

β_2 アドレナリン受容体のコ ミュニティ検出によるシグ ナル伝達機構の解析

名古屋大学 理学研究科 理学専攻
生物物理分野

学籍番号 262301298
氏名 齊藤 詩恋

指導教員: 倭剛久 准教授

第 1 章

序論

1.1 はじめに

1.1.1 アロステリー

タンパク質のアロステリー現象は、リガンド結合や外部刺激によって生じる構造変化が刺激受容部位から遠隔の活性部位に影響を及ぼす現象であり、そのメカニズム解明は生命現象の理解と創薬研究の中心的課題の一つである。アロステリーはタンパク質の機能を制御する重要な特性であり [1]、リガンド結合など外部刺激のシグナルが残基間相互作用を介してサブ Å ~ 数十 Å 離れた活性部位に伝達する仕組みを提供する。

1.1.2 グラフ理論

アロステリーの離れた場所間のコミュニケーションを解明する方法の一つとして、グラフ理論 [2] を用いたアプローチが注目されている。グラフ理論は、分子内の残基間の相互作用をネットワークとして表現し、複雑な動的挙動を解析するための強力なツールとして広く利用されてきた [3][4][5]。グラフ理論では、タンパク質のアミノ酸残基の相互作用を、ノードとエッジを用いたアミノ酸ネットワークとして表すことができる。

- ノード：アミノ酸残基（原子やタンパク質全体も可能）

- エッジ：ノード同士の相互作用

そしてエッジに重みづけをすることで、エッジを同じ距離 1 とするのではなく、各結合ごとに強度を振り分けることが可能である。

タンパク質間の相互作用ネットワークをグラフとしてモデル化することで、シグナル伝達機構を解釈しやすくなり、生命現象の解明や新薬の開発に役立っている。そのようなモデルは、ネットワーク内の最短経路の存在の重要性を強調 [6] しており、それらが離れた部位間での効率的な情報伝達に寄与していることを示している。また、その前提のもとで、高い中心性を持つ残基 [7] やネットワーク上における最短経路上によく現れる残基 [8] を解析することで、機能的残基を同定してきた。実際にこれらの残基はタンパク質の折りたたみにおける重要なアミノ酸や酵素ファミリーの活性部位残基であると関連付けられている。

1.1.3 本研究の位置づけ

ここまで、シグナルはある「経路」に沿って伝達するという仮説に基づいた解析について述べてきた。しかし、単純な「最短経路モデル」だけでは、アロステリーの情報伝達を十分に説明できない可能性がある。ピコ秒オーダーの残基間エネルギー移動過程 [9] とミリ秒以上のアロステリック遷移 [10] の時間スケールの違いを考慮すると、リガンド結合部位から活性部位まで情報が伝言ゲームのように最短経路上を通ると考えるにはあまりに単純すぎるのではないかと考えられるからだ。

実際にそれぞれのドメインの自由エネルギー的な安定性のみが問題であり、アロステリーは特定のエネルギー伝達の経路を必要としないと言及する先行研究 [11][12] が存在する。つまりこのモデルは再考の必要がある。

そうしたことを踏まえて、T. Ishikura, Y. Iwata, T. Hatano, T. Yamato は、アミノ酸間のエネルギーのやりとりの強弱をエネルギー交換ネットワークとして表現する Energy Exchange Network Model(EEN モデル) を確立 [13] した。このモデルを適用して PDZ ドメインの C 末端の α ヘリックスの除去前後で比較すると、 α ヘリックスがリガンド結合部位と他の領域のつ

ながりを保つ役割を果たしていたことが明らかになり、これは α ヘリックスの除去がリガンド親和性を 21 倍減少させる先行研究の見解と一致していた。つまり EEN の変化と機能の変化を結びつける画期的なモデルであることが証明された。

そこで本研究では、EEN モデルでも使用している「残基相互作用ネットワークの再編」の概念をベースにし、 β_2 アドレナリン受容体 (β_2 AR) という膜タンパク質の活性化による残基相互作用ネットワークの変化を比較することで、 β_2 AR のシグナル伝達機構を解析することを目的とする。

EEN モデルでは「残基」レベルのコミュニケーションに着目していたが、本研究では「コミュニティ (残基集団)」レベルの中間的スケールでの再編成に着目する。アロステリーのタンパク質全体のグローバルな動的性質と、局所的な残基間の相互作用の両方を効果的に取り入れることができるからだ。

また分子動力学 (MD) シミュレーションを用いて得られた inactive 状態と active 状態におけるトラジェクトリー解析から得られた平均構造を基に、ネットワークのエッジの重みとして用いる残基間最短距離の 2 乗逆数の平均 $\langle \frac{1}{d^2} \rangle$ を反映したネットワークを構築する。構造、ダイナミクス、相互作用を反映する物理的変数である残基間熱流の自己相関関数に基づく残基間熱伝導率の値を決定する上で残基間の最短距離が支配的な役割を果たしていることがわかっている [14] ため、この変数は残基間の挙動を十分に反映していると言える。また熱揺らぎなど分子の動的な挙動の性質を反映している平均構造を用いることで、より現実的なタンパク質の挙動を捉えることを可能する。

1.2 β_2 アドレナリン受容体 (β_2 AR)

基本情報

β_2 アドレナリン受容体 (β_2 AR) は、G タンパク質共役受容体 (GPCR) の一種である。GPCR は細胞膜に存在する膜タンパク質であり、ホルモンや神経伝達物質などの細胞外刺激を認識し、それを細胞内のシグナルに変換する役割を担っている。また、視覚、嗅覚、味覚といった感覚にも関与し、生体内の多様なシグナル伝達経路において中心的な役割を果たしている。GPCR

は7回膜貫通構造を持つことが特徴であり、各膜貫通ヘリックス (TM1 から TM7) は、細胞外ループ (ECL) と細胞内ループ (ICL) を介して他のヘリックスと連結されている。この構造により、細胞外でのリガンド認識と細胞内でのシグナル伝達を効率的に行うことが可能となる。

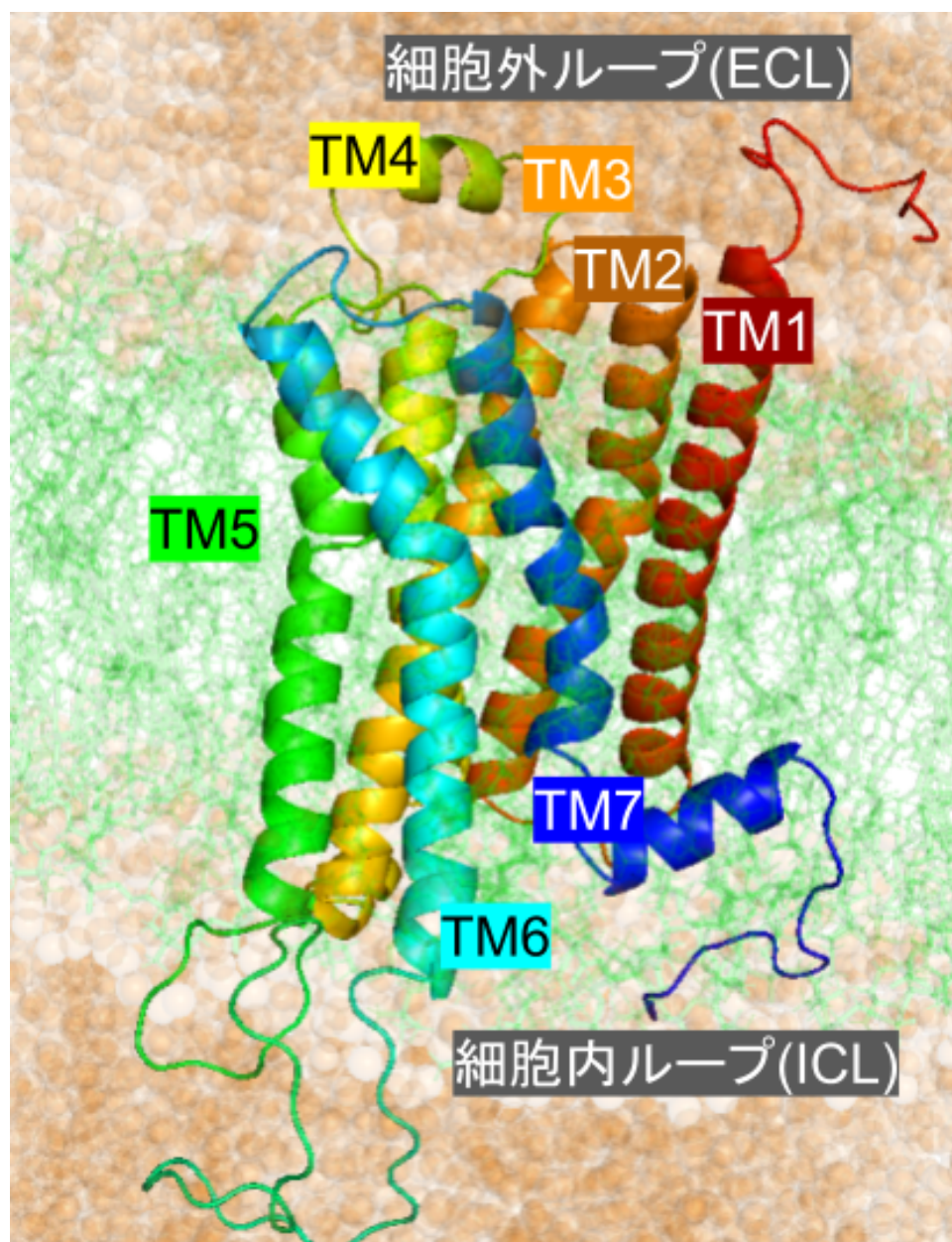


図 1.1: β_2 AR の膜貫通ヘリックス。

膜タンパク質の構造データは Protein Data Bank(PDB) から取得し、 β_2 AR の不活性状態として 2RH1[15], 活性状態として 3P0G[16] を用いる。

活性化に伴う構造的変化

β_2 AR の inactive 状態から active 状態への活性化は、リガンド結合部位に結合したアゴニストによって引き起こされる。すると受容体が活性化され、特に Gs といった G タンパク質の結合が促進される。Gs はヘテロ三量体型タンパク質であり、その活性化により細胞内でアデニル酸シクラーゼが刺激され、cAMP(サイクリック AMP) の産生が増加する。この過程 [17] を通じて、細胞内のシグナル伝達経路が活性化される。

β_2 AR のリガンド結合部位は膜貫通ヘリックス (TM) の間に位置し、細胞外の刺激を感知する。一方で、G タンパク質との結合は細胞内ループおよび細胞質側ドメインを介して行われる。

活性化に関わる重要な残基

β_2 AR の活性化において、特定のモチーフ [18][19] が重要な役割を果たしていることが知られている。

モチーフとは、タンパク質中の特定の機能に関連する保存されたアミノ酸配列のことであり、GPCR ではアロステリックシグナル伝達やコンフォメーション変化を介して受容体の機能を調節する。 β_2 AR を含むクラス A GPCR では、以下の 4 つの主要な保存モチーフ (DRY、CWxP、NPxxY、PIF) と「イオンロック」が重要な役割を果たしている。

- **DRY** モチーフ TM3 の細胞質側領域に位置し、Asp(D)-Arg(R)-Tyr(Y) から構成される。inactive 状態では TM6 の Glu とイオンロックを形成し、構造を安定化させている。一方、active 状態ではイオンロックが解除され、G タンパク質との結合が可能になる。
- **CWxP** モチーフ TM6 のリガンド結合ポケットの底部に位置し、Cys(C)-Trp(W)-任意の残基 (x)-Pro(P) から構成される。inactive 状態では Trp がリガンド結合ポケットを開いた状態を維持しているが、active 状態では Trp が「トグルスイッチ」として機能し、ポケットを

閉じる役割を果たす。

- **NPxxY** モチーフ TM7 の細胞質側に位置し、Asn(N)-Pro(P)-任意の残基 (xx)-Tyr(Y) から構成される。このモチーフにおける Tyr は、不活性状態と活性状態で異なる立体配座間を回転することで、構造の変化を媒介する。
- **PIF** モチーフ TM4、TM5、TM6 にまたがる位置に存在し、Pro(P)-Ile(I)-Phe(F) から構成される。このモチーフにおける Phe はスイッチとして機能し、活性化時に異なる立体配座間で方向を変えることで重要な役割を果たす。
- **イオンロック** TM3 と TM6 の間に位置している。不活性状態ではイオンロックが形成され、構造を安定化させているが、活性化時にはこのロックが解除され、GPCR の完全な活性化を促進する。

また水分子も、生物学的システムにおいて重要な役割を果たすことが知られており、特に GPCR の活性化メカニズムにも深く関与している。脂質二重層の誘電率は 2 から 5 程度であり水と比べると非常に低いため、電荷間の静電相互作用が水環境に比べてはるかに強くなる。すると脂質環境下では荷電残基が重要な役割を果たしており、それが近傍の水やイオンと強く相互作用する。

実際にロドプシンをはじめとする GPCR で高度に保存された水は、活性化過程においてアロステリックを仲介する機能を果たすことが示されている [20]。β₂AR においても、膜貫通ドメイン内にはいくつかの保存された水分子が確認され、これらの水分子はアロステリックシグナル伝達に関与していると考えられる。これらの水分子はアミノ酸残基と同じくらい受容体機能に重要であることが示されている。

そこで本研究でも、保存された結晶水をアミノ酸残基と同等の扱いとして露に取り入れることで、より現実的なタンパク質の動向を表現することを可能にした。

第 2 章

材料と方法

2.1 分子動力学シミュレーション

本研究では、大きく分けてモデリング、構造最適化、熱平衡化、サンプリング、解析の 5 つのステップを経た。

2.1.1 モデリング

モデリングでは、欠損部位の補完、結晶水の同定、膜タンパク質の構造準備の 3 つのステップを経た。

欠損部位の補完

PDB に登録されている膜タンパク質の構造データは、全ての残基位置が定まっているわけではない。なぜなら X 線結晶構造解析において、揺らぎが大きい部位は見えないからである。 β_2 AR では、MODELER ソフトウェアを用いて細胞内ループ 3(ICL3)、N 末端、C 末端の欠損部位を補完した。MODELER ではホモロジーモデリングを利用して、既存の PDB データから欠損部位の座標を予測した。

結晶水の同定

膜タンパク質内の結晶水が重要な役割を果たしていることから、DOWSER を用いてエネルギー的に安定な水分子を同定し、構造に含める作業を行った。DOWSER はエネルギー的に安定な水分子の位置を計算し、これを構造データに追加することが可能である。この手順により、機能的に重要な膜タンパク質の構造を構築した。

膜タンパク質の構造準備

初めに、膜タンパク質と脂質膜の複合体構造は、CHARMM-GUI を用いて作成した。複合体構築には置換法 (Replacement Method) を採用し、タンパク質を囲むように脂質二重膜を配置した。以下に CHARMM-GUI の設定手順を示す。

1. 初期設定: CHARMM-GUI サイトで *Protein/Membrane System* を選択し、準備した PDB ファイルをアップロード。
2. 分子選択: リガンドと結晶水を含めるように選択。
3. プロトン化とジスルフィド結合の設定: プロトン化残基やジスルフィド結合を指定。
4. 脂質二重膜の構築: Heterogeneous Lipid を選択し、脂質の XY 軸長を $80 \text{ \AA} \times 80 \text{ \AA}$ に設定。
5. イオン追加: NaCl 濃度を 0.15M に設定し、Na⁺ と Cl⁻ 追加。

分子動力学計算は Amber 力場で行うため、CHARMM-GUI で出力された pdb ファイルを Amber 形式に変換する。

続いて、リガンドの力場構築には、AmberTools の `antechamber` と `parmchk2` ツールを使用した。

続いて、実際にシミュレーションを行うシステムに対して、プロトン化状態の修正、ジスルフィド結合の形成を行う。プロトン化状態の修正に関しては、タンパク質構造が与えられた環境下で適切なプロトン化状態を取るように、残基の名前を変更する必要がある。本研究では、脂質環境下にある

Glu122 は GLH に変更し、膜表面側にある残りの GLu,Asp,Arg,Lys は中性 (pH=7) での状態を採用した。最後に、タンパク質-リガンド複合体の分子動力学シミュレーションで用いるためのシミュレーション用のファイルを準備した。シミュレーションに使用する力場関数に関して、タンパク質には ff12SB、脂質には lipid21、リガンドには gaff2、水分子には tip3p モデルを用いた。

系のユニットセルサイズは CHARMM-GUI で得られた情報に基づいて、以下のように指定する。

表 2.1: シミュレーションに用いた系の大きさ

モデル名	原子数	ボックスの大きさ (x, y, z)
inactive (3P0G)	84225	(80.2377, 80.2377, 113.064)
active (2RH1)	70608	(80.0239, 80.0239, 136.043)

2.1.2 構造最適化

構造最適化は 3 つのステップに分けて行った。全てのステップにおいて、最急降下法で最小化を開始し、200 ステップ後に共役勾配法 (CG 法) へ切り替えている。初めに、水素以外の全ての原子に位置制約を課し、タンパク質周辺の脂質や水分子の構造最適化を行った。続いて、膜タンパク質の主鎖原子、リガンドの重原子、脂質の head 部分に位置制約を課し、膜タンパク質の側鎖原子や水分子の構造最適化を行った。最後に、上記の束縛力を弱めて、全体の構造を最適化した。

2.1.3 熱平衡化

熱平衡化は 4 つのステップに分けて行った。初めに、NVT アンサンブルで系を徐々に加熱させた。初期速度は温度 $T = 0[\text{K}]$ のマクスウェル分布に従って与え、100 ps で温度を $T = 310[\text{K}]$ まで上昇させた。タンパク質とり

ガンドに対して $2.0 \text{ kcal/mol/\AA}^2$ の拘束を加えた。

続いて、NVT 条件下で 200 ps のシミュレーションを行った。langevin 法を用いた温度制御を行いながら、 $T = 310[\text{K}]$ で設定した。タンパク質とリガンドの位置拘束の重みを $0.1 \text{ kcal/mol/\AA}^2$ に設定した。

続いて、上記と同様の NVT 条件下で、今度は位置拘束を全て外して、200 ps のシミュレーションを行った。

最後に、NP γ T 条件下で 700 ps のシミュレーションを行った。langevin 法を用いた温度制御と Berendsen 法を用いた圧力制御を行いながら、 $T = 310[\text{K}]$ で設定した。位置拘束を外し、膜系のシミュレーション用に表面張力 $\gamma = 17 \text{ dyne/cm}$ で設定した。

2.1.4 サンプリング

ここでは NP γ T シミュレーションによるサンプリングを 56 ns 行った。設定は、熱平衡化で行った NP γ T シミュレーションと同様である。

2.1.5 保存された結晶水の同定

保存された結晶水を同定するために、サンプリングで得られたシミュレーションデータから水分子の位置とその密度を分析した。具体的には、シミュレーションの各タイムステップで得られる水分子の座標を格子状に集約し、サンプリングで得られたトラジェクトリの半分以上で保存されている水分子を同定した。

2.1.6 NVE

最後に NVE シミュレーションによるサンプリングを行った。初期状態として、NP γ T シミュレーションによるサンプリングで得られた座標と速度を用いた。保存した座標と速度のうち、2 つを使って $\Delta t = 0.5 \text{ fs}$ 、1000 ps のシミュレーションを実行した。最終的に、inactive 状態 active 状態双方で 1000ps トラジェクトリを計 10 本取得した。

2.2 グラフ理論

2.2.1 ネットワーク構築

本研究では、ノードをアミノ酸残基とし、以下の条件に基づいてエッジを形成した。1. 残基ペア間の最短距離が 3 \AA 未満である場合、エッジを形成する。2. アミノ酸配列上で隣接する残基間のエッジは削除する。ここで、残基ペア間の最短距離は、NVE トラジェクトリの全フレームの座標を時間平均することで得られた平均構造を用いて計算したものである。隣接する残基間のエッジを削除した理由は、本研究では native contact に焦点を当てるためである。この native contact はアンフォールド状態からフォールド状態への変化で形成される残基間相互作用であり、構造安定性や機能に関連する重要な情報が含まれている。本研究では inactive 状態から active 状態へのネットワークの再編に着目しているため、はじめから存在し続けている隣接残基間の共有結合は取り入れないことにした。

エッジの重みづけとして、残基間最短距離の 2 乗逆数の平均 $\langle \frac{1}{d^2} \rangle$ を用いた。

2.2.2 Louvain 法によるコミュニティ検出

Louvain 法は、ネットワークにおけるコミュニティ構造を検出するための効率的なアルゴリズムである。本研究では、Louvain 法 [21] を実装してネットワークデータを解析し、最適なコミュニティ分割とモジュラリティの評価を行った。

モジュラリティの定義

モジュラリティ Q は、ネットワーク内で検出されたコミュニティ構造の質を評価する尺度である。具体的には、同一コミュニティ内に存在するエッジの密度が、ランダムに生成されたネットワークと比較してどれだけ顕著に

密集しているかを示す。モジュラリティは次式で定義される：

$$Q = \sum_{c \in \text{コミュニティ}} \left(\frac{m_c}{m} - \left(\frac{K_c}{2m} \right)^2 \right), \quad (2.1)$$

ここで、 m_c はコミュニティ c 内のエッジの重み合計、 m はネットワーク全体のエッジの重み合計、 K_c はコミュニティ c 内のノードの次数の合計である。この定義により、同一コミュニティ内でのエッジ密度が高い場合にモジュラリティの値が大きくなり、ネットワークのコミュニティ構造がより明確であることを示す。

アルゴリズムの流れ

Louvain 法は以下の 2 つのステップを繰り返してコミュニティを検出する：

1. **局所移動ステップ**：各ノードを隣接するコミュニティに移動させ、モジュラリティの増加が最大となる配置を探索する。このプロセスにおいて、モジュラリティの変化量 ΔQ は以下で計算される：

$$\Delta Q = \Delta Q_{\text{削除}} + \Delta Q_{\text{挿入}}, \quad (2.2)$$

ここで、 $\Delta Q_{\text{削除}}$ はノードを元のコミュニティから削除した際のモジュラリティの変化量、 $\Delta Q_{\text{挿入}}$ はノードを新たなコミュニティに追加した際のモジュラリティの変化量である。

2. **集約ステップ**：検出されたコミュニティを 1 つのノードとして扱い、新しいネットワークを構築する。この操作を通じて、階層的なコミュニティ構造を得ることが可能である。

本研究では、Python を用いて Louvain 法を実装した。モジュラリティの増加が小さい場合に収束するよう、収束条件として閾値 $\epsilon = 10^{-7}$ を設定した。また、ノードの移動においてランダム性を導入し、複数回の試行によりモジュラリティの最適解を探索した。そして、100 回の試行の中で最大のモジュラリティおよびその標準偏差と標準誤差を算出した。

第 3 章

結果と考察

3.1 分子動力学 (MD) シミュレーション

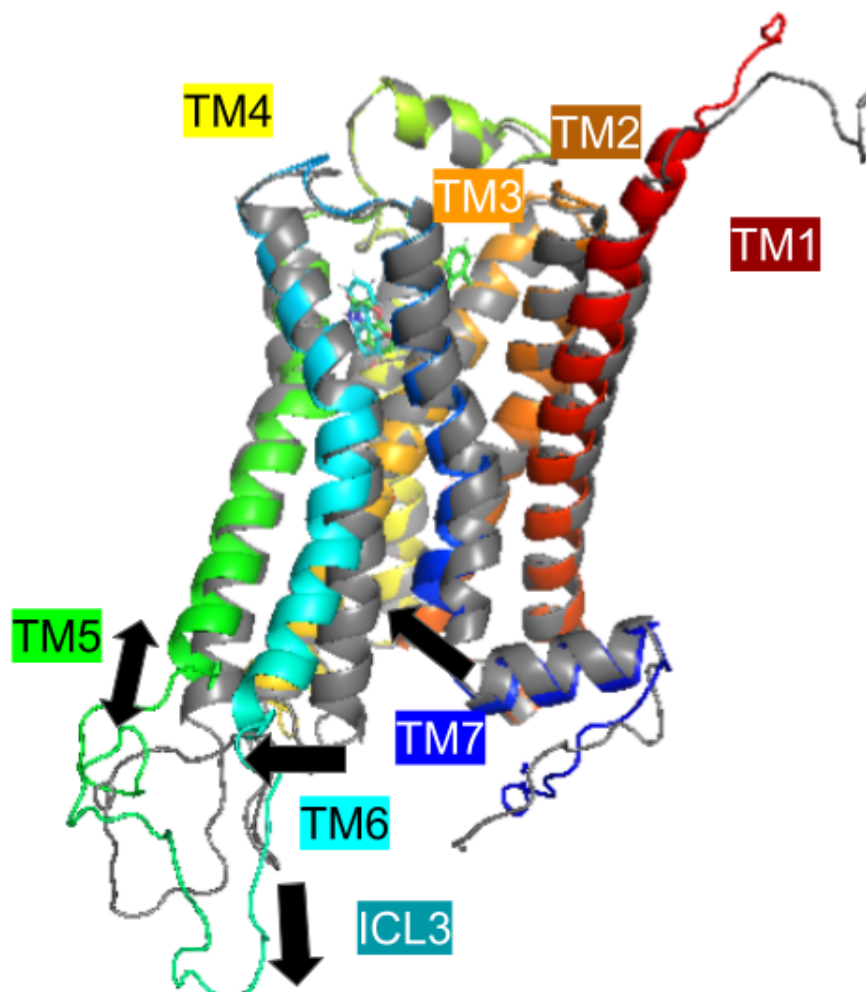
分子動力学シミュレーションにより、 β_2 AR の inactive 状態 active 状態双方で 1000ps トラジェクトリを計 10 本取得した。

β_2 AR の活性化による構造変化

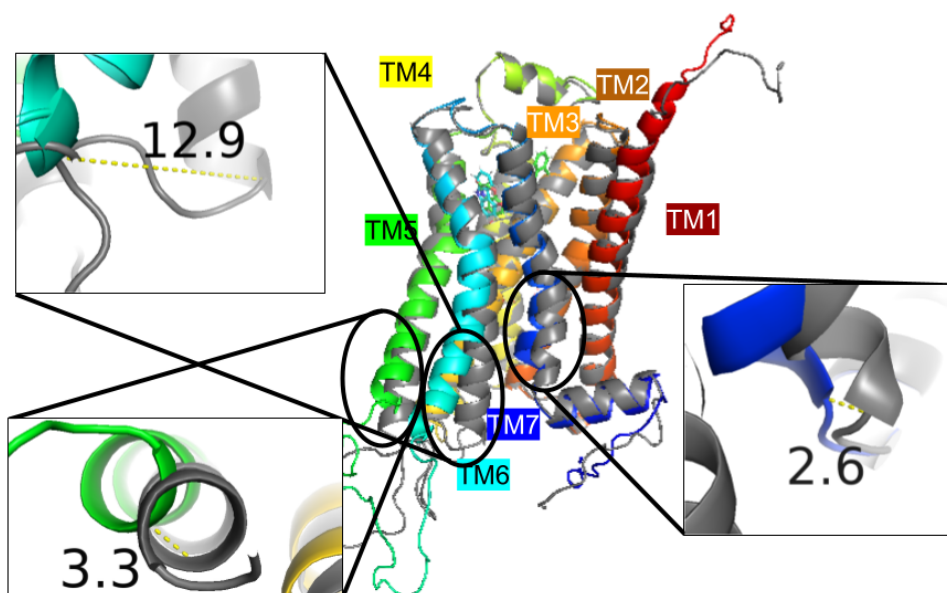
アゴニストが結合した β_2 AR は、顕著な構造変化を引き起こす [22][23] ことがわかっている。主な変化として以下が挙げられる：

1. **TM6** の外側への動き：TM6 の細胞質側末端が、約 14 Å 外側に移動する。
2. **TM5** の外側への動き：TM5 の細胞質側末端が、外側に移動する。
3. **TM7** の内側への動き：TM7 の細胞質側末端が、内側に移動する。

本研究で得られた、シミュレーション中の原子の平均的な配置を示した β_2 AR の inactive 構造と active 構造を重ね合わせたのが以下の図である。



(a) β_2 AR の TM5, TM6, TM7 の変化。



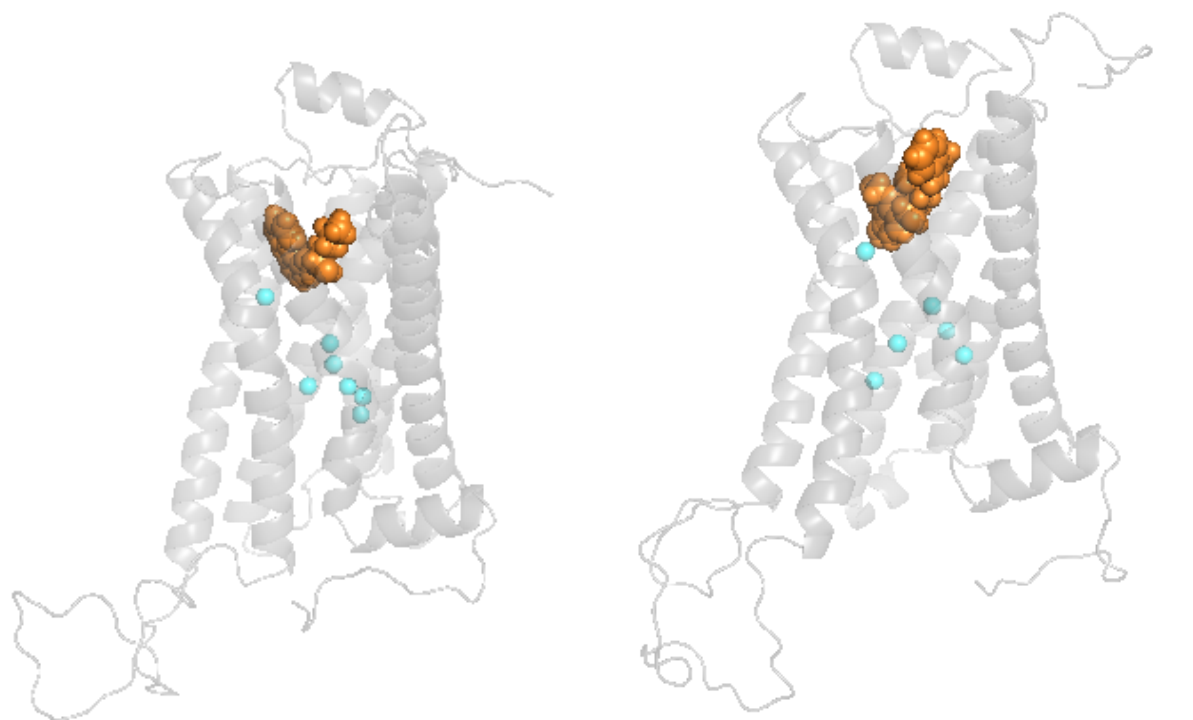
(b) β_2 AR の TM5, TM6, TM7 の変化の詳細。

図 3.1: β_2 AR の inactive 構造 (灰色) と active 構造 (色付き) の重ね合わせ。

TM6 の細胞質側末端にある残基の C_{α} 原子が、TM7 から離れる方向に 12.9 Å 移動する様子が観察された。TM5 の細胞質側末端部分は末端の C_{α} 原子が TM3 から離れる方向に 3.3 Å 移動する様子が観察された。TM7 の細胞質側末端にある残基の C_{α} 原子が、タンパク質の内側に近づく方向に 2.6 Å 移動する様子が観察された。TM5 と TM6 を結ぶ細胞内ループ 3(ICL3) においても、G タンパク質結合部位を開くように特定の方向に並んでいるような様子が観察された。

β_2 AR の保存された結晶水

DOWSER によって検出された、エネルギー的に安定な水分子は、inactive 状態で 14869 個、active 状態で 18876 個同定された。そのうち、シミュレーション中で保存されている結晶水は inactive 状態で 7 個、active 状態で 6 個同定された。



(a) β_2 AR の inactive 状態における保存された結晶水。

(b) β_2 AR の active 状態における保存された結晶水。

図 3.2: 保存された結晶水

シミュレーション中で保存されている inactive 状態の 7 個の結晶水、active 状態の 6 個の結晶水のうち、双方で共通の位置に存在していた結晶水は 4 つあったため、残基番号をそれぞれ (344,346,347,348) と定めた。それ以外の結晶水は、(355,356,357,358,359) の残基番号を振り分けた。

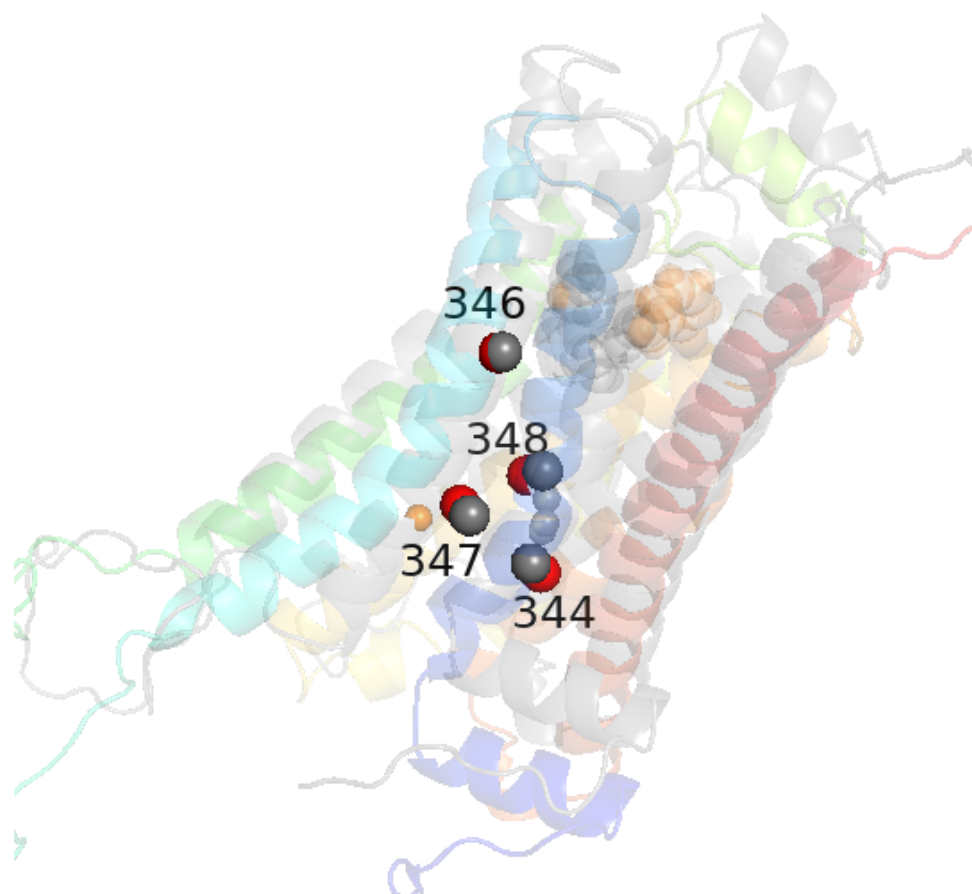


図 3.3: 保存された結晶水のうち、 β_2 AR の inactive 状態と active 状態の双方で共通の位置に存在する水分子。

3.2 β_2 AR の inactive および active 状態のコミュニティ検出

各トラジェクトリーから得られた平均構造の座標データを基に、ネットワークのエッジの重みとして用いる残基間最短距離の 2 乗逆数の平均 $\langle \frac{1}{d^2} \rangle$ を計算した。

エッジを以下の条件で形成させた後、エッジの重みを加え、inactive 状態 active 状態双方のネットワークを構築した。1. 残基ペア間の最短距離が 3 Å 未満である場合、エッジを形成する。2. アミノ酸配列上で隣接する残基感のエッジは削除する。

以下が構築されたネットワークの詳細である。

表 3.1: ネットワーク上のノード数と、残基ペア間の最短距離が 3 Å 未満のエッジ数

モデル名	ノード数	エッジ数
inactive(2RH1)	351	16332
active(3P0G)	349	15433

構築されたネットワークを基に、コミュニティ検出を行った。

Louvain 法の信頼性

ネットワークにおけるコミュニティ構造を検出するために用いた Louvain 法の結果の信頼性は、モジュラリティ Q の値を用いて評価される。モジュラリティ Q の値は通常 -1 から 1 の範囲を取り、以下のように解釈される：

- Q が負：分割がネットワーク構造と一致しておらず、不適切な分割である。
- Q が 0 に近い：ネットワークがランダム構造に近い。
- Q が正：ネットワーク内にコミュニティ構造が存在する。

本研究の解析対象である $\beta_2\text{AR}$ の inactive および active 状態のコミュニティ検出におけるモジュラリティ Q の最良値および標準偏差を計算した。

■ 試行回数 100 回におけるモジュラリティ値の平均および標準偏差

- inactive 状態: $Q = 0.5163 \pm 0.0054$
- active 状態: $Q = 0.5239 \pm 0.0044$

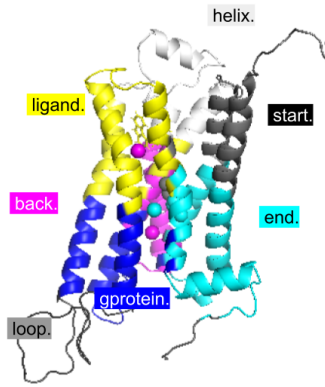
さらに、モジュラリティ値の標準誤差は以下の通りです：

- inactive 状態: ± 0.0005
- active 状態: ± 0.0004

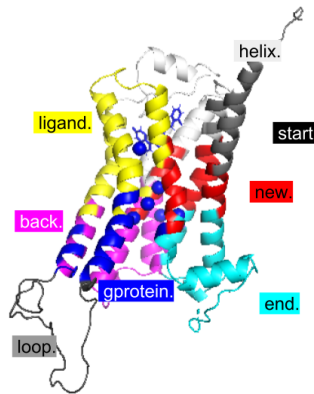
inactive 状態と active 状態のモジュラリティ値はどちらも正の値を示しており、ネットワーク内に明確に分割されたコミュニティ構造があることを示唆している。

3.2.1 Louvain 法で検出されたコミュニティ

以下に $\beta_2\text{AR}$ の inactive 構造と active 構造で検出されたコミュニティを示す。



(a) inactive 状態において検出されたコミュニティ構造を色分けして示した図。各色と名付けたコミュニティは以下に表す：黒 (start)、シアン (end)、白 (helix)、ピンク (back)、グレー (loop)、黄色 (ligand)、青 (gprotein)。



(b) active 状態において検出されたコミュニティ構造を色分けして示した図。各色と名付けたコミュニティは以下に表す：黒 (start)、シアン (end)、白 (helix)、ピンク (back)、グレー (loop)、黄色 (ligand)、青 (gprotein)、赤 (new)。

図 3.4: β_2 AR の Louvain 法で検出されたコミュニティ構造

まず双方のコミュニティに共通することとして、リガンド結合部位と活性部位である G タンパク質結合部位に対応する独立したコミュニティが、黄色 (ligand)、青 (gprotein) として検出された。また inactive 状態から active 状態への変化として、青 (gprotein) コミュニティの再編成が起きていることと、赤 (new) コミュニティが新規に形成されたことが挙げられる。

表 3.2: 活性化に伴うコミュニティの再編成と、コミュニティを構成している残基群の比較

活性化に伴うコミュニティの再編成	inactive での構成要素	active での構成要素
青 (gprotein) コミュニティの再編成	TM3, TM5, TM6 の細胞質側残基	TM6 の細胞質側残基、リガンド、6 つの保存された結晶水
赤 (new) コミュニティの新規形成	なし	TM7, TM1, TM2 の中間付近に位置する残基

青 (gprotein) コミュニティの再編成が起きたのは、TM5, TM6 の細胞質側がタンパク質の中心から離れるように外側に動いたことで、もともと同じ青 (gprotein) コミュニティに所属していた TM3 と別のコミュニティに分かれたためだと考えられる。実際に TM6 の顕著な動きにより、TM5 と TM6 が協調して動くことが確認されている。しかしその独立した TM6 の青 (gprotein) コミュニティに、active 状態に存在している保存された結晶水 6 つ全てとりガンドが含まれたことは興味深い結果であり、これと、赤 (new) コミュニティの新規形成が、シグナル伝達経路にどう影響を与えるかは以後の分析で考察していくこととする。

ここで、inactive 状態と active 状態で示されているネットワークの全体密度を計算する。全体エッジ密度 D_{global} の計算式は以下のように表される。

$$D_{\text{global}} = \frac{\sum_{i=1}^N \sum_{j=i+1}^N w(i, j)}{\tilde{\omega} \times N(N-1)} \quad (3.1)$$

ここで $\sum_{i=1}^N \sum_{j=i+1}^N w(i, j)$ はネットワーク内の全てのエッジ重みの総和、 $\tilde{\omega}$ はネットワーク全体のエッジの平均重み (全エッジの重みの総和 ÷ 全エッジ数)、 N はネットワーク内の全ノード数、 $N(N-1)$ は理論的に存在しうるグラフ (完全グラフ) での最大エッジ数を示している。

分子はネットワーク内の実際のエッジ重みの総和を表しており、分母はネットワーク内の理論的に考えられる全エッジが、ネットワークの平均的な重みを持つと仮定した時に期待される重みの総和を表している。

この式に基づいて、inactive 構造と active 構造のネットワークの全体エッジ密度を計算すると、それぞれ以下ようになる。

- inactive 状態 : $D = 0.133$

- active 状態 : $D =: 0.127$

active 状態ではネットワーク全体の密度が若干低下した。

3.2.2 コミュニティ内およびコミュニティ間のエッジ密度

active 状態で再編成されたコミュニティや新しく検出されたコミュニティの役割を定量的に分析するために、ネットワークの全体密度 D_{global} の計算で用いた密度の概念を用いたさらなる計算を行った。続いて、inactive 状態と active 状態双方においてそれぞれコミュニティ内のエッジ密度、コミュニティ間のエッジ密度を計算した。

コミュニティ内エッジ密度

コミュニティ内エッジ密度 D_{internal} の計算式は以下のように表される。

まずコミュニティごとにサブグラフを作成する。ただしコミュニティ間のエッジは削除し、独立したコミュニティを表現している。続いてコミュニティ内エッジ密度を計算する。

$$D_{\text{internal}} = \frac{\sum_{(u,v) \in E_C} w_{uv}}{\tilde{\omega} \cdot n_c(n_c - 1)} \quad (3.2)$$

ここで E_C はコミュニティ C 内の全エッジの集合、 w_{uv} はコミュニティ C 内のノード u とノード v の間の実際のエッジ重み、 $\tilde{\omega}$ はネットワーク全体のエッジの平均重み (全エッジの重みの総和 ÷ 全エッジ数)、 n_c はコミュニティ C 内の全ノード数、 $n_c(n_c - 1)$ は理論的に存在しうるコミュニティ内のグラフ (完全グラフ) での最大エッジ数

分子はコミュニティ内の実際のエッジ重みの総和を表しており、分母はコミュニティ内の理論的に考えられる全エッジが、ネットワークの平均的な重みを持つと仮定した時に期待される重みの総和を表している。

コミュニティ内エッジ密度 D_{internal} の値は以下のように解釈される：

- D_{internal} が 0 に近い：コミュニティ内のノード間での接続が少ないか、エッジの重みが非常に小さい。相互作用が弱く、コミュニティと

しての結束が薄い。

- D_{internal} が 1 に近い：コミュニティ内のノード間の接続が非常に密であり、実際のエッジ重みが最大理論値に近い。コミュニティ内の相互作用が非常に強い。

上記の式に基づいて、inactive 構造と active 構造のコミュニティ内エッジ密度を計算すると、それぞれ以下ようになる。

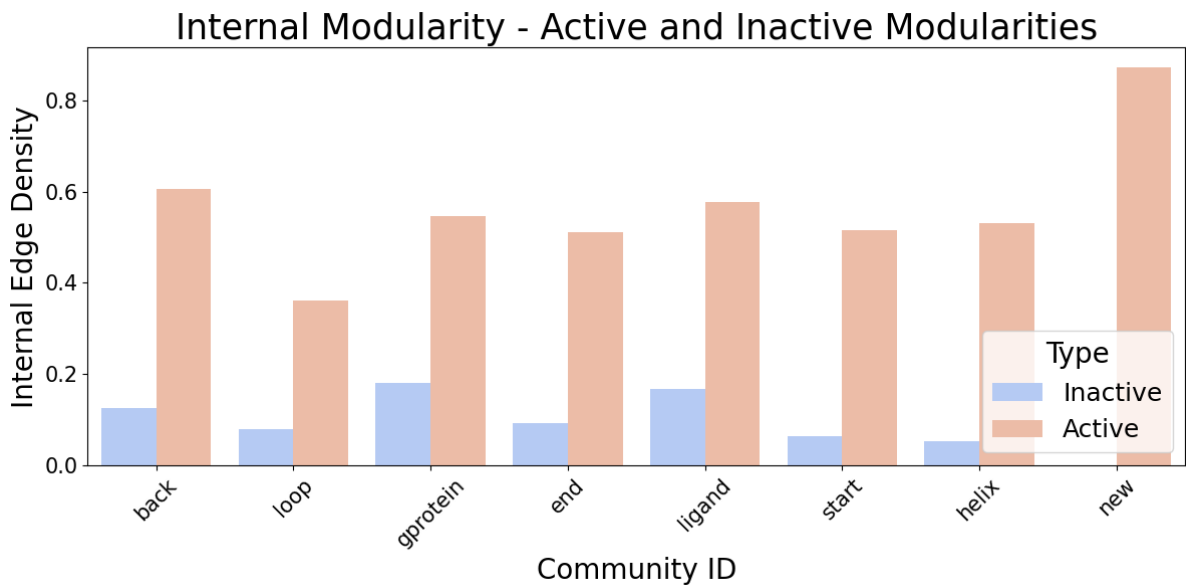


図 3.5: $\beta_2\text{AR}$ の inactive 状態と active 状態において検出されたコミュニティのコミュニティ内エッジ密度を示した図。コミュニティの名称は、コミュニティ構造で名付けたものである。

inactive 構造では、全てのコミュニティ内エッジ密度が、0.2 より低い値を示した。しかし active 構造で全てのコミュニティのコミュニティ内エッジ密度が増加し、新たに出現した new コミュニティは 0.87 という高い値を示した。活性化に伴って、すべてのコミュニティの内部の相互作用が強い結びつきを形成している構造に変化したことがわかる。

コミュニティ間エッジ密度

コミュニティ間エッジ密度 D_{inter} の計算式は以下のように表される。

$$D_{\text{inter}} = \frac{\sum_{\substack{u \in C_i \\ v \in C_j}} w_{uv}}{\tilde{\omega} \cdot n_c n_{c'}} \quad (3.3)$$

ここで C_i と C_j は異なるコミュニティ i と j 、 w_{uv} はコミュニティ C_i 内のノード u とコミュニティ C_j 内のノード v の間の実際のエッジ重み、 $\tilde{\omega}$ はネットワーク全体のエッジの平均重み (全エッジの重みの総和 ÷ 全エッジ数)、 n_c と $n_{c'}$ はコミュニティ C_i 内とコミュニティ C_j 内の全ノード数、 $n_c n_{c'}$ は理論的に存在しうるコミュニティ C_i と C_j 間のグラフ (完全グラフ) での最大エッジ数

分子はコミュニティ間の実際のエッジ重みの総和を表しており、分母はコミュニティ間の理論的に考えられる全エッジが、ネットワークの平均的な重みを持つと仮定した時に期待される重みの総和を表している。

コミュニティ内エッジ密度 D_{inter} の値は以下のように解釈される：

- D_{inter} が 0 に近い：異なるコミュニティ間の接続がほとんどなく、エッジの重みが小さい。コミュニティ間での相互作用がほぼ存在しないか非常に弱い。
- D_{inter} が 1 に近い：異なるコミュニティ間の接続が非常に密であり、実際のエッジ重みが理論値に近い。異なるコミュニティ間で強い相互作用や影響のやりとりが行われている。

上記の式に基づいて、inactive 構造と active 構造のコミュニティ間エッジ密度を計算すると、それぞれ以下ようになる。

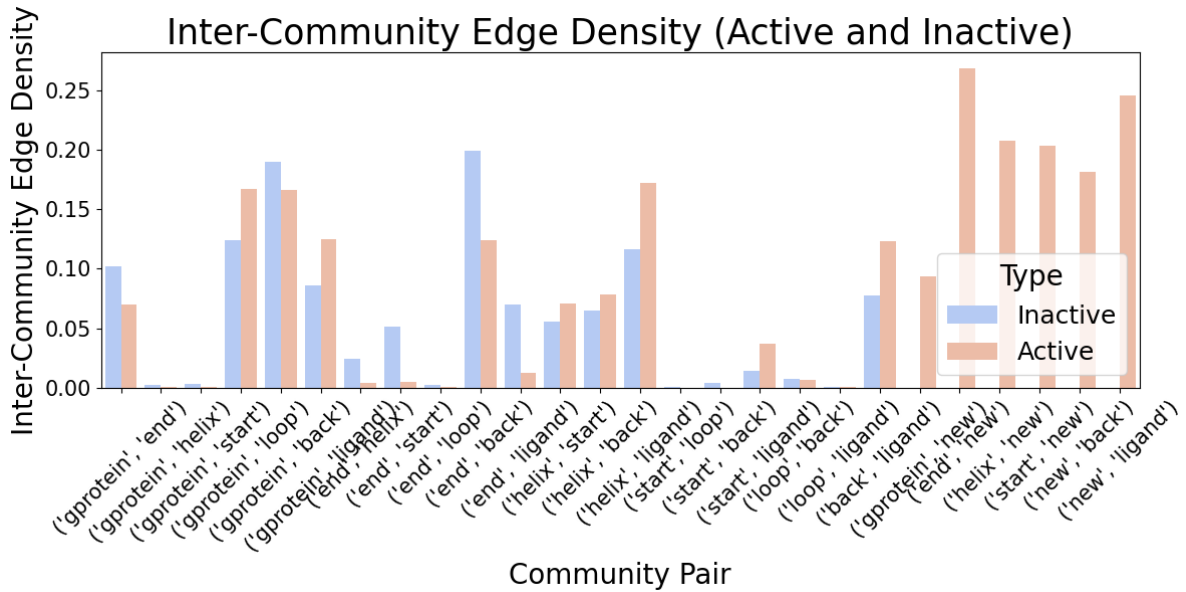


図 3.6: β_2 AR の inactive 状態と active 状態において検出されたコミュニティのコミュニティ間エッジ密度を示した図。コミュニティの名称は、コミュニティ構造で名付けたものである。

ただし、inactive 状態と active 状態の双方でコミュニティ間エッジ密度が 0 だったコミュニティペアは、図から除外している。

コミュニティ間エッジ密度 Q の値は以下のように解釈される：

- Q が 0 に近い：コミュニティ間でエッジが少なく、各コミュニティがほぼ独立している。
- Q が 1 に近い：コミュニティ間で多数のエッジが存在し、コミュニティ間の結びつきが強い。

active 状態になると (gprotein,loop)(gprotein,ligand)(helix,ligand)(back,ligand) といったコミュニティペアが増加し、new に関わる 5 つのコミュニティペアが相対的に高い値を示した。

inactive 状態と active 状態の全体密度、コミュニティ密度の考察

■密度に関わる 3 つの変数における、inactive 状態から active 状態への変化
ここまで inactive 状態と active 状態のコミュニティ内エッジ密度、コミュ

ニティ間エッジ密度、全体エッジ密度の変化を通じて、コミュニティ内およびコミュニティ間の相互作用を定量的に評価してきた。それぞれ、以下のことが示唆される。

1. 全体エッジ密度の微減: active 状態への遷移に伴い、ネットワーク全体のエッジ密度がわずかに減少した。
2. コミュニティ内エッジ密度の増加: active 状態では、すべてのコミュニティ内エッジ密度が増加し、特に new コミュニティが高い値を示した。
3. コミュニティ間エッジ密度の増加: active 状態において、ligand や gprotein に関わるコミュニティペアは増加した。新しく生成された new コミュニティに関連するコミュニティペアは高い値を示した。

これらの結果は、active 状態への遷移に伴う分子全体のネットワーク再編成が、コミュニティ内部の相互作用を強化させるとともに、情報伝達に関わる分子間の情報伝達の効率化を支える重要なメカニズムであることを示している。特に、new コミュニティが形成されたことで、リガンド結合部位と活性部位間の情報伝達が促進されている点を考慮すると、リガンド結合部位や活性部位である G タンパク質結合部位間の情報伝達を促進する導管として働いている可能性があることが示唆された。このことは、アロステリック遷移における協調的な相互作用の重要性を示唆しており、分子内情報伝達の動的メカニズム解明に貢献するものと考えられる。

(密度変化や new コミュニティの重要性や生理的・機能的意義を、先行研究と交えて)

ここまで、コミュニティ内外のエッジ密度の変化に基づき、active 状態への遷移に伴う分子全体のネットワーク再編成について議論してきた。しかし、コミュニティ単位の相互作用だけでは、各ノードが分子全体の接続性や情報伝達にどのように寄与しているのかを十分に把握することはできない。そこで、次にノード単位でのコミュニティに与える影響を詳細に分析し、分子内のネットワークの構造や情報伝達効率への貢献度を評価する。

3.3 ノード削除による active ネットワーク接続性への影響

あるノードが active ネットワーク全体または所属するコミュニティの接続性に与える影響を定量化に評価するために、impact score という変数を導入した。これは、特定のノードを active ネットワークから削除した際に生じる全体エッジ密度とコミュニティ内エッジ密度の変化量を基に計算した。全体エッジ密度とコミュニティ内エッジ密度はそれぞれ前述の式に従っている。以降、全体エッジ密度の変化量を global impact score、コミュニティ内エッジ密度の変化量を community impact score と表現する。

また、global impact score と community impact score について、Z スコアを算出することで影響の大小を統計的に評価する。Z スコアは、あるデータ点で平均からどれだけ標準偏差の単位で離れているかを示す指標である。Z スコアの計算式は以下のとおりである。

$$z = \frac{x - \mu}{\sigma} \quad (3.4)$$

ここで x はデータ点の値、 μ はデータセットの平均、 σ はデータセットの標準偏差を示している。Z スコアの解釈は以下のとおりである。

- $z < 0$: データ点は平均よりも小さい値である。
- $z = 0$: データ点は平均と一致している。
- $z > 0$: データ点は平均よりも大きい値である。
- $z > 2$: データ点は平均から 2 標準偏差以上離れており、全体の約 2.5% にあたるくらい高い値である。

global impact score

global impact score の変化量が大きかったノードのうち、Z スコアが 2 以上であったノードを昇順に並び替えたのが以下の表である。

Node	Global Impact Score	Community	Node type
66	3.327×10^{-3}	new	Motif
62	3.203×10^{-3}	new	Other
273	3.033×10^{-3}	ligand	Motif
303	3.027×10^{-3}	new	Ligand-Site
69	2.952×10^{-3}	new	Motif
111	2.945×10^{-3}	new	Other
356	2.872×10^{-3}	gprotein	X-Water
96	2.860×10^{-3}	helix	Ligand-Site
107	2.847×10^{-3}	new	Motif

表 3.3: Top 9 Nodes by Impact Score

ここで Community の名称は、コミュニティ構造を示した時に名付けたものである。

上位 9 このうち 6 このノードが new コミュニティに属しており、active 構造で検出された新しいコミュニティがタンパク質全体に重要な役割を果たしていることが示唆された。また、モチーフに関連するノードも上位 9 このうち 4 こ存在しており、これらの領域が全体のネットワーク密度に強い影響を与えていることが示唆される。

community impact score

community impact score が大きかったノードのうち、Z スコアが 2 以上であったノードを昇順に並び替えたのが以下の表である。

Node	Community Impact Score	Community	Node type
259	2.084×10^{-1}	gprotein	Other
66	1.974×10^{-1}	new	Motif
348	1.882×10^{-1}	gprotein	X-Water
356	1.844×10^{-1}	gprotein	X-Water
347	1.797×10^{-1}	gprotein	X-Water
69	1.652×10^{-1}	new	Motif
355	1.649×10^{-1}	gprotein	X-Water
262	1.600×10^{-1}	gprotein	Gprotein-Site
344	1.588×10^{-1}	gprotein	X-Water
343	1.584×10^{-1}	gprotein	Ligand
34	1.553×10^{-1}	new	Other
346	1.507×10^{-1}	gprotein	X-Water
23	1.470×10^{-1}	start	Other
264	1.253×10^{-1}	gprotein	Other
24	1.249×10^{-1}	start	Other
261	1.250×10^{-1}	gprotein	Gprotein-Site
258	1.239×10^{-1}	gprotein	Gprotein-Site
311	1.211×10^{-1}	new	Motif
1	1.210×10^{-1}	start	Other

表 3.4: Top 19 Nodes by Impact Score

19 このうち 12 このノードが gprotein コミュニティに属しており、活性化による構造変化が見られた TM5, TM6, TM7 の細胞質側領域と ICL3 がここに含まれていることから、gprotein コミュニティに属する残基がコミュニティ内の重要な結束性に寄与している可能性があることが示唆された。また、保存された結晶水 6 つ全てが含まれており、保存された結晶水が G タンパク質結合領域内で強い影響力を持つことを示している。

global impact score、community impact score の考察

global impact score が目立った残基は、ネットワーク全体の構造を保つ役割を果たしていると推測される。つまり、global impact score が高かった new コミュニティに属する残基やモチーフは、ネットワーク全体の「骨格」としてリガンド結合によるネットワーク全体の再編成を保っており、アロステリックな影響を広げる重要な起点となっていることが示唆される。一方で community impact score が目立った残基は、局所的なコミュニティ構造を保つ役割を果たしていると推測される。つまり、gprotein,loop コミュニティに属する残基や、保存された結晶水は局所的な「足場」を提供し、Gタンパク質の結合やシグナル伝達の効率化を高めていることが示唆される。

第 4 章

まとめ

本研究では、 $\beta 2AR$ の inactive 状態および active 状態におけるネットワーク構造を解析し、Louvain 法を用いて検出されたコミュニティの特性を比較した。まず検出されたコミュニティを比較すると、G タンパク質結合部位が再編成され、新規のコミュニティが形成されていることが明らかになった。続いてそれぞれのネットワークの全体エッジ密度、コミュニティ内エッジ密度、コミュニティ間エッジ密度を比較した。全体エッジ密度に関しては、inactive 状態では 0.133、active 状態では 0.127 であった全てのコミュニティでコミュニティ内エッジ密度が増加し、新たに出現した new コミュニティは 0.87 という高い値を示した。コミュニティ間エッジ密度の解析では、(gprotein,loop)(gprotein,ligand)(helix,ligand)(back,ligand) といったコミュニティペアの相互作用が強化されており、特に新しく生成された new コミュニティに関連するコミュニティペアは高い値を示した。これらの結果は、active 状態への遷移に伴う分子全体のネットワーク再編成が、コミュニティ内部の相互作用を強化させるとともに、情報伝達に関わる分子間の情報伝達の効率化を支える重要なメカニズムであることを示している。特に、new コミュニティが形成されたことで、リガンド結合部位や活性部位である G タンパク質結合部位間の情報伝達を促進する導管として働いている可能性があることが示唆された。

最後にノード削除が全体エッジ密度とコミュニティ内エッジ密度に与え

る影響の解析では、リガンド結合部位や new コミュニティに属しているモチーフに関するノードが全体エッジ密度に、リガンドや g タンパク質結合部位、保存された結晶水がコミュニティ内エッジ密度に与える影響が大きかった。前者はネットワーク全体の「骨格」としてアロステリックな影響を広げる重要な起点となっていることが、後者は局所的な「足場」を提供し G タンパク質の結合やシグナル伝達の効率化を高めていることが示唆された。

4.1 今後の展望

本研究ではノード間の距離を重みとした構造ネットワークを用いた。しかし活性化によってダイナミクスや相互作用のみが変化したノードに関しては、構造ネットワークではその変化を捉えることが困難である。そのため、構造のみならず、ダイナミクスや相互作用も反映した変数である熱伝導度 [24] を重みとした物理的な熱ネットワークの構築により、 $\beta 2AR$ のアロステリー機構の解明をより詳細に理解することが期待できる。また、Louvain 法によるコミュニティ検出では、特定の時間スケールでの 1 つのコミュニティ分割しか検出しておらず、異なる時間スケールでの過渡現象を観察できない。そのため、マルコフ安定性 [25] のようなネットワーク内のさまざまなスケールに存在する多層コミュニティ構造を識別できる動力学ベースのマルチスケール方の導入が必要である。

謝辞

学部3年次の生物物理セミナーからはじめ、3年間の研究の遂行にあたり、指導教官として終始多大なご指導を賜った、倭剛久先生先生に深謝致します。研究室選びに迷っている中で学部3年次に出席した生物物理分野に興味を持ち、急遽研究室見学を受け入れてくださった時から今まで、生物物理に関する深い知識と経験に基づく多くの有益な助言をいただきました。木村明洋先生にも、講義やセミナー、研究室での議論を通じて、多くのご指導をいただきました。またB研の皆様には、本研究の遂行にあたり多大なご助言、ご協力頂きました。特に王婷婷氏、杉浦航氏、在田陽一氏には、右も左もわからなかった自分に、生物物理やシミュレーションの基礎をはじめ、私の疑問や悩みをたくさん解決していただきました。大久保雄大氏、吉村風汰氏、大櫃丈氏には、研究室内発表時の議論を通じて多くの刺激をもらいました。

最後に、研究ができる環境を提供していただき、精神的な支えにもなっていただいた両親や祖父、姉に深く感謝いたします。

参考文献

- ¹Q. Cui and M. Karplus, “Allostery and cooperativity revisited”, *Protein Sci.* **17**, 1295–1307 (2009).
- ²H.-X. Zhou and X. Pang, “Electrostatic interactions in protein structure, folding, binding, and condensation”, *Chemical Reviews* **118**, 4–16 (2018).
- ³N. T. Doncheva, K. Klein, F. S. Domingues, and M. Albrecht, “Analyzing and visualizing residue networks of protein structures”, *Trends Biochem. Sci.* **36**, 179–182 (2011).
- ⁴A. J. M. Martin, M. Vidotto, F. Boscariol, T. D. Domenico, I. Walsh, and S. C. E. Tosatto, “Ring: networking interacting residues, evolutionary information and energetics in protein structures”, *Bioinformatics* **27**, 2003–2005 (2011).
- ⁵N. T. Doncheva, Y. Assenov, F. S. Domingues, and M. Albrecht, “Topological analysis and interactive visualization of biological networks and protein structures”, *Nat. Protoc.* **7**, 670–685 (2012).
- ⁶A. Ghosh and S. Vishveshwara, “A study of communication pathways in methionyl-trna synthetase by molecular dynamics simulations and structure network analysis”, *Proc. Natl. Acad. Sci.* **104**, 15711–15716 (2007).
- ⁷G. Amitai, A. Shemesh, E. Sitbon, M. Shklar, D. Netanel, I. Venger, and S. Pietrokovski, “Network analysis of protein structures identifies functional residues”, *J. Mol. Biol.* **344**, 1135–1146 (2004).

- ⁸A. Del Sol, H. Fujihashi, D. Amoros, and R. Nussinov, “Residues crucial for maintaining short paths in network communication mediate signaling in proteins”, *Mol. Syst. Biol.* **2** (2006).
- ⁹M. Lim, T. A. Jackson, and P. A. Anfinrud, “Femtosecond near-ir absorbance study of photoexcited myoglobin: dynamics of electronic and thermal relaxation”, *J. Phys. Chem.* **100**, 12043–12051 (1996).
- ¹⁰J.-P. Changeux and S. J. Edelstein, “Allosteric mechanisms of signal transduction”, *Science* **308**, 1424–1428 (2005).
- ¹¹V. J. Hilser and E. B. Thompson, “Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins”, *Proceedings of the National Academy of Sciences* **104**, 8311–8315 (2007).
- ¹²H. Pan, J. C. Lee, and V. J. Hilser, “Binding sites in *Escherichia coli* dihydrofolate reductase communicate by modulating the conformational ensemble”, *Proceedings of the National Academy of Sciences* **97**, 12020–12025 (2000).
- ¹³T. Ishikura, Y. Iwata, T. Hatano, and T. Yamato, “Energy exchange network of inter-residue interactions within a thermally fluctuating protein molecule: a computational study”, *Journal of Computational Chemistry* **36**, 1709–1718 (2015).
- ¹⁴T. Y. T. Wang and W. Sugiura, “Thermal energy transport through nonbonded native contacts in protein”, *The Journal of Physical Chemistry B* **128** (2024).
- ¹⁵V. Cherezov, D. M. Rosenbaum, M. A. Hanson, S. G. F. Rasmussen, F. S. Thian, T. S. Kobilka, H.-J. Choi, P. Kuhn, W. I. Weis, and R. C. Stevens, “High-resolution crystal structure of an engineered human β 2-adrenergic g protein - coupled receptor”, *Science* **318**, 1258–1265 (2007).
- ¹⁶S. G. F. Rasmussen, H.-J. Choi, J. J. Fung, E. Pardon, P. Casarosa, P. S. Chae, B. T. DeVree, D. M. Rosenbaum, F. S. Thian, T. S. Kobilka, A. Schnapp, I. Konetzki, R. K. Sunahara, S. H. Gellman, A. Pautsch,

- J. Steyaert, W. I. Weis, and B. K. Kobilka, “Structure of a nanobody-stabilized active state of the β 2 adrenoceptor”, *Nature* **469**, 175–180 (2011).
- ¹⁷F. Philip, P. Sengupta, and S. Scarlata, “Signaling through a g protein-coupled receptor and its corresponding g protein follows a stoichiometrically limited model”, *J. Biol. Chem.* **282**, 19203–19216 (2007).
- ¹⁸R. Nygaard, T. M. Frimurer, B. Holst, M. M. Rosenkilde, and T. W. Schwartz, “Ligand binding and micro-switches in 7tm receptor structures”, *Trends Pharmacol. Sci.* **30**, 249–259 (2009).
- ¹⁹Y. Lee, S. Choi, and C. Hyeon, “Mapping the intramolecular signal transduction of g-protein coupled receptors”, *Proteins: Structure, Function, and Bioinformatics* **82**, 727–743 (2014).
- ²⁰T. E. Angel, M. R. Chance, and K. Palczewski, “Conserved waters mediate structural and functional activation of family a (rhodopsin-like) g protein-coupled receptors”, *Proceedings of the National Academy of Sciences* **106**, 8555–8560 (2009).
- ²¹V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, “Fast unfolding of communities in large networks”, *J. Stat. Mech.* **2008**, P10008 (2008).
- ²²S. G. F. Rasmussen, B. T. DeVree, Y. Zou, A. C. Kruse, K. Y. Chung, T. S. Kobilka, F. S. Thian, P. S. Chae, E. Pardon, D. Calinski, J. M. Mathiesen, S. T. A. Shah, J. A. Lyons, M. Caffrey, S. H. Gellman, J. Steyaert, G. Skiniotis, W. I. Weis, R. K. Sunahara, and B. K. Kobilka, “Crystal structure of the β 2 adrenergic receptor - gs protein complex”, *Nature* **477**, 549–555 (2011).
- ²³H. Poudel and D. M. Leitner, “Activation-induced reorganization of energy transport networks in the β 2 adrenergic receptor”, *J. Phys. Chem. B* **125**, 6490–6498 (2021).

- ²⁴T. Yamato, T. Wang, W. Sugiura, O. Lapr  v  te, and T. Katagiri, “Computational study on the thermal conductivity of a protein”, *J. Phys. Chem. B* **126**, 3489–3495 (2022).
- ²⁵B. Amor, S. N. Yaliraki, R. Woscholski, and M. Barahona, “Uncovering allosteric pathways in caspase-1 using markov transient analysis and multiscale community detection”, *Mol. BioSyst.* **10**, 2247–2258 (2014).