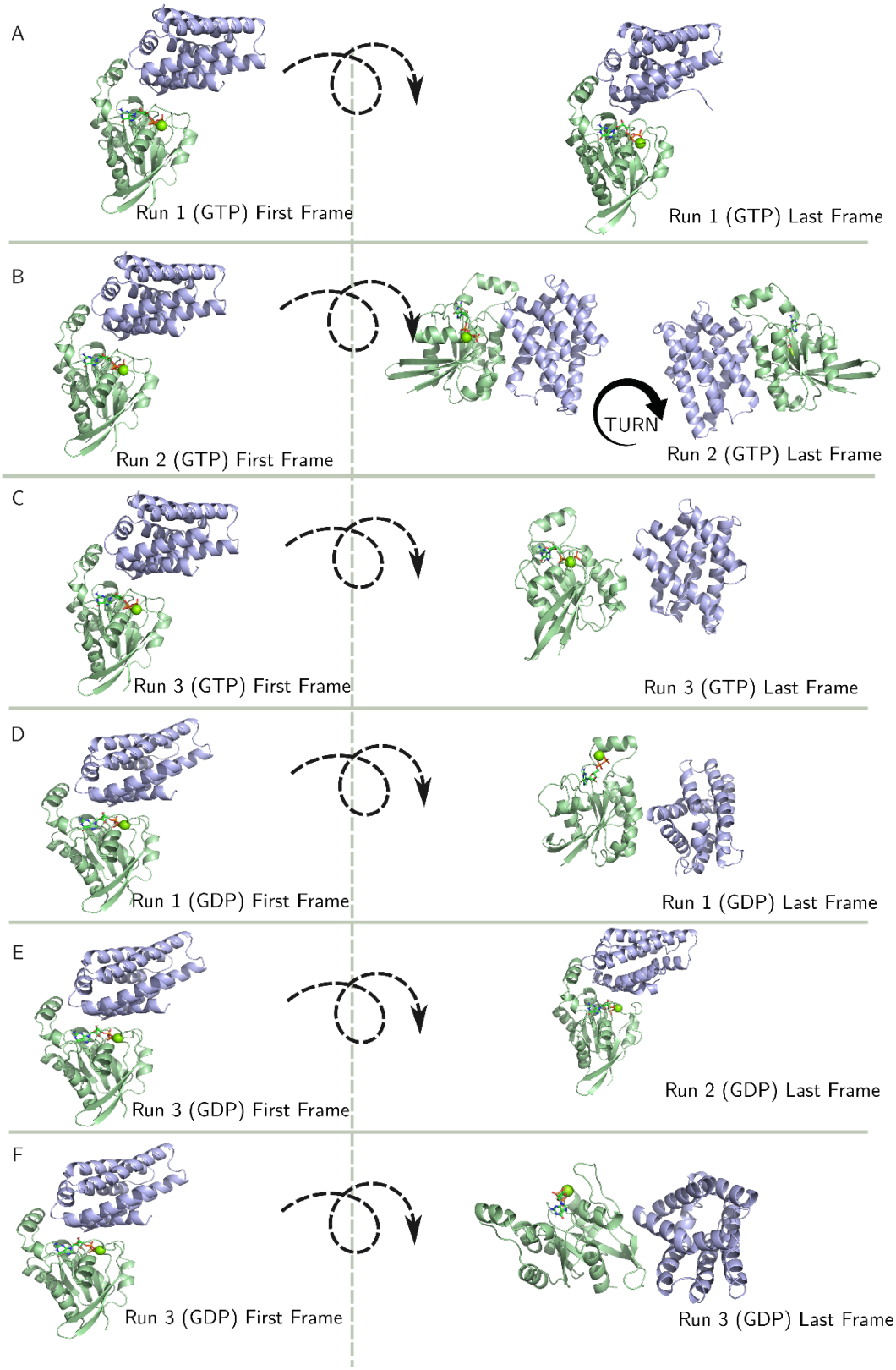


Figure 5.8: Panels A to C show the first and last frames of the three runs of Rac1(GTP) and TIPE2, while panels D to F show the first and last frames of the three runs of Rac1(GDP) and TIPE2

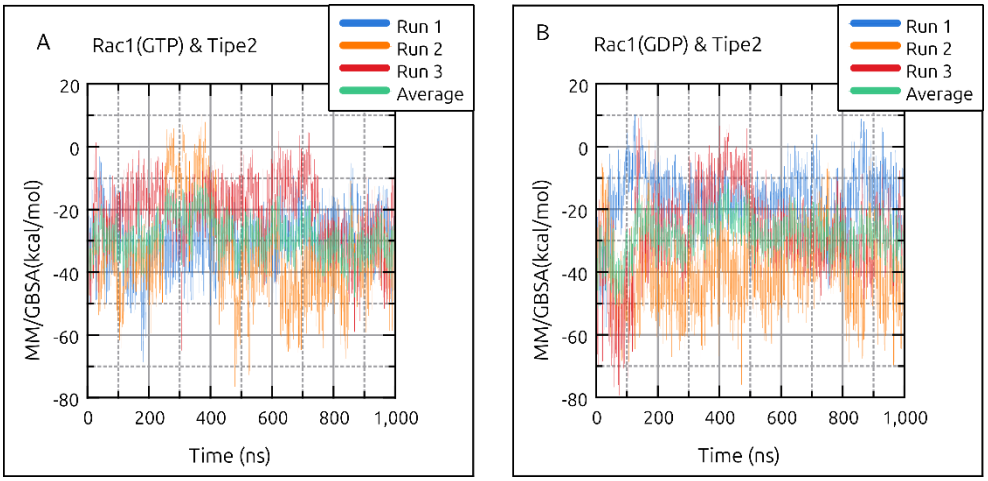


Figure 5.9: RMSD plots of Rac1(GTP)–TIPE2 and Rac1(GDP)–TIPE2 simulations

5.2.2 RMSF

The RMSF data from the MD simulations of Rac1 alone are summarized as follows:

- Wild-type Rac1: The GDP-bound form shows increased flexibility in the Switch I region and around the Mg2+/GDP-GTP binding site compared to the GTP-bound form (Figure 5.10 A1, A2 and Figure 5.11 A1, A2).
- F28L mutant: The GTP-bound form exhibits no major differences compared to the wild-type, while the GDP-bound form shows significantly higher flexibility in the Switch I region (Figure 5.10 B1, B2 and Figure 5.11 B1, B2).
- P29S mutant: In contrast to F28L, the GTP-bound form appears considerably more stable than the wild-type, whereas the GDP-bound form shows no significant deviation (Figure 5.10 C1, C2 and Figure 5.11 C1, C2).
- Q61L mutant: No substantial differences are observed in either binding state when compared to the wild-type (Figure 5.11 A1, A2 and Figure 5.11 A1, A2).
- Q61R mutant: Increased flexibility in the Switch I region is detected for both GTP- and GDP-bound forms relative to the wild-type (Figure 5.11 B1, B2 and Figure 5.11 B1, B2).
- T17N mutant: In both GTP- and GDP-bound forms, a marked increase in flexibility is observed in the Switch I region compared to the wild-type. Additionally, the GDP-bound form also displays increased fluctuations in

the Switch II region relative to the wild-type (Figure 5.11 C1, C2 and Figure 5.11 C1, C2)

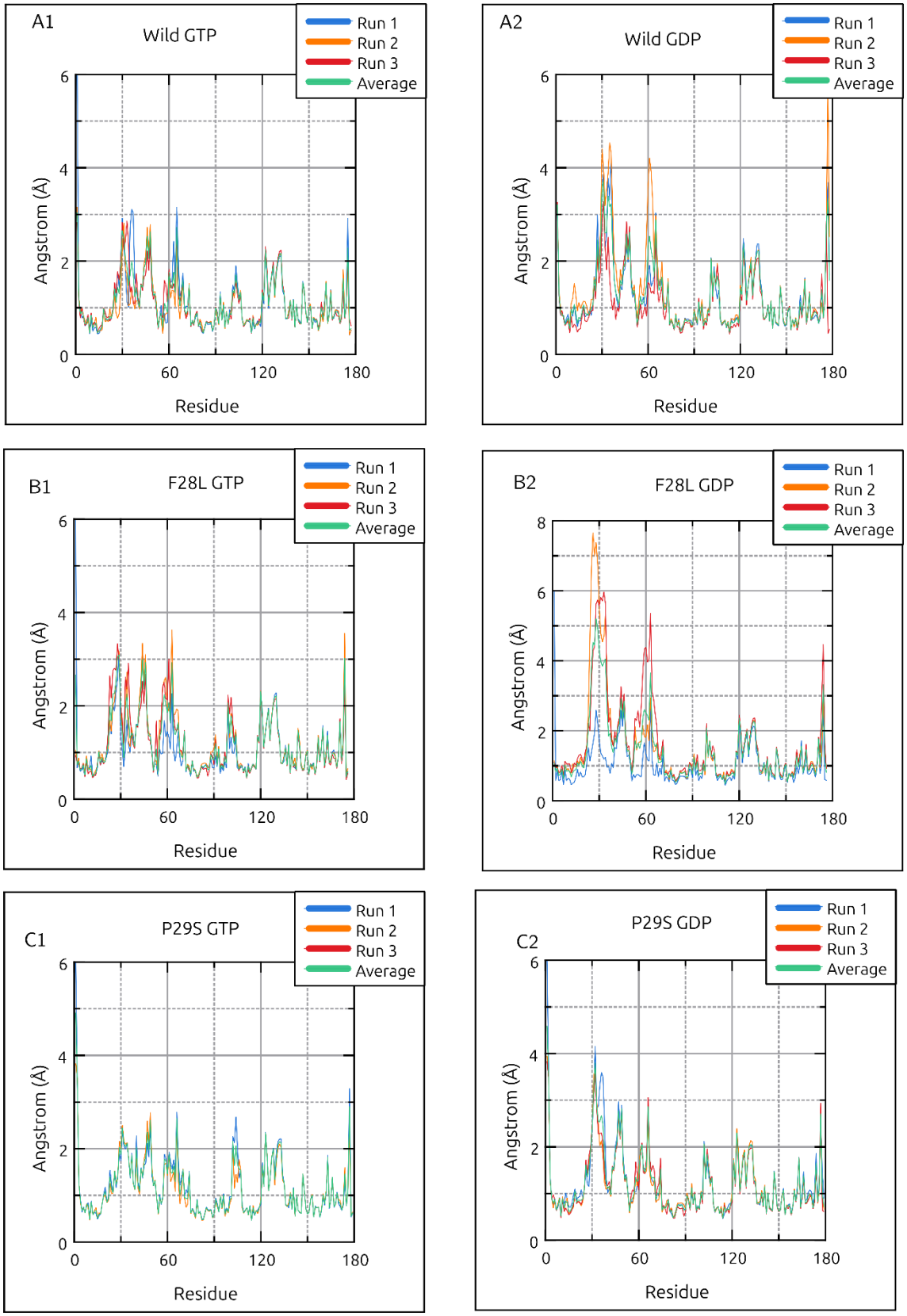


Figure 5.10: RMSF plots of Rac1 trajectories across various mutations and wild type (F28L,P29S)

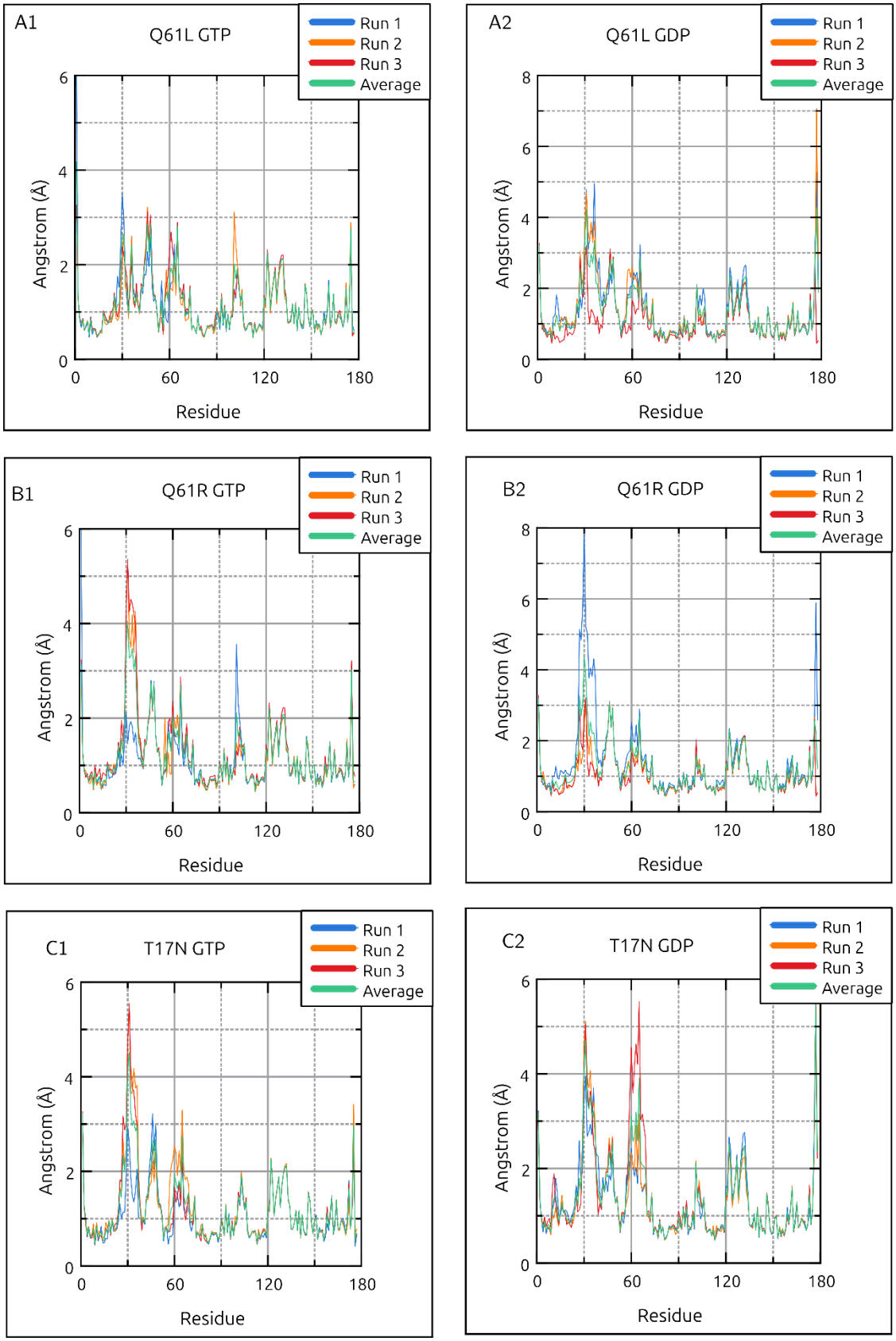


Figure 5.11: RMSF plots of Rac1 trajectories across various mutations (Q61R, Q61L, T17N)

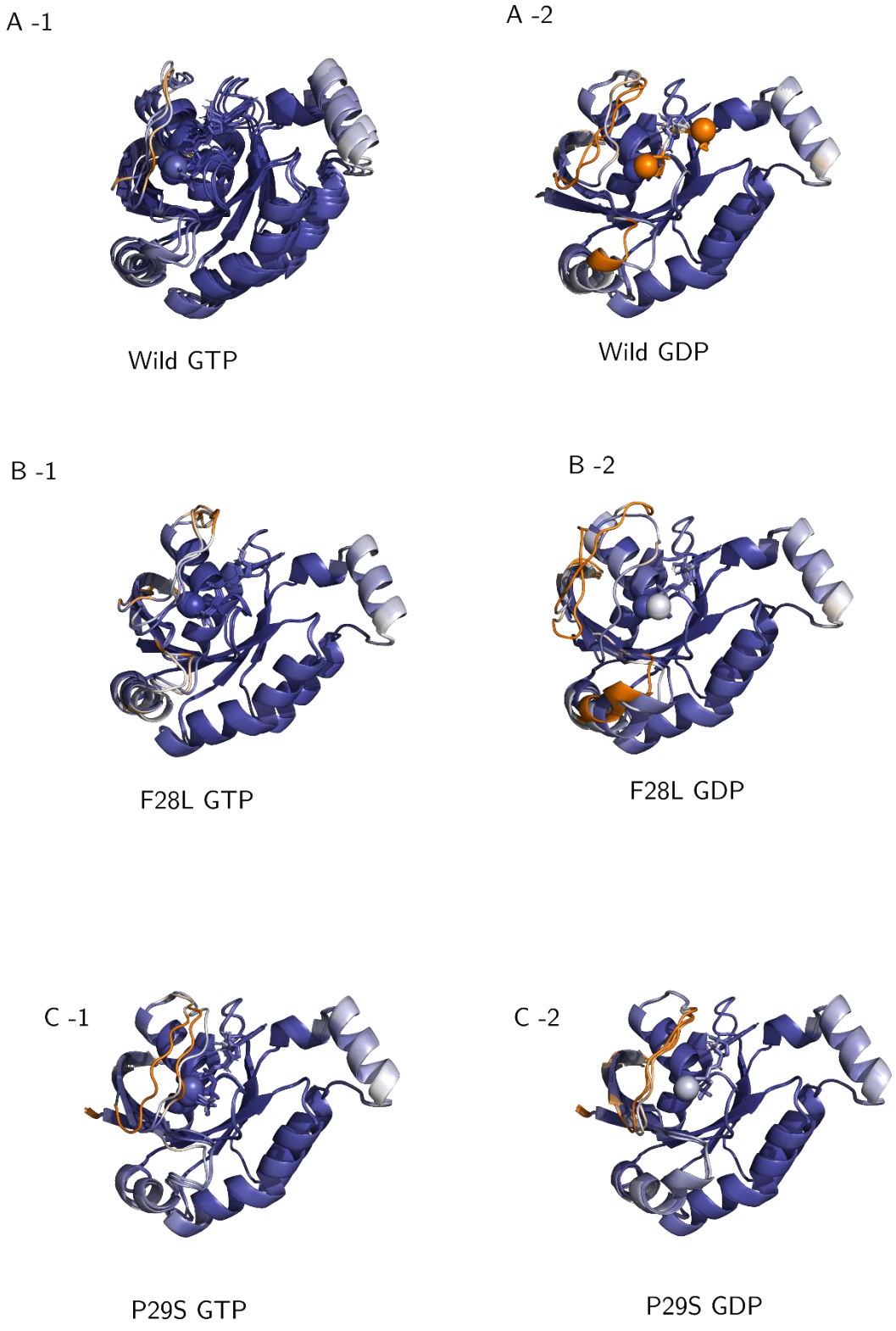


Figure 5.12: RMSF plots of Rac1 trajectories across various mutations and wild type (F28L,P29S)

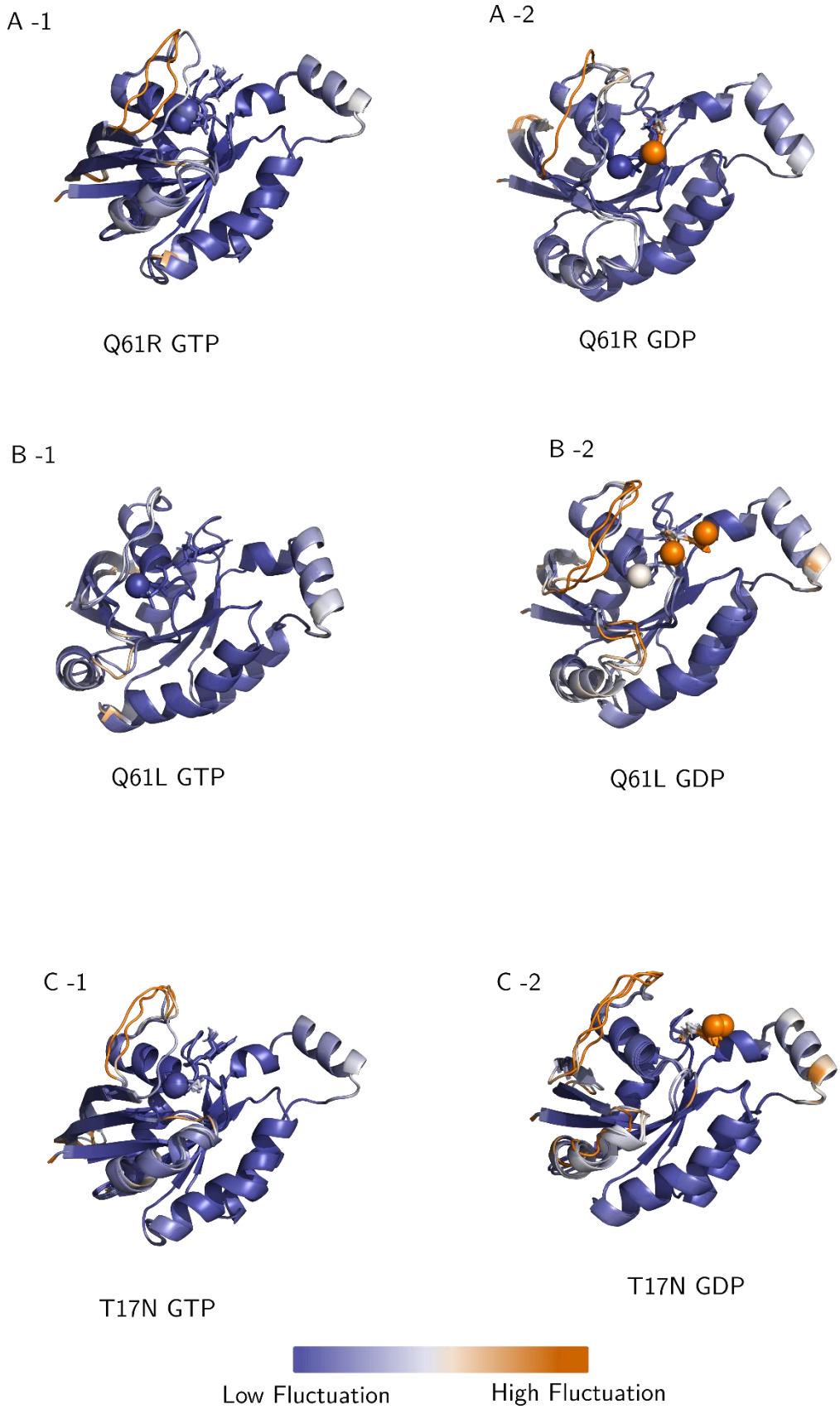


Figure 5.13: RMSF plots of Rac1 trajectories across various mutations (Q61R, Q61L)

5.2.3 Binding Dynamics Analysis with MM/GBSA and Binding Free Energy

Binding dynamics of wt/mutant Rac1 with the nucleotides GTP/GDP

When examining the complexes of Rac1 with the GTP/GDP nucleotides, various MM/GBSA scores for were observed. Figures 5.14 and 5.15 are the MM/GBSA value plots for GTP-/GDP-bound wild type and mutant Rac1s. Starting with the wild type, the observed energies and changes are as follows:

- For the wild type, an average energy of approximately −75 kcal/mol was observed for GTP, while GDP showed a highly fluctuating energy between −30 and −40 kcal/mol (Figure 5.14).
- For the F28L mutant, energy values ranged from −50 to −80 kcal/mol for GTP with an average around −65 kcal/mol. For GDP, the values varied between −35 and −70 kcal/mol with an average of about −55 kcal/mol (Figure 5.14). Although there appears to be a difference between GTP and GDP averages, due to high variance between runs, the difference is not statistically significant.
- For the P29S mutant, energy values ranged between −40 and −60 kcal/mol for GTP with an average around −50 kcal/mol. For the GDP version, the values fluctuated between −25 and −40 kcal/mol, with an average of approximately −35 kcal/mol, indicating a statistically significant difference between GTP and GDP. Compared to the wild type, similar values were observed for GDP, while slightly lower values were noted for GTP (Figure 5.14).
- For the Q61R mutant, energy values ranged from −60 to −90 kcal/mol for GTP with an average around −70 kcal/mol. In the GDP version, especially toward the end of the simulation (notably in Run 1), energies rose above −10 kcal/mol, resulting in an average of about −40 kcal/mol. The overall average energy for GDP was around −30 kcal/mol. Similar to the wild type, a distinction between GTP and GDP was observed (Figure 5.14).
- For the Q61L mutant, energy values ranged between −40 and −70 kcal/mol for GTP. Although the average energy was approximately −60 kcal/mol, the GBSA energy values showed significant variation among runs. In the GDP version, Run 3 showed energies between −40 and −50 kcal/mol toward the end of the simulation, whereas Runs 1 and 2 were close to each other around −15 kcal/mol. The average energy for the GDP version was around −25 kcal/mol. Similar to the wild type, changes between GTP and GDP were observed (Figure 5.14).

- For the T17N mutant, both GTP and GDP versions showed very little variance between runs in GBSA energy values. According to these data, GTP energies ranged from −40 to −60 kcal/mol with an average around −50 kcal/mol. For GDP, these values varied between −10 and −15 kcal/mol (Figure 5.14)

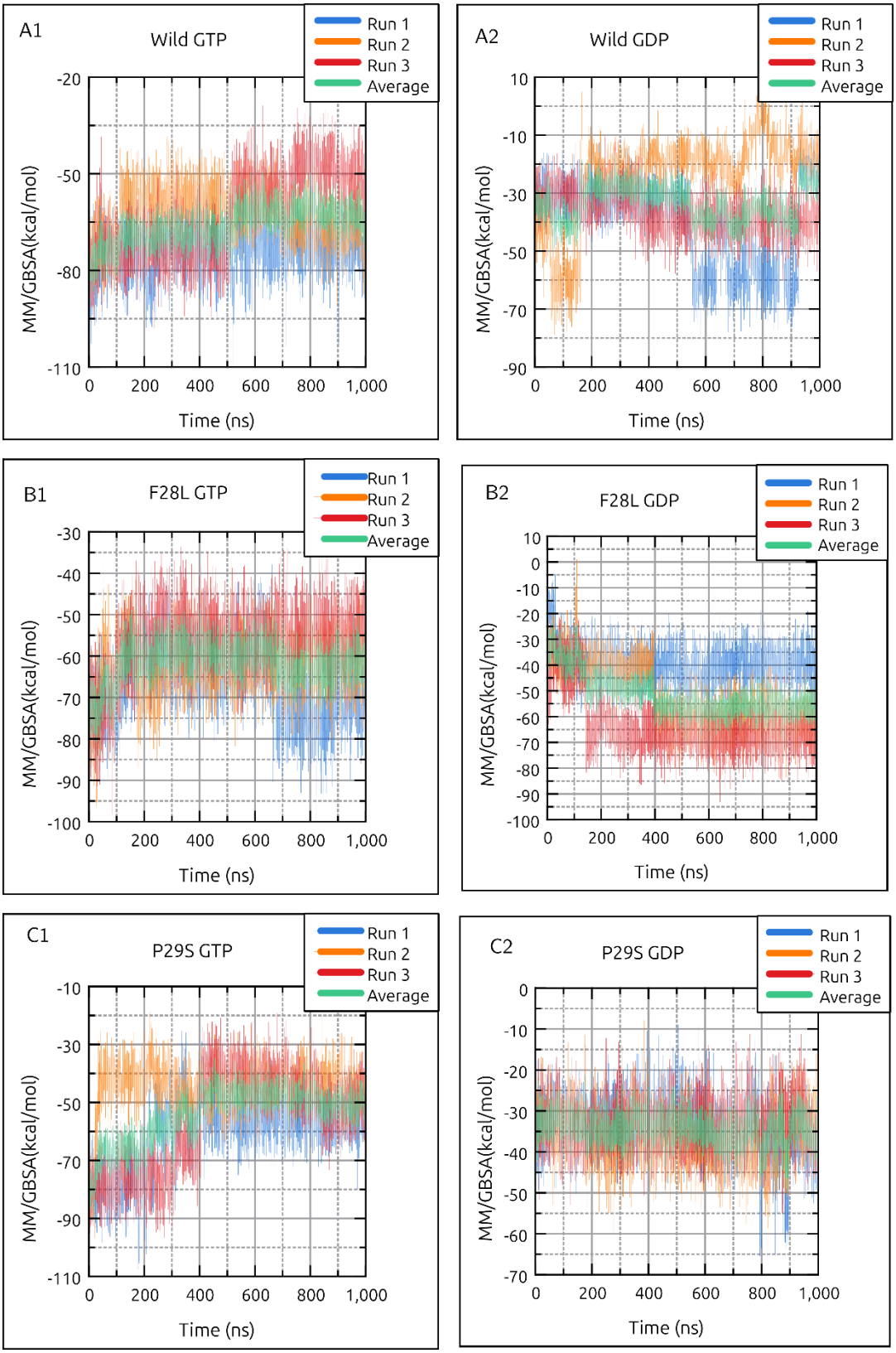


Figure 5.14: MMPBSA Energy Data for GTP- and GDP-bound Rac1 Wild Type, F28L, and P29S Mutants

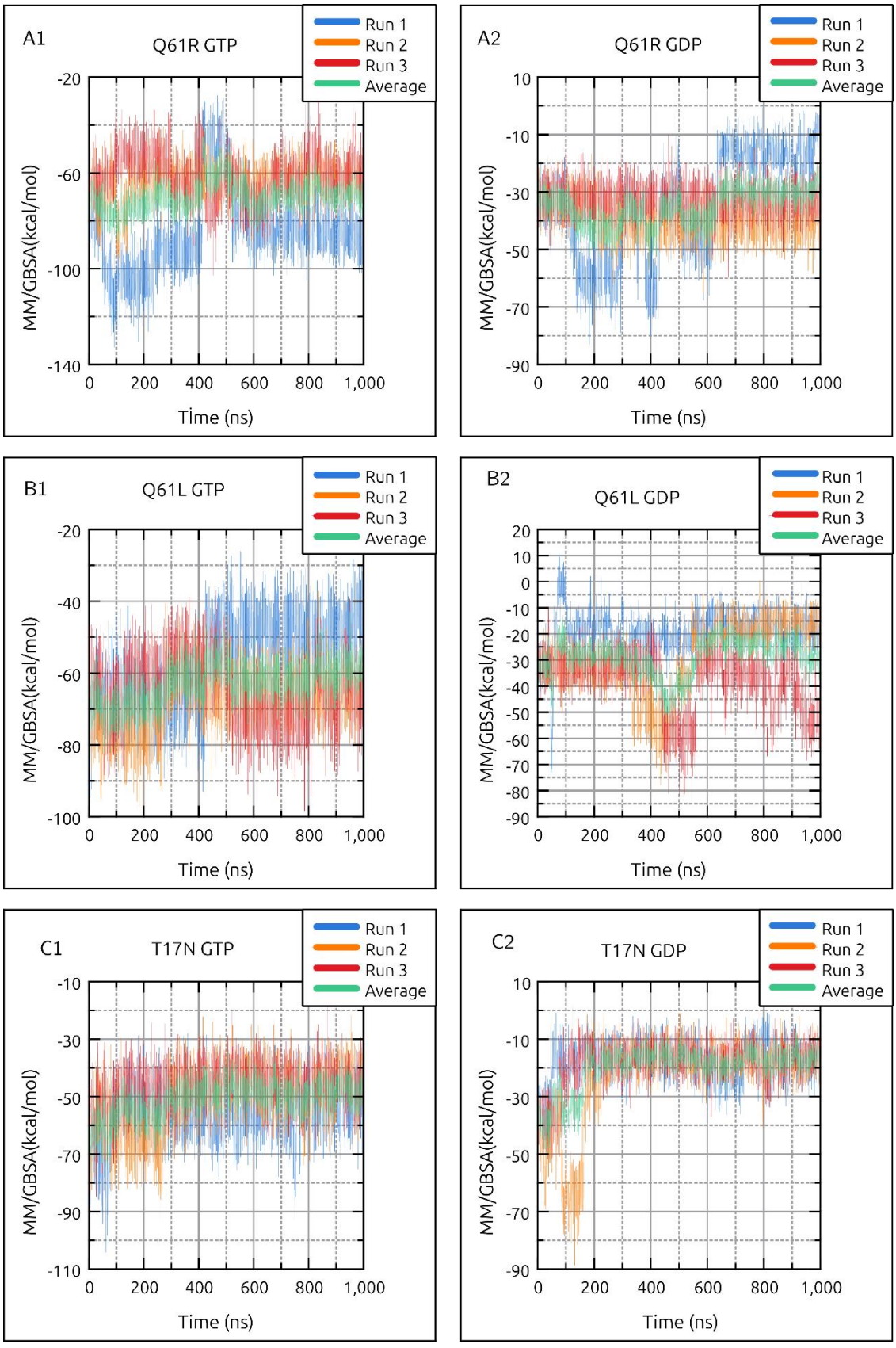


Figure 5.15: MMPBSA Energy Data for GTP- and GDP-bound Rac1 Q61R, Q61L and T17N Mutants

Entropy Analysis for the Rac1-GTP/GDP Complexes

Entropy calculations were performed to carry out binding energy estimations, and the translational, rotational, and vibrational entropy energy components along with the total entropy values for each run are presented in Table 5.17. Specifically, the entropy values for both GDP- and GTP-bound states were calculated separately to observe the differences in binding dynamics between these two nucleotide forms. The averages of the obtained entropy values and their comparison against the wild type, together with related interpretations, are provided in Table 5.18.

Table 5.17: Translational, Rotational, and Vibrational Entropy Contributions and Total $-T\Delta S$ Values (kcal/mol) for Protein-Ligand (GTP/GDP) Complexes. Values Obtained from Three Replicates for Each Complex.

Protein-Ligand	**Translational**	**Rotational**	**Vibrational**	**Total**
Wild (GTP)-1	-13.3130	-11.2903	-12.9636	-37.5672
Wild (GTP)-2	-13.3130	-11.2844	-13.7376	-38.3349
Wild (GTP)-3	-13.3130	-11.2760	-13.0807	-37.6701
Wild (GDP)-1	-13.1773	-10.9892	-15.9069	-40.0734
Wild (GDP)-2	-13.1773	-10.9999	-14.0011	-38.1784
Wild (GDP)-3	-13.1773	-11.0533	-12.6377	-36.8680
F28L (GTP)-1	-13.3127	-11.3109	-14.5736	-39.1972
F28L (GTP)-2	-13.3127	-11.3055	-14.7838	-39.4020
F28L (GTP)-3	-13.3127	-11.2886	-13.3944	-37.9962
F28L (GDP)-1	-13.1770	-10.9916	-12.1484	-36.3168
F28L (GDP)-2	-13.1770	-10.9850	-12.2602	-36.4220
F28L (GDP)-3	-13.1770	-10.9856	-12.5098	-36.6724
P29S (GTP)-1	-13.3130	-11.2900	-13.7206	-38.3239
P29S (GTP)-2	-13.3130	-11.2796	-14.4328	-39.0254
P29S (GTP)-3	-13.3130	-11.2343	-15.1910	-39.7383
P29S (GDP)-1	-13.1773	-11.0697	-14.9164	-39.1635
P29S (GDP)-2	-13.1773	-11.0477	-14.3365	-38.5615
P29S (GDP)-3	-13.1773	-11.0444	-12.4549	-36.6763
Q61L (GTP)-1	-13.3127	-11.2939	-13.4466	-38.0535
Q61L (GTP)-2	-13.3127	-11.2868	-13.8067	-38.4065
Q61L (GTP)-3	-13.3127	-11.2897	-13.9820	-38.5851
Q61L (GDP)-1	-13.1773	-10.8813	-15.8195	-39.8782
Q61L (GDP)-2	-13.1773	-10.9597	-16.6889	-40.8263
Q61L (GDP)-3	-13.1773	-11.0286	-12.5038	-36.7094
Q61R (GTP)-1	-13.3130	-11.2897	-12.7373	-37.3400
Q61R (GTP)-2	-13.3130	-11.2832	-14.4937	-39.0901
Q61R (GTP)-3	-13.3130	-11.2889	-14.7996	-39.4017
Q61R (GDP)-1	-13.1773	-11.0318	-16.0035	-40.2133
Q61R (GDP)-2	-13.1773	-11.0614	-12.9859	-37.2243
Q61R (GDP)-3	-13.1773	-11.0617	-12.6967	-36.9360
T17N (GTP)-1	-13.3130	-11.3103	-13.3005	-37.9244
T17N (GTP)-2	-13.3130	-11.2850	-14.1416	-38.7392
T17N (GTP)-3	-13.3130	-11.2763	-13.8440	-38.4330
T17N (GDP)-1	-13.1773	-10.9737	-16.9564	-41.1074
T17N (GDP)-2	-13.1773	-10.9952	-14.9955	-39.1677
T17N (GDP)-3	-13.1773	-10.9505	-17.1725	-41.3009

Table 5.18: Entropy Contribution $-T\Delta S$ Values of the Groups, with Wild Type Mean and Comparative Analysis

Grup	**Mean $-T\Delta S$ (kcal/mol)**	**Wild Type Mean (kcal/mol)**	**Comparison and Commentary**
Wild (GTP)	-37.86 ± 0.40	-37.86	Reference group
Wild (GDP)	-38.37 ± 1.61	-38.37	Reference group
F28L (GTP)	-38.86 ± 0.76	-37.86	More negative → decreased stability due to entropy
F28L (GDP)	-36.47 ± 0.18	-38.37	More positive → increased stability due to entropy
P29S (GTP)	-39.03 ± 0.73	-37.86	More negative → decreased stability due to entropy
P29S (GDP)	-38.13 ± 1.34	-38.37	No significant difference → slight entropy decrease
Q61L (GTP)	-38.35 ± 0.29	-37.86	More negative → decreased stability
Q61L (GDP)	-39.14 ± 2.34	-38.37	More negative → decreased stability (High ± value)
Q61R (GTP)	-38.61 ± 1.04	-37.86	More negative → decreased stability
Q61R (GDP)	-38.12 ± 1.81	-38.37	No significant difference → slight entropy decrease
T17N (GTP)	-38.36 ± 0.41	-37.86	More negative → decreased stability
T17N (GDP)	-40.53 ± 1.23	-38.37	More negative → decreased stability due to entropy

Binding Free Energy Analysis for Rac1-GTP/GDP Complexes

- The binding energy values obtained by including entropy are shown in Table 5.19. Starting from the wild type, the changes observed in the wild type and mutants are as follows:
- In the wild type, GTP binds to Rac1 much better compared to GDP.
- For the P29S mutation, although GTP binds better than GDP, when compared to the wild type, the binding energy of GTP is higher while the energy of GDP is generally lower on average.
- In F28L, despite the presence of both low and high energy values among the runs for both GTP and GDP versions, the average binding energy difference between GTP and GDP disappeared since it is similar to the wild type on average.
- For Q61R, except for the extremely low energy observed in Run 1 of GTP, the binding free energies of GTP and GDP generally showed a similar pattern to the wild type.
- For Q61L, except for the high energy in Run 1 of GTP, the binding free energies of GTP and GDP generally showed a similar pattern to the wild type.

- For T17N, unlike the wild type, GTP was observed to bind to Rac1 much more weakly compared to the wild type. No significant difference was observed for GDP compared to wild type Rac1.

Table 5.19: Binding free energy values of GTP/GDP obtained from three replicate runs and their averages

Grup	Run 1 (kcal/mol)	Run 2 (kcal/mol)	Run 3 (kcal/mol)	Mean (kcal/mol)
Wild GTP	-29.1726	-34.8818	-16.3894	-26.8146
Wild GDP	20.4112	22.8804	-2.0245	13.7557
P29S GTP	-18.5448	-7.9114	-5.7778	-10.7447
P29S GDP	-4.9331	6.0022	0.7755	0.6149
F28L GTP	-31.2019	-22.5391	-13.2315	-22.3242
F28L GDP	-2.8429	-27.1299	-34.3269	-21.4332
Q61R GTP	-53.2822	-24.6174	-23.5046	-33.8014
Q61R GDP	30.6781	1.8698	2.3416	11.6298
Q61L GTP	-0.7745	-26.6507	-26.6612	-18.0288
Q61L GDP	25.1427	24.9845	-15.4262	11.5670
T17N GTP	-18.3836	-5.1725	-4.7596	-9.4386
T17N GDP	23.1526	23.2267	25.8520	24.0771

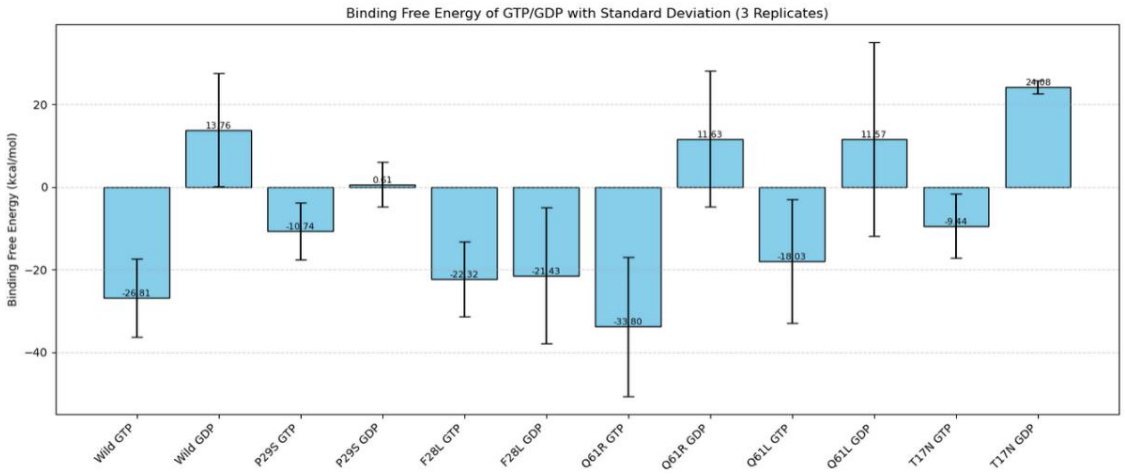


Figure 5.16: Binding free energy graph for Rac1. Binding free energy values (in kcal/mol) of Rac1 in complex with GTP or GDP, calculated from three replicate molecular dynamics simulations. Bars represent the mean values, and error bars denote standard deviations. Variants of Rac1, including wild type and mutants (P29S, F28L, Q61R, Q61L, T17N), are shown for both GTP and GDP bound forms

MM/GBSA Energy Analysis for Rac1-PAK1 Complex

The MM/GBSA binding energy scores between Rac1 and PAK1 are summarized as follows:

- Wild-type Rac1: The GTP-bound form exhibits an average binding energy of approximately –75 kcal/mol, indicating stronger binding, while the GDP bound form shows a higher (weaker) average binding energy of about –60 kcal/mol (Figure 5.17 A1, A2).
- F28L mutant: The GTP-bound form does not differ significantly from the wild-type. However, the GDP-bound form demonstrates enhanced binding affinity, with an average energy around –80 kcal/mol, indicating stronger binding than the wild-type GDP-bound form (Figure 5.17 B1, B2).
- P29S mutant: No substantial changes in binding energy are observed in either the GTP- or GDP-bound forms compared to the wild-type (Figure 5.17 C1, C2).
- Q61L mutant: The GTP-bound form remains comparable to the wild-type, while the GDP-bound form shows a notably stronger binding profile, reaching an average energy of about –80 kcal/mol toward the end of the simulation (Figure 5.18 A1, A2).
- Q61R mutant: Both the GTP- and GDP-bound forms exhibit unstable binding energy profiles and lower binding affinities relative to the wild-type, indicating impaired interaction with PAK1 (Figure 5.18 B1, B2) .
- T17N mutant: The GTP-bound form is similar to the wild-type, whereas the GDP bound form displays a slightly improved binding affinity toward the end of the simulation (Figure 5.18 C1, C2).

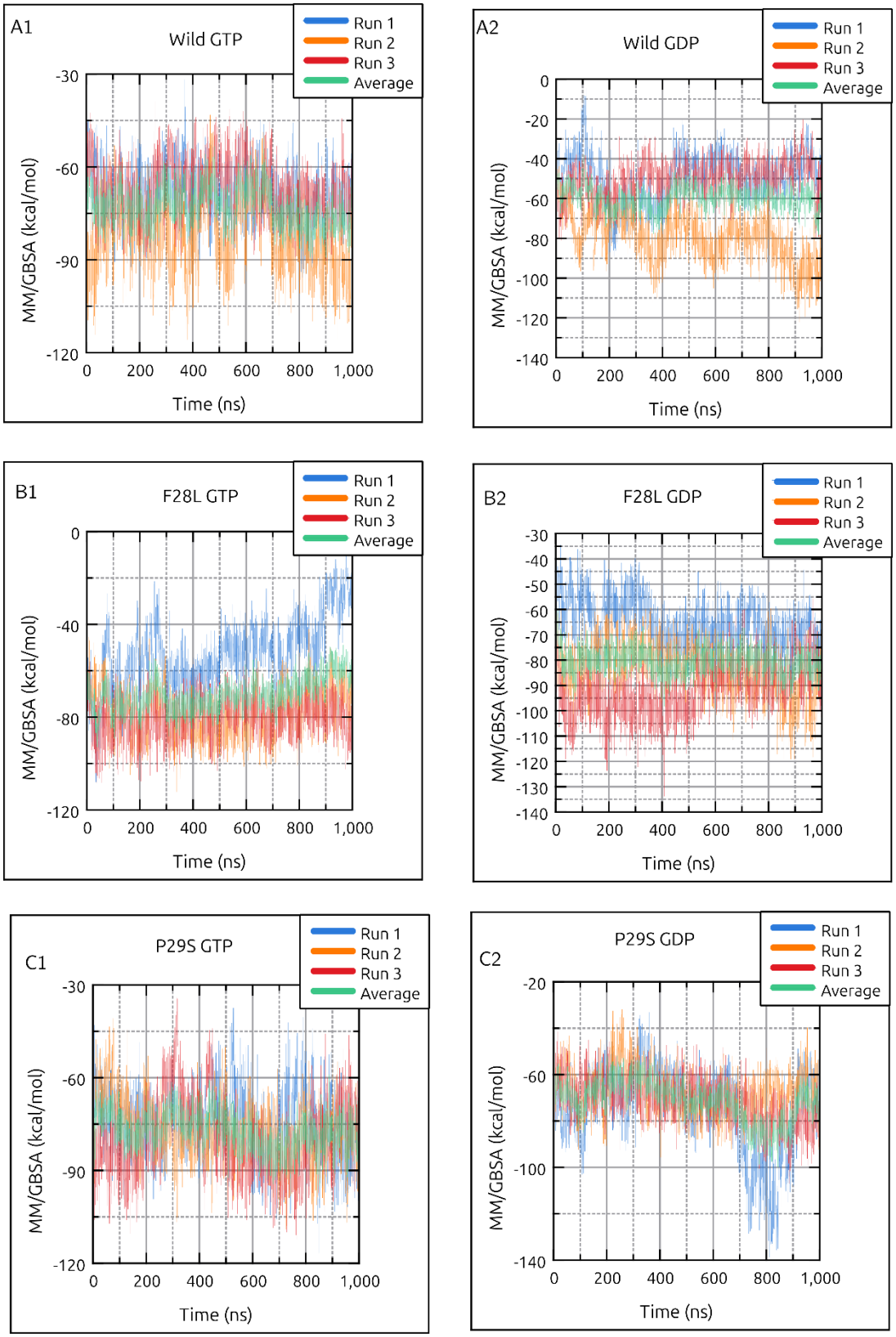


Figure 5.17: MM/GBSA plots of trajectories for docking structures of Rac1 and PAK1 across wild type and various mutations (F28L, P29S)

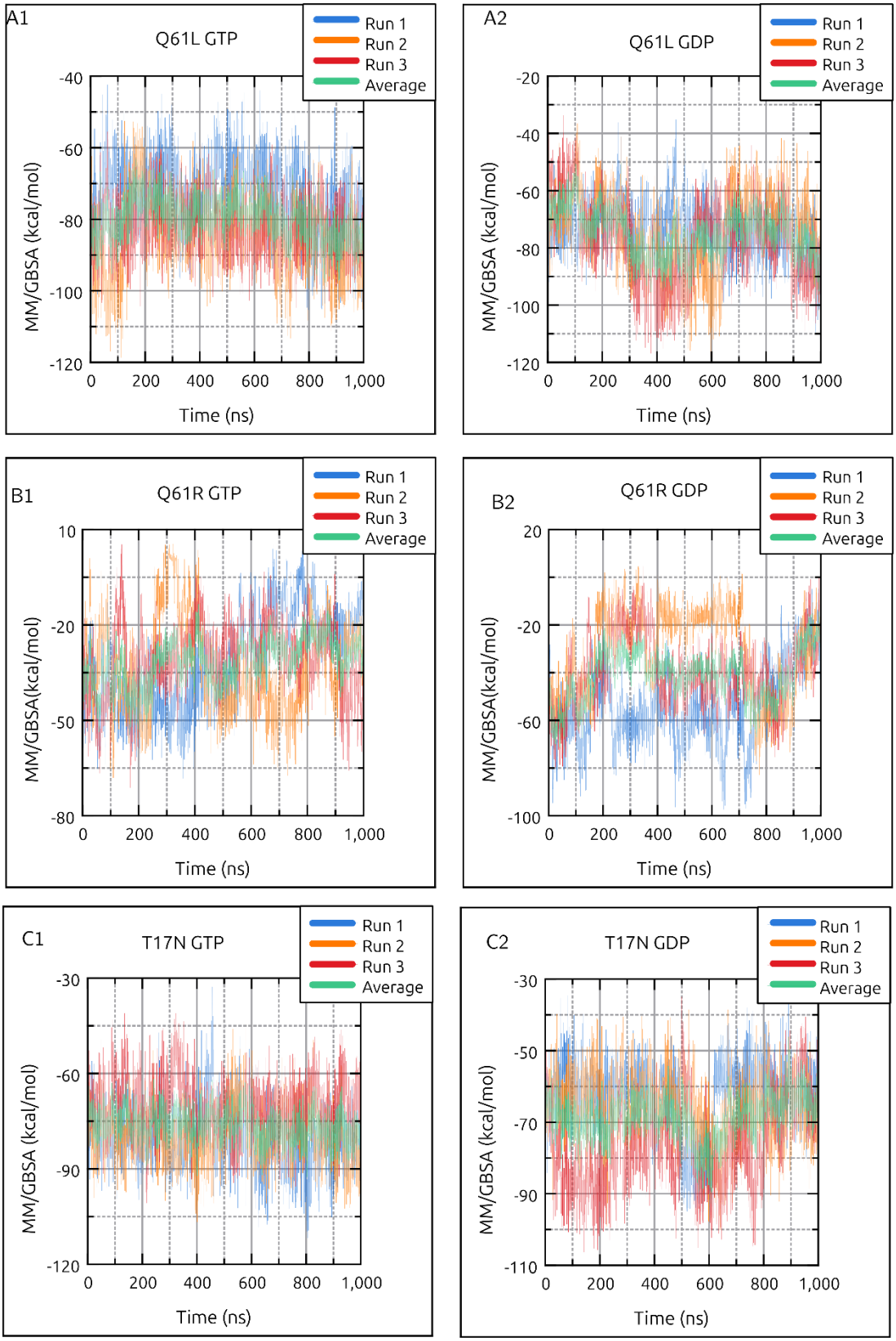


Figure 5.18: MM/GBSA plots of trajectories for docked Rac1 and PAK1 structures across various mutations (Q61L, Q61R, T17N)

Entropy Analysis for Rac1-PAK1 Complex

Table 5.20: Translational, Rotational, and Vibrational Entropy Contributions and Total $-T\Delta S$ Values (kcal/mol) for Protein-Protein (Rac1/PAK1) Complexes. Values Obtained from Three Replicates for Each Complex.

Protein-Ligand	Translational	Rotational	Vibrational	Total
Wild (GTP)-1	-14.6323	-14.0742	-26.4307	-55.1372
Wild (GTP)-2	-14.6323	-14.3091	-27.0195	-55.9610
Wild (GTP)-3	-14.6323	-14.2820	-28.4483	-57.3629
Wild (GDP)-1	-14.6326	-13.8643	-26.8058	-55.3027
Wild (GDP)-2	-14.6326	-14.1192	-26.0398	-54.7913
Wild (GDP)-3	-14.6326	-14.2393	-26.3890	-55.2606
F28L (GTP)-1	-14.6317	-14.0107	-29.0416	-57.6840
F28L (GTP)-2	-14.6317	-14.4782	-27.7616	-56.8715
F28L (GTP)-3	-14.6317	-14.3455	-26.3693	-55.3468
F28L (GDP)-1	-14.6311	-14.3240	-26.7035	-55.6586
F28L (GDP)-2	-14.6311	-14.4481	-28.4533	-57.5322
F28L (GDP)-3	-14.6311	-14.4394	-25.8350	-54.9052
P29S (GTP)-1	-14.6332	-14.3899	-28.1567	-57.1798
P29S (GTP)-2	-14.6332	-14.3825	-26.5553	-55.5710
P29S (GTP)-3	-14.6332	-14.4126	-28.2440	-57.2901
P29S (GDP)-1	-14.6326	-14.5673	-28.4784	-57.6786
P29S (GDP)-2	-14.6326	-14.4794	-27.7703	-56.8820
P29S (GDP)-3	-14.6326	-14.2486	-26.2482	-55.1294
Q61L (GTP)-1	-14.6323	-14.4299	-26.4659	-55.5284
Q61L (GTP)-2	-14.6323	-14.3503	-25.8773	-54.8596
Q61L (GTP)-3	-14.6323	-14.3720	-26.3502	-55.3545
Q61L (GDP)-1	-14.6326	-14.3670	-27.2411	-56.2406
Q61L (GDP)-2	-14.6326	-14.1654	-28.3448	-57.1428
Q61L (GDP)-3	-14.6326	-14.2868	-27.2059	-56.1252
Q61R (GTP)-1	-14.6323	-13.5360	-26.1138	-54.2824
Q61R (GTP)-2	-14.6323	-13.7614	-27.2429	-55.6366
Q61R (GTP)-3	-14.6323	-13.8243	-27.4537	-55.9100
Q61R (GDP)-1	-14.6320	-13.8425	-27.9229	-56.3975
Q61R (GDP)-2	-14.6320	-14.0679	-28.5994	-57.3000
Q61R (GDP)-3	-14.6320	-13.9633	-27.8928	-56.4878
T17N (GTP)-1	-14.6323	-14.3607	-26.5342	-55.5272
T17N (GTP)-2	-14.6323	-14.3950	-26.7718	-55.7991
T17N (GTP)-3	-14.6323	-14.3574	-27.6868	-56.6768
T17N (GDP)-1	-14.6326	-13.9972	-27.8594	-56.4893
T17N (GDP)-2	-14.6326	-14.3172	-27.5476	-56.4976
T17N (GDP)-3	-14.6326	-14.3392	-26.5172	-56.4976

Table 5.21: Entropy Contribution $-T\Delta S$ Values of the Groups (Rac1/PAK1), with Wild Type Mean and Comparative Analysis

Grup	**Mean Total $-T\Delta S$ (kcal/mol)**	**Wild Type Ortalama (kcal/mol)**	**Comparison and Commentary**
Wild (GTP)	-56.15 ± 1.12	-56.15	Reference group
Wild (GDP)	-55.78 ± 0.27	-55.78	Reference group
F28L (GTP)	-56.63 ± 1.42	-56.15	More negative → decrease in entropic stability (DIFFERENCE did not exceed standard deviation)
F28L (GDP)	-56.70 ± 1.37	-55.78	More negative → decrease in entropic stability (DIFFERENCE did not exceed standard deviation)
P29S (GTP)	-56.01 ± 1.00	-56.15	Similar values → no significant difference (DIFFERENCE did not exceed standard deviation)
P29S (GDP)	-56.56 ± 1.29	-55.78	More negative → decrease in entropic stability
Q61L (GTP)	-55.91 ± 0.35	-56.15	Similar values → no significant difference
Q61L (GDP)	-56.50 ± 0.48	-55.78	More negative → decrease in entropic stability
Q61R (GTP)	-55.94 ± 0.80	-56.15	Similar values → no significant difference
Q61R (GDP)	-56.73 ± 0.50	-55.78	More negative → decrease in entropic stability
T17N (GTP)	-56.00 ± 0.59	-56.15	Similar values → no significant difference
T17N (GDP)	-56.49 ± 0.03	-55.78	More negative → decrease in entropic stability

Entropy calculations were performed to carry out binding energy estimations, and the translational, rotational, and vibrational entropy energy components along with the total entropy values for each run are presented in Table 5.20. In addition to the GTP and GDP bound states, entropy contributions for the PAK1 protein were also calculated to better understand its role in the binding process. The averages of the obtained entropy values and their comparison against the wild type, together with related interpretations, are provided in Table 5.21.

Binding Free Energy Analysis for Rac1-PAK1 Complex

The MM/GBSA binding energy scores between Rac1 and PAK1 based on three replicate simulations for each case are summarized as follows:

- Wild-type Rac1: The GTP-bound form shows an average binding energy of approximately –21 kcal/mol, despite one replicate exhibiting relatively weak interaction. In the GDP-bound form, two replicates yield weak energies (around

- –4 and –2 kcal/mol), while one replicate presents a much stronger value of –40 kcal/mol. Excluding this outlier, the GDP-bound form demonstrates substantially weaker binding compared to the GTP-bound form (Table 5.22).
- P29S mutant: There is no significant difference between the GTP- and GDP- bound forms. However, the GTP-bound form exhibits slightly stronger binding. Additionally, the GDP-bound P29S binds better than the GDP-bound wild-type (Table 5.22).
- F28L mutant: The GTP-bound form displays variability, with one replicate showing poor binding (+39 kcal/mol) and another yielding weak binding (–4 kcal/mol). The GDP-bound form generally shows favorable binding, with an average of –21 kcal/mol, excluding one moderate value (–8 kcal/mol). This suggests enhanced binding compared to the GDP-bound wild-type (Table 5.22).
- Q61R mutant: All replicates yield positive binding energy values in both GTP- and GDP-bound forms, indicating a lack of interaction between Rac1 and PAK1 for this mutant (Table 5.22).
- Q61L mutant: No distinct difference is observed between the GTP- and GDP-bound forms. However, the GDP-bound form exhibits stronger binding relative to the GDP-bound wild-type (Table 5.22).
- T17N mutant: A clear distinction is observed between the GTP- and GDP-bound forms. The GTP-bound form resembles the wild-type, while the GDP bound form shows reduced binding affinity in comparison to the GDP-bound wild-type (Table 5.22)

Table 5.22: Binding free energy values and averages of PAK1 obtained from three replicate runs

Grup	Run 1 (kcal/mol)	Run 2 (kcal/mol)	Run 3 (kcal/mol)	Mean (kcal/mol)
Wild GTP	-17.1377	-38.4841	-7.5452	-21.0557
Wild GDP	-4.8579	-40.6910	-2.7747	-16.1079
P29S GTP	-21.6897	-34.5136	-13.3645	-23.1893
P29S GDP	-12.2450	-27.6195	-23.1686	-21.0110
F28L GTP	39.0345	-4.7923	-23.2784	3.6546
F28L GDP	-8.0689	-26.0684	-30.1984	-21.4452
Q61R GTP	40.4703	30.9191	18.3327	29.9074
Q61R GDP	18.7849	40.9211	33.4745	31.0602
Q61L GTP	-22.3464	-37.5838	-29.6724	-29.8675
Q61L GDP	-40.6721	-15.9392	-36.7172	-31.1095
T17N GTP	-27.1421	-32.5240	-6.1337	-21.9333
T17N GDP	-7.4680	-11.5425	-7.8446	-8.9517

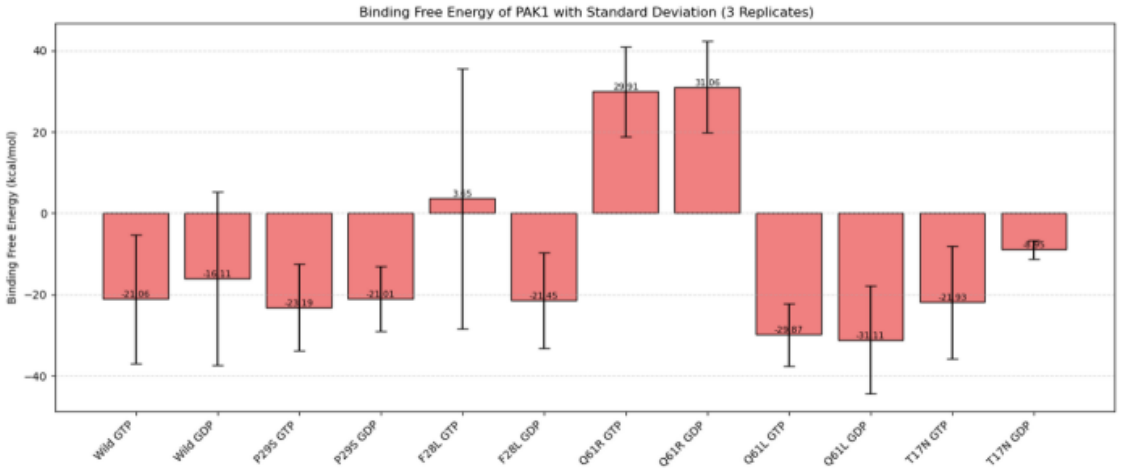


Figure 5.19: Binding free energy graph for PAK1. Binding free energy values (in kcal/mol) of PAK1 in complex with Rac1-bound GTP or GDP, based on three replicate simulations. Mean binding energies are shown as bars, with standard deviations indicated by error bars. The impact of Rac1 mutations on the binding affinity of PAK1 is visualized across different nucleotide states

MM/GBSA Energy Analysis for Rac1-TIPE2 Complex

When examining the MM/GBSA energy data for the interaction between Rac1 (GTP) and TIPE2, although the values vary significantly, energy values averaging between -20 and -40 kcal/mol were observed towards the end (Figure 5.20). In particular, Run 2 showed slightly lower energy values compared to the other GTP bound runs. Run 3, on the other hand, exhibited high energy values for a long period until it stabilized again after dissociation. For the interaction between Rac1 (GDP) and TIPE2, the energy values generally ranged between 0 kcal/mol (including frames with occasional positive values) and -60 kcal/mol (Figure 5.20). Overall, Run 2 showed lower energy compared to the other runs and was the run in which no dissociation event was observed based on the conformation.

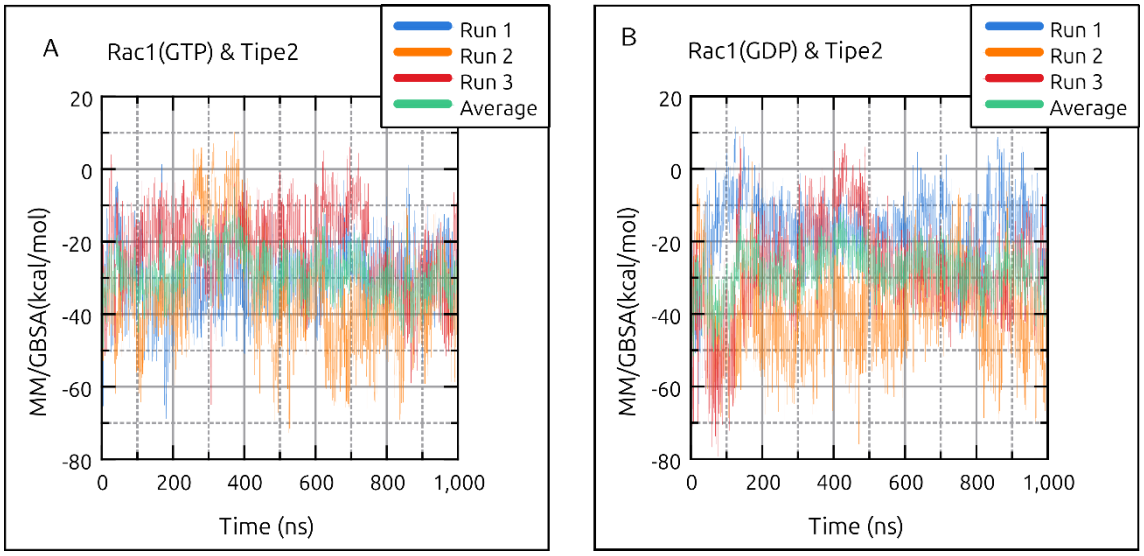


Figure 5.20: MM/GBSA Values for Rac1-TIPE2 Complex

Entropy Analysis for Rac1-TIPE2 Complex

Table 5.23 shows the entropy values obtained from a total of 6 runs between Rac1 (GDP/GTP) and TIPE2

Table 5.23: Translational, Rotational, and Vibrational Entropy Contributions and Total $-T\Delta S$ Values (kcal/mol) for Rac1 TIPE2 Complexes. Values Obtained from Three Replicates for Each Complex.

Protein-Ligand	**Translational**	**Rotational**	**Vibrational**	**Total**
Rac1(GTP) and TIPE2-1	-15.8568	-15.5038	-30.1695	-61.5295
Rac1(GTP) and TIPE2-2	-15.8568	-15.5408	-30.2402	-61.6374
Rac1(GTP) and TIPE2-3	-15.8568	-15.3097	-28.9838	-60.1503
Rac1(GDP) and TIPE2-1	-15.8574	-15.4164	-29.8374	-61.1112
Rac1(GDP) and TIPE2-2	-15.8574	-15.5438	-30.0610	-61.4624
Rac1(GDP) and TIPE2-3	-15.8574	-15.4152	-30.3588	-61.6306

Binding Free Energy Analysis for Rac1-TIPE2 Complex

The binding energy scores (without considering entropy) obtained from the MD runs are shown in Table 5.24. According to the binding energy data, no negative scores were observed for either GTP- or GDP-bound repeats. For the GTP-bound repeats, Run 2 showed a slightly better result compared to the other runs, while for the GDP-bound runs, Run 2 exhibited a lower binding energy compared to the other GDP-bound runs.

Table 5.24: Binding energy data obtained from MD simulations of Rac1 (GTP and GDP) with TIPE2

Grup	Run 1 (kcal/mol)	Run 2 (kcal/mol)	Run 3 (kcal/mol)	Mean (kcal/mol)
Rac1(GTP)- TIPE2	28.4191	23.7240	43.0669	31.7367
Rac1(GDP)- TIPE2	49.0880	6.0325	26.4648	27.1951

DISCUSSION AND CONCLUSION

Simulations based on Rac1 have yielded numerous meaningful results for both TIPE2 and the Rac1-PAK1 interaction, including various mutant Rac1-PAK1 complexes. Among these findings, one particularly notable observation is the substantial conformational changes depending on whether GDP or GTP is bound. Below, the implications of several mutations and the results obtained in this context are presented in order.

For the F28L mutation, an increase in binding energy to PAK1 was observed in the GDP-bound form, which was evaluated alongside increased flexibility in the Switch I region. These findings suggest that the F28L mutation may enhance the binding capability of GDP-bound Rac1—which is typically considered to be in a ″closed″ or ″inactive″ conformation—to PAK1. This observation challenges the traditional notion that GDP-bound Rac1 is largely incapable of engaging in productive interactions with downstream effectors. The F28L mutant, by increasing local flexibility, might be facilitating transient conformational states that are more favorable for binding, thereby enabling a non-canonical interaction mechanism with PAK1. Such a mechanism may have significant implications for understanding the functional plasticity of Rac1 in different signaling contexts.

In the case of the Q61L mutation, a strengthened binding to PAK1 in the GDP- bound form was also identified. Q61L is a well-known ″activating″ mutation, and the observed increase in binding energy suggests that this mutation may promote interaction with PAK1, potentially leading to sustained cellular signal transduction and consequently increased cell proliferation or motility. Given the oncogenic relevance of Q61L in various cancers, its ability to promote effector binding even in the GDP bound state may reflect an altered nucleotide cycling behavior, where the mutant protein remains in a signaling-competent conformation for extended periods. This reinforces the importance of evaluating such mutants not only in the GTP-bound form but also considering their GDP-associated dynamics and interactions.

In contrast, the Q61R mutation resulted in decreased binding energy in both GDP- and GTP-bound forms, indicating a weakened interaction. Moreover, increased flexibility in the Switch I region was recorded, which may be related to conformational variability at the binding interface and may thus hinder stable binding. These results indicate that the Q61R mutation may impair Rac1-PAK1

interaction and cause disruptions in cellular signaling pathways, leading to functionally detrimental effects. The observed loss of interaction could be attributed to the arginine side chain introducing steric hindrance or electrostatic repulsion at critical interaction sites, thereby destabilizing the Rac1-PAK1 complex. Additionally, the heightened flexibility in Switch I may prevent the adoption of the structural arrangement required for effective effector engagement, thereby emphasizing the delicate balance between conformational rigidity and plasticity in protein-protein interactions.

Overall, the assessment of protein dynamics revealed that the flexibility of the Switch I region plays a critical role in PAK1 binding. Increased flexibility in this region is generally associated with reduced binding affinity (as observed particularly in the Q61R mutant); however, the F28L mutant represents an exception where increased flexibility still coincides with enhanced PAK1 binding. In contrast, more stable and less flexible conformations tend to promote stronger binding. This duality suggests that while excessive flexibility can disrupt critical contacts necessary for binding, a moderate degree of flexibility might enable adaptive recognition, particularly in mutants that alter the local energy landscape in a favorable manner. Such nuanced effects underline the importance of considering dynamic ensembles, rather than static structures, when interpreting the functional consequences of mutations.

Additionally, the nucleotide-bound state (GTP vs. GDP) serves as a key factor in determining the effects of mutations. Mutations alter binding energies and structural stability in different ways depending on the nucleotide state. While some mutations (F28L, Q61L) enhance binding, others (Q61R) weaken it, thereby leading to distinct consequences in cellular signaling pathways. This highlights the critical interplay between nucleotide state and structural dynamics in modulating effector recognition. Furthermore, it supports the notion that mutations can differentially influence Rac1's interaction network depending on its activation state, which may contribute to altered signaling specificity or pathway rewiring in disease contexts.

In another part of this thesis focusing on the interaction between Rac1 and TIPE2, it was concluded that further investigation is necessary. The analysis of TIPE2 simulations showed that while system stability was preserved under certain conditions, structural shifts occurred in others. Among the simulations displaying structural shifts, the second study performed in the GTP-bound form demonstrated geometrically promising conformations. However, when entropy contributions were considered, all six simulations yielded positive binding free

energies, indicating that the interaction is thermodynamically unfavorable. This suggests that although enthalpic contributions may occasionally support transient association, the overall entropy-driven cost outweighs potential binding benefits under the simulated conditions. Such results imply that Rac1-TIPE2 binding might be highly context-dependent, potentially requiring specific post-translational modifications or additional co-factors in vivo to stabilize the interaction. Moreover, the observed structural shifts could reflect a conformational sampling behavior that, while unfavorable in terms of free energy, might still hold regulatory relevance under certain cellular environments

These findings suggest that although some conformations may hold potential functional significance, there is not yet sufficient or conclusive evidence regarding the overall binding stability. Further refinement of simulation protocols, extended sampling times, or the inclusion of experimentally derived restraints may provide deeper insights into the nature of this interaction. Such approaches could help elucidate whether Rac1-TIPE2 engagement occurs under specific physiological stimuli or within particular subcellular compartments.

FUTURE WORK

Understanding the conformational changes induced by mutations in Rac1 is crucial for developing strategies to reverse their effects. Particularly, focusing on the differences observed between the wild-type and mutant forms in both GTP- and GDP-bound states—such as alterations present in the wild type but absent or exaggerated in mutants—can provide valuable insights relevant to cancer and immunerelated diseases.

Future studies should involve detailed residue-residue network and communication pathway analyses to identify disrupted or aberrant signaling routes caused by mutations. These altered interaction networks could serve as targets for designing small-molecule therapeutics aimed at restoring normal signaling by stabilizing or compensating for the disrupted pathways.

Similarly, mapping residue-residue interactions in the Rac1-PAK1 complex may facilitate understanding how mutations affect binding and allow the design of interventions to restore mutant behavior toward that of the wild type

Regarding the interaction between Rac1 and TIPE2, the intracellular environment presents a highly complex and dynamic context. The central cavity of TIPE2, characterized by sufficient size and favorable physicochemical properties, appears capable of binding various small molecules and peptides. Targeting this cavity with stabilizing ligands could enhance TIPE2's structural stability and, consequently, its interaction with Rac1.

Investigating such molecular modulators and their effects on the Rac1-TIPE2 interaction may yield novel insights into the regulation of this complex and un- cover potential therapeutic avenues for modulating signaling pathways implicated in disease.

BIBLIOGRAPHY

[Acuner et al., 2021] Acuner, S. E., Sumbul, F., Torun, H., and Haliloglu, T. (2021). Oncogenic mutations on rac1 affect global intrinsic dynamics underlying gtp and pak1 binding. Biophysical journal, 120(5):866–876.

[Bokoch, 2003] Bokoch, G. M. (2003). Biology of the p21-activated kinases. Annual review of biochemistry, 72(1):743–781.

[Bordoloi et al., 2018] Bordoloi, D., Banik, K., Shabnam, B., Padmavathi, G., Monisha, J., Arfuso, F., Dharmarajan, A., Mao, X., Lim, L. H., Wang, L., et al. (2018). Tipe family of proteins and its implications in different chronic diseases. International journal of molecular sciences, 19(10):2974.

[Cao et al., 2013a] Cao, X., Zhang, L., Shi, Y., Sun, Y., Dai, S., Guo, C., Zhu, F., Wang, Q., Wang, J., Wang, X., et al. (2013a). Human tumor necrosis factor (tnf)-alpha-induced protein 8-like 2 suppresses hepatocellular carcinoma metastasis through inhibiting rac1. Molecular cancer, 12:1–10.

[Cao et al., 2013b] Cao, X., Zhang, L., Shi, Y., Sun, Y., Dai, S., Guo, C., Zhu, F., Wang, Q., Wang, J., Wang, X., et al. (2013b). Human tumor necrosis factor (tnf)-alpha-induced protein 8-like 2 suppresses hepatocellular carcinoma metastasis through inhibiting rac1. Molecular cancer, 12:1–10.

[Chaudhury et al., 2011] Chaudhury, S., Berrondo, M., Weitzner, B. D., Muthu, P., Bergman, H., and Gray, J. J. (2011). Benchmarking and analysis of protein docking performance in rosetta v3. 2. PloS one, 6(8):e22477.

[Chaudhury and Gray, 2008] Chaudhury, S. and Gray, J. J. (2008). Conformer selection and induced fit in flexible backbone protein–protein docking using computational and nmr ensembles. Journal of molecular biology, 381(4):1068–1087.

[Ching et al., 2003] Ching, Y.-P., Leong, V. Y., Wong, C.-M., and Kung, H.-F. (2003). Identification of an autoinhibitory domain of p21-activated protein kinase 5. Journal of Biological Chemistry, 278(36):33621–33624.

[Conway et al., 2014] Conway, P., Tyka, M. D., DiMaio, F., Konerding, D. E., and Baker, D. (2014). Relaxation of backbone bond geometry improves protein energy landscape modeling. Protein science, 23(1):47–55.

[Davis et al., 2013] Davis, M. J., Ha, B. H., Holman, E. C., Halaban, R., Schlessinger, J., and Boggon, T. J. (2013). Rac1p29s is a spontaneously activating cancer-associated gtpase. Proceedings of the National Academy of Sciences, 110(3):912–917.

[Ferrandez et al., 2017] Ferrandez, Y., Zhang, W., Peurois, F., Akendengu´e, L., Blangy, A., Zeghouf, M., and Cherfils, J. (2017). Allosteric inhibition of the guanine nucleotide exchange factor dock5 by a small molecule. Scientific reports, 7(1):14409.

[Gray et al., 2003] Gray, J. J., Moughon, S., Wang, C., Schueler-Furman, O., Kuhlman, B., Rohl, C. A., and Baker, D. (2003). Protein–protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. Journal of molecular biology, 331(1):281–299.

[Gu et al., 2020] Gu, Z., Cui, X., Sun, P., and Wang, X. (2020). Regulatory roles of tumor necrosis factor-α-induced protein 8 like-protein 2 in inflammation, immunity and cancers: A review. Cancer management and research, pages 12735–12746.

[Ha et al., 2015] Ha, B. H., Morse, E. M., Turk, B. E., and Boggon, T. J. (2015). Signaling, regulation, and specificity of the type ii p21-activated kinases. Journal of Biological Chemistry, 290(21):12975–12983.

[Hendlich et al., 1997] Hendlich, M., Rippmann, F., and Barnickel, G. (1997). Ligsite: automatic and efficient detection of potential small molecule-binding sites in proteins. Journal of Molecular Graphics and Modelling, 15(6):359–363.

[Honorato et al., 2021] Honorato, R. V., Koukos, P. I., Jim´enez-Garc´ıa, B., Tsaregorodtsev, A., Verlato, M., Giachetti, A., Rosato, A., and Bonvin, A. M. (2021). Structural biology in the clouds: the wenmr-eosc ecosystem. Frontiers in molecular biosciences, 8:729513.

[Honorato et al., 2024] Honorato, R. V., Trellet, M. E., Jim´enez-Garc´ıa, B., Schaarschmidt, J. J., Giulini, M., Reys, V., Koukos, P. I., Rodrigues, J. P., Karaca, E., van Zundert, G. C., et al. (2024). The haddock2. 4 web server for integrative modeling of biomolecular complexes. Nature protocols, 19(11):32193241.

[Hopkins et al., 2015] Hopkins, C. W., Le Grand, S., Walker, R. C., and Roitberg, A. E. (2015). Long-time-step molecular dynamics through hydrogen mass repartitioning. Journal of chemical theory and computation, 11(4):1864–1874.

[Khatib et al., 2011] Khatib, F., Cooper, S., Tyka, M. D., Xu, K., Makedon, I., Popovi´c, Z., Baker, D., and Players, F. (2011). Algorithm discovery by protein folding game players. Proceedings of the National Academy of Sciences, 108(47):18949–18953.

[Krauthammer et al., 2012] Krauthammer, M., Kong, Y., Ha, B. H., Evans, P., Bacchiocchi, A., McCusker, J. P., Cheng, E., Davis, M. J., Goh, G., Choi, M., et al. (2012). Exome sequencing identifies recurrent somatic rac1 mutations in melanoma. Nature genetics, 44(9):1006–1014.

[Kumar et al., 2013] Kumar, A., Rajendran, V., Sethumadhavan, R., and Purohit, R. (2013). Molecular dynamic simulation reveals damaging impact of rac1 f28l mutation in the switch i region. PloS one, 8(10):e77453.

[Kumar et al., 2006] Kumar, R., Gururaj, A. E., and Barnes, C. J. (2006). p21-activated kinases in cancer. Nature Reviews Cancer, 6(6):459–471.

[Laskowski, 2007] Laskowski, R. A. (2007). Enhancing the functional annotation of pdb structures in pdbsum using key figures extracted from the literature. Bioinformatics, 23(14):1824–1827.

[Li and Merz Jr, 2014] Li, P. and Merz Jr, K. M. (2014). Taking into account the ion-induced dipole interaction in the nonbonded model of ions. Journal of chemical theory and computation, 10(1):289–297.

[Li et al., 2016] Li, Z., Guo, C., Liu, X., Zhou, C., Zhu, F., Wang, X., Wang, Q., Shi, Y., Wang, J., Zhao, W., et al. (2016). Tipe2 suppresses angiogenesis and non-small cell lung cancer (nsclc) invasiveness via inhibiting rac1 activation and vegf expression. Oncotarget, 7(38):62224.

[Liu et al., 2021] Liu, H., Liu, K., and Dong, Z. (2021). The role of p21-activated kinases in cancer and beyond: where are we heading? Frontiers in cell and developmental biology, 9:641381.

[Ma et al., 2023] Ma, N., Xu, E., Luo, Q., and Song, G. (2023). Rac1: a regulator of cell migration and a potential target for cancer therapy. Molecules, 28(7):2976.

[Maguire et al., 2021] Maguire, J. B., Haddox, H. K., Strickland, D., Halabiya, S. F., Coventry, B., Griffin, J. R., Pulavarti, S. V. K., Cummins, M., Thieker, D. F., Klavins, E., et al. (2021). Perturbing the energy landscape for improved packing during computational protein design. Proteins: Structure, Function, and Bioinformatics, 89(4):436–449.

[Maier et al., 2015] Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015). ff14sb: improving the accuracy of protein side chain and backbone parameters from ff99sb. Journal of chemical theory and computation, 11(8):3696–3713.

[Marze et al., 2017] Marze, N. A., Jeliazkov, J. R., Roy Burman, S. S., Boyken, S. E., DiMaio, F., and Gray, J. J. (2017). Modeling oblong proteins and water-mediated interfaces with rosettadock in capri rounds 28–35. Proteins: Structure, Function, and Bioinformatics, 85(3):479–486.

[Marze et al., 2018] Marze, N. A., Roy Burman, S. S., Sheffler, W., and Gray, J. J. (2018). Efficient flexible backbone protein–protein docking for challenging targets. Bioinformatics, 34(20):3461–3469.

[Meagher et al., 2003] Meagher, K. L., Redman, L. T., and Carlson, H. A. (2003). Development of polyphosphate parameters for use with the amber force field. Journal of computational chemistry, 24(9):1016–1025.

[Miller III et al., 2012] Miller III, B. R., McGee Jr, T. D., Swails, J. M., Homeyer, N., Gohlke, H., and Roitberg, A. E. (2012). Mmpbsa. py: an efficient program for end-state free energy calculations. Journal of chemical theory and computation, 8(9):3314–3321.

[Molli et al., 2009] Molli, P. R., Li, D.-Q., Murray, B., Rayala, S. K., and Kumar, R. (2009). Pak signaling in oncogenesis. Oncogene, 28(28):2545–2555.

[Niv´on et al., 2013] Niv´on, L. G., Moretti, R., and Baker, D. (2013). A pareto-optimal refinement method for protein design scaffolds. PloS one, 8(4):e59004.

[Oxford and Theodorescu, 2003] Oxford, G. and Theodorescu, D. (2003). Ras superfamily monomeric g proteins in carcinoma cell motility. Cancer letters, 189(2):117–128.

[Rane and Minden, 2014] Rane, C. K. and Minden, A. (2014). P21 activated kinases: structure, regulation, and functions. Small GTPases, 5(1):e28003.

[Roe and Cheatham III, 2013] Roe, D. R. and Cheatham III, T. E. (2013). Ptraj and cpptraj: software for processing and analysis of molecular dynamics trajectory data. Journal of chemical theory and computation, 9(7):3084–3095.

[Rojas et al., 2012] Rojas, A. M., Fuentes, G., Rausell, A., and Valencia, A. (2012). The ras protein superfamily: evolutionary tree and role of conserved amino acids. Journal of Cell Biology, 196(2):189–201.

[Saraste et al., 1990] Saraste, M., Sibbald, P. R., and Wittinghofer, A. (1990). The

p-loop—a common motif in atp-and gtp-binding proteins. Trends in biochemical sciences, 15(11):430–434.

[Schrödinger, LLC, 2015a] Schr¨odinger, LLC (2015a). The AxPyMOL molecular graphics plugin for Microsoft PowerPoint, version 1.8.

[Schrödinger, LLC, 2015b] Schr¨odinger, LLC (2015b). The PyMOL molecular graphics system, version 1.8.

[Sells et al., 1997] Sells, M. A., Knaus, U. G., Bagrodia, S., Ambrose, D. M., Bokoch, G. M., and Chernoff, J. (1997). Human p21-activated kinase (pak1) regulates actin organization in mammalian cells. Current biology, 7(3):202–210.

[Steinkellner et al., 2009] Steinkellner, G., Rader, R., Thallinger, G. G., Kratky, C., and Gruber, K. (2009). Vasco: computation and visualization of annotated protein surface contacts. BMC bioinformatics, 10:1–11.

[Sun et al., 2008] Sun, H., Gong, S., Carmody, R. J., Hilliard, A., Li, L., Sun, J., Kong, L., Xu, L., Hilliard, B., Hu, S., et al. (2008). Tipe2, a negative regulator of innate and adaptive immunity that maintains immune homeostasis. Cell, 133(3):415–426.

[Szklarczyk et al., 2023] Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A. L., Fang, T., Doncheva, N. T., Pyysalo, S., et al. (2023). The string database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic acids research, 51(D1):D638–D646.

[Tanabe et al., 2000] Tanabe, K., Tachibana, T., Yamashita, T., Che, Y. H., Yoneda, Y., Ochi, T., Tohyama, M., Yoshikawa, H., and Kiyama, H. (2000). The small gtp-binding protein tc10 promotes nerve elongation in neuronal cells, and its expression is induced during nerve regeneration in rats. Journal of Neuroscience, 20(11):4138–4144.

[Tyka et al., 2011] Tyka, M. D., Keedy, D. A., Andr´e, I., DiMaio, F., Song, Y., Richardson, D. C., Richardson, J. S., and Baker, D. (2011). Alternate states of proteins revealed by detailed energy landscape mapping. Journal of molecular biology, 405(2):607–618.

[Ugochukwu et al., 2007] Ugochukwu, E., Yang, X., Elkins, J., Burgess-Brown, N., Bunkoczi, G., Knapp, S., and Doyle, D. (2007). Crystal structure of human rac3 in

complex with crib domain of human p21-activated kinase 1 (pak1). The Protein Data Bank.

[Vetter, 2014] Vetter, I. R. (2014). The structure of the g domain of the ras superfamily. Ras superfamily small G proteins: biology and mechanisms 1: general features, signaling, pages 25–50.

[Wang et al., 2007] Wang, C., Bradley, P., and Baker, D. (2007). Protein–protein docking with backbone flexibility. Journal of molecular biology, 373(2):503–519.

[Wang et al., 2005] Wang, C., Schueler-Furman, O., and Baker, D. (2005). Improved side-chain modeling for protein–protein docking. Protein Science, 14(5):1328–1339.

[Wang et al., 2019] Wang, E., Sun, H., Wang, J., Wang, Z., Liu, H., Zhang, J. Z., and Hou, T. (2019). End-point binding free energy calculation with mm/pbsa and mm/gbsa: strategies and applications in drug design. Chemical reviews, 119(16):9478–9508.

[Wang et al., 2011] Wang, J., Wu, J.-W., and Wang, Z.-X. (2011). Structural insights into the autoactivation mechanism of p21-activated protein kinase. Structure, 19(12):1752–1761.

[Weng et al., 2019] Weng, G., Wang, E., Wang, Z., Liu, H., Zhu, F., Li, D., and Hou, T. (2019). Hawkdock: a web server to predict and analyze the protein–protein complex based on computational docking and mm/gbsa. Nucleic acids research, 47(W1):W322–W330.

[Xu et al., 2024] Xu, L., Pan, F., and Guo, Z. (2024). Tipe2: a candidate for targeting antitumor immunotherapy. The Journal of Immunology, 212(5):755–763.

[Yao et al., 2020] Yao, D., Li, C., Rajoka, M. S. R., He, Z., Huang, J., Wang, J., and Zhang, J. (2020). P21-activated kinase 1: Emerging biological functions and potential therapeutic targets in cancer. Theranostics, 10(21):9741.

[Zhang et al., 2009] Zhang, X., Wang, J., Fan, C., Li, H., Sun, H., Gong, S., Chen, Y. H., and Shi, Y. (2009). Crystal structure of tipe2 provides insights into immüne homeostasis. Nature structural & molecular biology, 16(1):89–90.