

課題名 (タイトル) :

データ駆動型分子動力学計算で解明する細胞環境の生体分子ダイナミクス

利用者氏名 :

○松岳 大輔

優 乙石

杉田 有治

George A. Pantelopulos

理研での所属研究室名 :

杉田理論分子科学研究室

本課題では、細胞環境下における生体分子のダイナミクスを調べることを主題の1つとして、課題 A, C では、細胞膜に埋まった膜貫通ペプチドについての分子動力学(MD)シミュレーション、課題 B では、細胞中の分子混雑にある環境を模倣した系での MD シミュレーションを行った。それぞれの結果を下に課題別に記す。

また、計算が実験データをどの程度再現するのかも、シミュレーションにおける重要な点であるから、その点も十分に注意しながら計算を行っている。本報告書ではその点についても記載する。

### 【課題 A】EGFR 膜近傍領域の機械学習アルゴリズムを用いた構造解析 (担当: 松岳)

#### 1. 本課題の研究の背景、目的、関係するプロジェクトとの関係

上皮細胞増殖因子受容体(EGFR)は、細胞表面に存在するシグナル伝達蛋白質 receptor tyrosin kinase (RTK)の1種であり、細胞の増殖や分化などに関与している。RTKは、ホモ二量体のタンパク質であり、各モノマーは細胞外ドメインと細胞内ドメインが1回膜貫通ヘリックスを介して結合した構造を示す。膜貫通(TM)および膜近傍(JM)領域は、他のタンパク質と直接相互作用することのない領域ではあるが、その領域は、細胞外ドメインに与えられたシグナルを細胞質側へ伝達する役割を有しており、EGFRの機能に不可欠な領域である。

EGFR TM-JM 領域の脂質二重膜中の二量体構造は、溶液 NMR によって解かれている (PDB ID: 2M20)。その構造では JM 領域は $\alpha$ ヘリックス構造を示し、そのヘ

リックスは反平行な二量体を形成している。その構造は活性状態での構造を表すと考えられている。

JM 領域に存在するスレオニン残基 T654 がリン酸化された場合、EGFR の受容体活性が低下することが知られている。そこで、EGFR JM 領域の構造についての知見を得るために、理化学研究所・佐甲細胞情報研究室において、EGFR TM-JM 領域からなるペプチドの C 末端部位に蛍光分子である cy3, cy5 を付加し、それを脂質二重膜に埋めて FRET を測定する実験が行われた。その実験では、FRET 効率分布が、T654 側鎖がリン酸化されている時(pT654)とリン酸化されていない時(T654)で大きく異なっていた。この実験結果から、T654 がリン酸化されていないときは、JM 領域は helix が反平行な二量体を形成していると考えられるが、リン酸化されているときは、JM 領域は二量体を形成していない可能性が考えられた(Figure. 1)。

しかしながら、FRET 測定実験は蛍光分子間距離についての情報は与えるものの、その起源となった詳細な3次元構造については何の情報を与えないので、他の手法によって FRET 測定結果を再現する構造を取得することが必要である。近年、MD シミュレーションのトラジェクトリからマルコフ状態モデルを作成し、機械学習で使われるアルゴリズムを援用して FRET 測定実験結果を再現する構造アンサンブルを取得するような手法が提案されている。本課題でも、その手法の使用を検討していたが、現時点ではまだ計算の途上であり、マルコフ状態モデルの計算をするに十分な統計を得るには至っていない。しかしながら、現時点でも FRET 実験結果に一致するようなシミュレーション結果は得られているので、その結果を報告する。

なお本課題は、理研・独創的研究課題“脂質の統合

的理解”の1サブプロジェクト『脂質—膜タンパク質相互作用』に関連する課題の1つとして取り組んでいるものである。

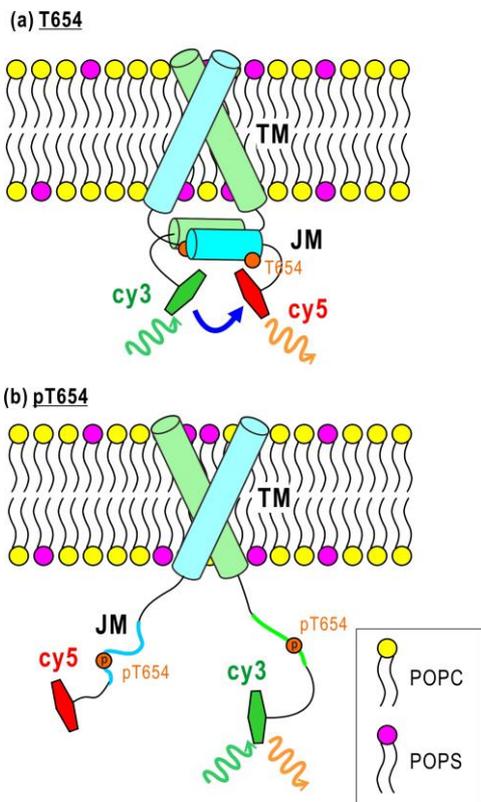


Fig 1. EGFR TM-JM 領域の構造の概念図。

円筒モデルは $\alpha$ ヘリックスを表す。(a) T654 がリン酸化されていないときは、高 FRET 効率が得られることから $\alpha$ ヘリックスは反平行な二量体を形成していると示唆される。(b) 一方 T654 がリン酸化されているとき (pT654) では、FRET 効率が低いことから、JM 領域は二量体を形成していない可能性がある。

## 2. 具体的な利用内容、計算方法

### 2.1. cy3, cy5 の力場モデルの作成

佐甲研で行われた実験を再現するにあたり、本シミュレーションでも EGFR TM-JM 領域からなるペプチドの C 末端に cy3, cy5 分子を付加することとした。しかし、信頼できる cy3, cy5 の力場パラメータが存在しなかったため、シミュレーションの前に force field Toolkit および Gaussian を援用して、cy3, cy5 の力場パラメータを作成した。Gaussian の計算には、GW-ACSL もしくは GW-ACSG を使用した。

### 2.2. 脂質二重膜中での EGFR TM-JM 領域のシミュレーション

現在のシミュレーション時間内では、JM 領域が二量体を形成した状態から、解離した状態へと変化するのを自然なシミュレーションによって再現するのは困難であるため、JM 領域が二量体を形成したモデル (JM\_contact) と、JM 領域が二量体を形成していない状態 (JM\_separate) それぞれを初期モデルとしてシミュレーションをスタートさせた。佐甲研究室の実験では、T654 のリン酸化の有無のほかにも、脂質二重膜の成分を変えた時の影響も調べているので、脂質の成分を実験に合わせて変えたものなど計 12 個を準備し、そのそれぞれに対して GW-MPC および BW-MPC を使用して、シミュレーションを行っている。現時点では、それぞれの系に対して 200 ns のシミュレーションが終了している。すべての MD シミュレーションは、当研究室で開発を行っているプログラム "GENESIS" を用いて行われた。

## 3. 結果

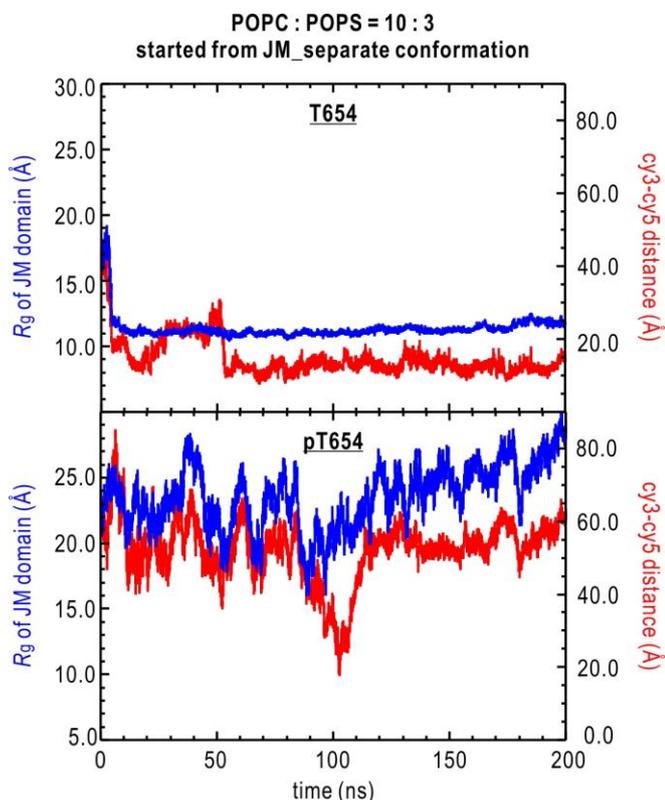
JM 領域が二量体を形成した状態から出発したシミュレーション (JM\_contact) では、T654 のリン酸化の有無にかかわらず、200 ns のシミュレーション中で、二量体が壊れることはなかった。したがって、 $\alpha$ ヘリックスが二量体を形成した構造は非常に安定であると考えられる。

一方、JM 領域が互いに離れた状態からスタートしたシミュレーション (JM\_separate) では、T654 側鎖のリン酸化の有無によって結果が大きく異なっていた。T654 側鎖がリン酸化されていないときは、シミュレーション開始から早い時点で、JM 領域のヘリックスは二量体を形成していた (Fig. 2)。対して、T654 がリン酸化されているときは、JM 領域のヘリックスは 200 ns の間、二量体を形成することはなかった (Fig. 2)。Fig. 2 では、POPC:POPS=10:3 の時の結果を示すが、他の脂質成分の時でも同様の結果が得られていた。

## 4. まとめ

まず、cy3, cy5 について Gaussian などを用いて、信頼できる力場パラメータの作製を行い、それを使用して、脂質二重膜中の EGFR TM-JM 領域のペプチド二量体のシミュレーションを行った。JM 領域が二量体

を形成した状態から出発したシミュレーションでは、初期構造からの有意な変化は見られなかったが、JM 領域が離れた構造から出発した場合、T654 残基がリン酸化されている時と、リン酸化されていない時で、JM 構造の違いに大きな違いが現れた。この結果は、FRET 測定実験結果と合致するものである。



**Figure 2.** JM 領域のヘリックスが離れている状態から出発した時のシミュレーション結果。青線は JM 領域の慣性半径  $R_g$  の時間変化を、赤線は cy3-cy5 間距離の時間変化を表す。

## 5. 今後の計画・展望

現時点では、T654 のリン酸化の有無が、JM 領域の二量体構造の結果に影響を及ぼすことを示唆する結果が得られているが、一方でそれぞれ 1 本のシミュレーションのみしか流していないので、初期乱数の影響を排除できていない。そこで、脂質分子の初期配置や初期乱数などを変えて、さらに 2 本程度シミュレーションを流すことで、現在得られている結果の補強を行う予定である。またマルコフ状態モデルを計算するのに十分な計算時間を得るため、現在のシミュレーションについても継続して計算を行っていく予定である。そのトラジェクトリーを得たうえで、T654 のリン酸化の有無が、なぜ JM 領域の二量体構造に影響を与えた

のか、種々の解析を通じて明らかにする予定である。

また、脂質の構成成分を変えたシミュレーションでは、FRET 測定実験で得られた結果を十分に説明するような結果を得ることができていないので、計算条件を検討したうえで、更なるシミュレーションを実行し、その結果を踏まえて、膜成分が TM-JM 領域の構造に与える影響について考察したい。

## 【課題 A-2】 FGFR3 のレプリカ交換法による膜貫通ヘリックスの構造サンプリング (担当: 松岳)

本年度申請した課題に加えて、松岳は G16021 の課題 4 の継続として、FGFR3 膜貫通領域のシミュレーションも行ったので合わせて報告する。

### 1. 本課題の研究の背景、目的、関係するプロジェクトとの関係

繊維芽細胞増殖因子 FGFR3 もまた RTK の一種であって、細胞増殖や分化に重要な役割を果たす。このタンパク質も他の RTK と同様、細胞外ドメインにリガンドが結合すると一連の構造変化が生じ、細胞内キナーゼドメインにあるチロシン残基がリン酸化されることによって、細胞内のシグナル伝達を引き起こす。

FGFR3 膜貫通(TM)領域に存在する病原性変異体 G380R は、リガンド非結合時においても活性を持つことが知られており、TM 領域の活性時の二量体構造を調べるためのモデルとして用いられている。京都薬科大学の佐藤教授らによって行われた脂質二重膜中での固体 NMR 測定によって、野生型(WT)と G380R 変異体で、TM 領域の二量体構造が異なることが示唆する結果が得られたが、その詳細な構造については今なお不明である。

そこで前年度から、レプリカ交換法の一つである REST 法を用いて TM 領域の二量体構造に対するサンプリングを行っていた。今年度も継続して計算を行い、その中で、去年得たシミュレーション結果を踏まえ、計算条件の見直しを行った結果、下記のシミュレーションの結果を得ることに成功した。

なお本課題も、理研・独創的研究課題“脂質の統合的理解”の 1 サブプロジェクト『脂質—膜タンパク質相互作用』に関連する課題の 1 つとして取り組んでいるものである。

## 2. 具体的な利用内容、計算方法

本研究では、FGFR3 TM-JM 領域からなる膜貫通ペプチド二量体を作成し、それを DMPC/DMPG 混合膜からなる脂質二重膜に埋め込んだ系を用意して、シミュレーションを行った。ペプチド二量体の初期構造は PREDDIMER プログラムを使用して作成した。

膜中での二量体構造を効率よくサンプリングするために、ペプチド分子および脂質分子を“溶質”として扱った REST シミュレーションを行った。“溶質”の温度は、310 K - 350 K の範囲で変化させた。WT および G380R 変異体それぞれに対して 230 ns のシミュレーションを行った。上記のシミュレーションおよびその解析は、本研究室で開発しているプログラム“GENESIS”を用いた。

## 3. 結果

REST シミュレーションによって、実験結果と一致する結果を得ることに成功した。すなわち、

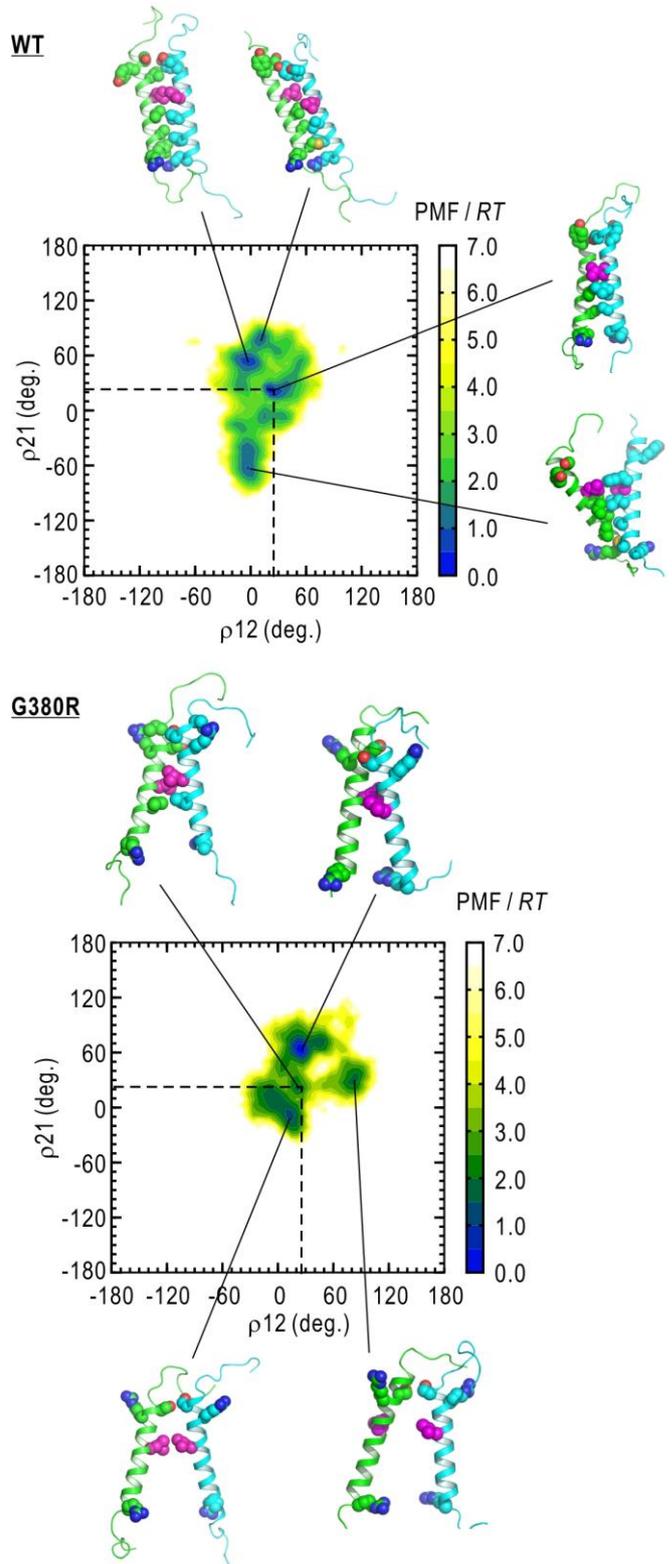
- 1) G380R 変異体の TM ヘリックスの重心位置は、WT のものと比べて 2 Å 程度 z 軸上方に移動していた。
- 2) G380R 変異体のヘリックス慣性主軸の膜方面に対する角度が、WT のものと比べて小さい。
- 3) WT では L385 側鎖が、二量体界面にある確率が高いが、たいして、G380R 変異体では、L385 側鎖が二量体界面にある確率は小さい。

REST シミュレーションで得られた二量体構造の特徴は WT と G380R 変異体で大きく異なっていた。WT では、2 本のヘリックスが平行に並んだ構造 (parallel dimer) もしくは TM 領域の C 末端側で相互作用した二量体 (C-terminal dimer) が有意な構造として現れた。一方、G380R 変異体では、TM 領域の N 末端側で相互作用した変異体 (N-terminal dimer) が主たる構造として現れた (Fig. 3)。この二量体構造の違いは、TM ヘリックスの傾き角や重心位置の違いによって、WT では膜内部に存在している疎水性残基が、変異体では溶媒露出し、そのため細胞質側の JM 領域の相互作用形態が WT と変異体で異なることの結果ではないかと考えている。

## 4. まとめ

膜貫通ペプチド二量体のサンプリングのために、脂

質とペプチドを“溶質”として扱った REST シミュレーションを行い、実験結果と一致する結果を得ることに成功した。現在、この結果は論文としてまとめているところである。



**Figure 3.** REST シミュレーションによってサンプリングされた二量体構造分布およびその代表構造。

## 【課題 B】 細胞環境中の蛋白質ダイナミクス

(担当：優，杉田)

## 【背景】

細胞質中では生体分子が体積の 3 割を占め、蛋白質などは常に隣接する分子と非特異的な相互作用を繰り返している。このような環境の動態を MD 計算で再現する場合、大きな問題となるのが生体高分子の凝集である。通常分子間相互作用パラメータを用いた場合、蛋白質間の引力的相互作用を過剰に評価してしまい、本来起き得ない凝集を生じてしまう。我々は生体高分子と水分子の間に働く分子間力を調節することでこの問題の解決を試みている。現在までに、小型の蛋白質を用いた分子混雑シミュレーション(課題番号 G15029 にて実施)によって、NMR(核磁気共鳴)法で測定された蛋白質の回転緩和データを再現するような相互作用パラメータの絞り込みが出来ている。本課題においては、新しいパラメータを用いて、複数の蛋白質を含む中規模細胞質モデルの MD シミュレーションや、小型蛋白質 villin で構成される分子混雑環境の MD シミュレーションを実行した。

## 【計算・結果】

新たな水-蛋白質相互作用パラメータを用いて、複数の蛋白質と RNA を含む細胞質の中規模モデル(約 100 万原子)(Figure 4)を作成し、GW-MPC を使って、全原子 MD シミュレーションを実行した。計算には、杉田研究室で開発を行っているプログラム”GENESIS”を使用した。従来のパラメータを用いて実行した MD シミュレーションと比較して、蛋白質の凝集や並進・回転の拡散係数がどのように変化するかを調査した。改良されたパラメータを用いた場合においても、時間経過とともに蛋白質の凝集が進むものの、その凝集体形成の速度が大幅に遅くなる事が明らかになった。

蛋白質 villin の分子混雑環境については、蛋白質の柔軟性が凝集体形成に与える影響を調査するために、64 個の villin から構成される分子混雑環境について全原子 MD シミュレーションを実行した(Figure 5)。蛋白質の主鎖立体構造を固定して構造を固くしたモデルと、固定しない通常のフレキシブルモデルを比較した結果、villin の多量体(オリゴマー)の分布状態に大きな差異がある事が明らかになった。これらの結果は、細胞内に生じるタンパク質のドロップレット(液

滴)形成の微視的解明に寄与するものである。

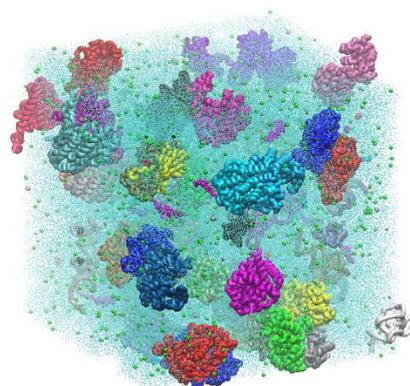


Figure 4. 細胞質の中規模モデル

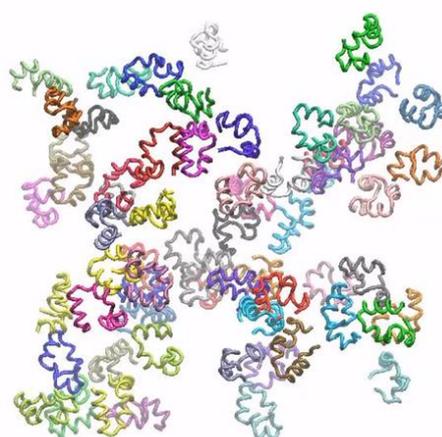


Figure 5. 小型蛋白質 villin による分子混雑環境モデル。蛋白質の主鎖のみを表示。

## 【今後の計画】

改良した水-蛋白質相互作用によって、小型蛋白質 villin で構成される分子混雑環境の動態が大幅に改善されることが確認できたが、細胞質のような、不均一でより混雑率の高い環境ではその効果が弱まる事が明らかになった。今後は細胞質モデルを構成する生体分子の種類や配置、および周期境界の大きさが蛋白質凝集に与える影響を調査する。ドロップレット形成に関しては、villin だけではなく、実験的に細胞内でドロップレット形成が確認されている蛋白質をモデルに形成メカニズムの解明を進める。

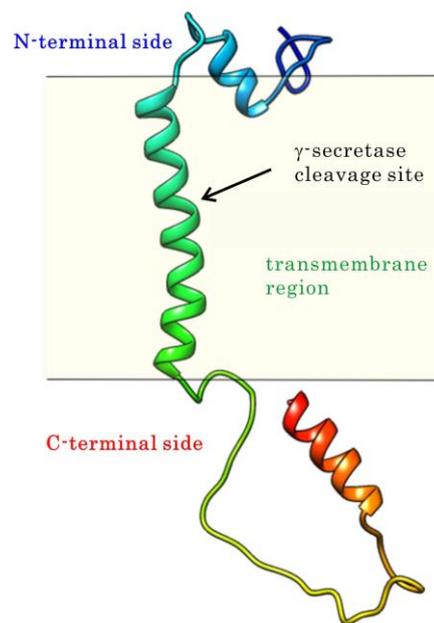
**【課題 C】 Characterizing monomer and dimer structures of the C99 peptide and enhancing simulation methods for complex membrane environments** (担当: Pantelopulos)

**【Background】**

Amyloid beta ( $A\beta$ ) protein is implicated in the onset of Alzheimer's disease (AD) and dementia, and it has been the subject of vast body of experimental, theoretical, and computational work.  $A\beta$  is thought to trigger the AD mechanism via oligomerization and fibrilization.  $A\beta$  is formed from the cleavage of a 99-residue fragment of the Amyloid precursor protein C-terminus (C99). C99 can be cleaved at many sites in its transmembrane (TM) domain by  $\gamma$ -secretase (Figure 6), forming a variety of lengths of  $A\beta$ . Only the longest  $A\beta$  are implicated in AD, but the mechanism of C99 cleavage leading to formation of particular  $A\beta$  lengths are not understood. Because of this, there have been several studies to understand the structure of the TM domain of C99, however, the structure of extramembrane residues has received relatively little attention. The extramembrane domains interact with many cytoplasmic proteins involved in apoptosis signaling, and the N-terminal side of C99 interacts with  $\gamma$ -secretase as well as  $A\beta$  outside cells. Additionally, the structure of C99 and the propensity of C99 cleavage to form long  $A\beta$  fragments are known to be sensitive to the composition and the structure of the cell membrane.

Enhanced sampling methods for rapid equilibration of complex mixtures of molecules in the lipid membrane, which is a process far beyond the capability of current MD simulation methods, is critical to modeling realistic membrane environments. Development and testing of soft-core nonbonded interactions between lipids in cell membrane simulations is a possible approach for enhancing dynamics of lipids by creating more free volume and permitting lipids to more easily move

past each other in the membrane.



**Figure 6. Illustration of C99 structure**

**【Usage of the system and calculation method】**

BW-MPC was used to perform analysis of previously performed implicit membrane MD simulations of C99 in a variety of membrane thicknesses. BW-MPC was used to perform MD simulations of C99 in explicitly detailed lipid bilayers initiated from initial structures found from the analysis of the implicit membrane simulations. Additionally, GW-MPC was used for validation and testing of soft-core nonbonded interaction code in “GENESIS”, also used for aforementioned C99 simulations on BW-MPC.

**【Result】**

A study of full-sequence C99 structure in implicit membrane was completed and has been submitted for peer review. This work provides the first-ever view of the C99 extramembrane domain structural ensembles, which are shown to be intrinsically disordered but sensitive to membrane thickness.

Comparison of equilibrium simulation data against experimental observation is displayed in Figure 7, accompanied by a visualization of C99 structure in

different membranes after coming to equilibrium. This validation of equilibrium ensembles of C99 permitted us confidence in interpreting the intrinsically disordered structural ensembles of extramembrane C99 domains.

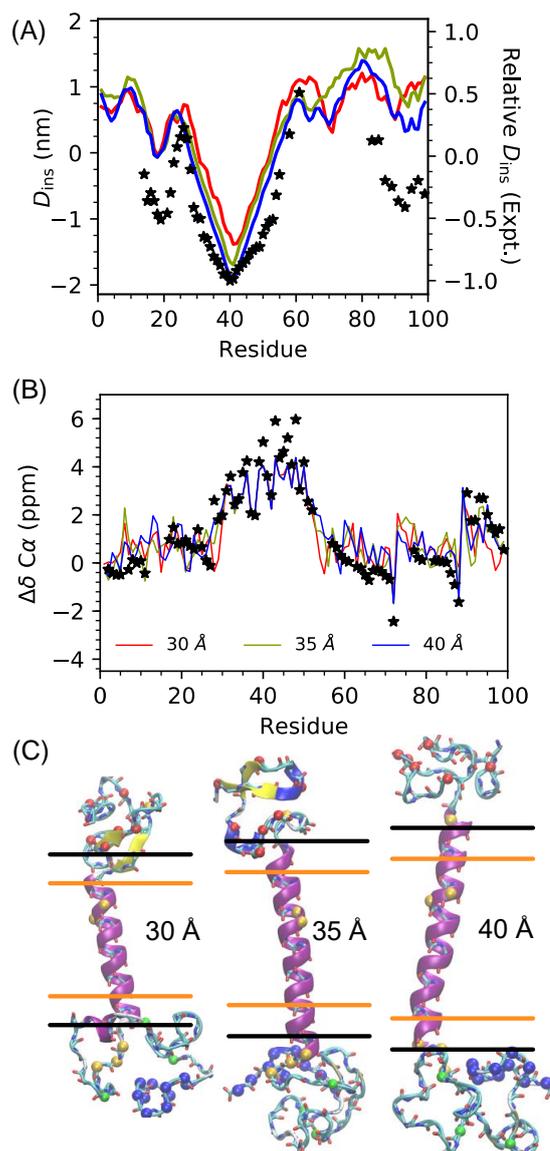
Explicit solvent simulations of C99, initiated from C99 structural states identified in implicit membrane simulations, were tested with a variety of force field parameter sets. The results of this are incomplete, but they suggest that the CHARMM36 force field combined with TIP3P water model are reasonable for achieving ensembles of C99 structures observed in experiments.

Soft-core nonbonded interactions were successfully implemented in “GENESIS”. Modifying nonbonded interactions in this made it possible to achieve a significant enhancement to the dynamics of lipids in model membrane simulations while maintaining lipid bilayer structure.

### 【Conclusion】

We performed the first-ever characterization of full-sequence C99 structure via simulation, revealing structural ensembles of C99 extramembrane domains which reproduce ensemble-averaged experimental measurements on C99 structure.

We also found soft-core nonbonded interactions to provide significant enhancement dynamics in lipid bilayer simulations without disturbing the structure of membranes.



**Figure 7.** (A) Mean and standard deviation residue insertion depths to the membrane defined by  $C\alpha$  positions ( $D_{ins}$ ). Shading represents relative signal for experimental hydrophilic (cyan) and hydrophobic (blue) residue probes, and stars are insertion depths from EPR experiments. (B)  $C\alpha$  chemical shifts at equilibrium predicted using SHIFTX2 and compared with micelle experiments. (C) Representative structures of C99 once equilibrated in implicit membranes of various thicknesses.

### 【Schedule and prospect for the future】

Explicit solvent model simulations of C99 are continuing to be tested, initiated from implicit membrane-derived structures of C99 in early 2018. Discussion of current results with experimental

experts is underway to guide future investigations of C99 structure.

Simulations of the C99 dimer, which is thought to be important to enhancing production of A $\beta$ , will be performed first using implicit membrane models in early 2018, and refined using explicit membrane models in late 2018.

Comparison of soft-core nonbonded potentials in enhancement of lipid bilayer dynamics and effect on structure against other proposed methods for enhancement in complex membrane mixture equilibration will be approached in early 2018. Refinement of the soft-core nonbonded potential code in GENESIS will be performed in late 2018 to prepare for release for use by the general public.

平成 29 年度 利用研究成果リスト

【論文発表】

Po-hung Wang, Isseki Yu, Michael Feig, and Yuji Sugita.

“Influence of protein crowder size on hydration structure and dynamics in macromolecular crowding.”

*Chem. Phys. Lett.*, **671**, 63-70 (2017).

【国際会議での口頭発表】（○が発表者）

Daisuke Matsuoka○, Motoshi Kamiya, Yuji Sugita

“Dimer conformation sampling of FGFR3 transmembrane domain using MD simulation”

会議名： *International seminar on biophysics and chemical biology of biomembrane and lipid bilayers*

場所：大阪大学豊中キャンパス、発表日：2017年10月10日

【国内学会などでの口頭発表】（○が発表者）

松岳 大輔○、神谷 基司、杉田 有治

“Conformational Sampling of FGFR3 transmembrane dimer using REMD simulations.”

会議名：日本物理学会 2017 年秋季大会、

場所：岩手大学、発表日：2017 年 9 月 22 日

松岳 大輔○、神谷 基司、杉田 有治

“分子動力学シミュレーションによる FGFR3 膜貫通ヘリックス部位の構造サンプリング”

会議名：理研シンポジウム 「細胞システムの動態と論理 IX」、

場所：理化学研究所（和光）、発表日：2017 年 4 月 14 日