Vol.2022-BIO-69 No.6
2022/3/10

機械学習によるタンパク質のアポ構造からホロ構造の予測手法の開発

李 明曄[1,a)]　安尾 信明[2]　関嶋 政和[1]

概要：蛋白質の立体構造を基にした薬物設計では、タンパク質のホロ構造が必要だが、ホロ構造の構造決定が難しい。そこで、本研究ではタンパク質のアルファカーボンの距離行列を用いて、アポタンパク質の距離行列からホロタンパク質の距離行列を予測する手法の開発を行った。その結果、学習データに対してはホロ構造の距離行列を生成できたが、テストデータでは生成できておらず、モデルが過学習を起こしている可能性が示唆された。本研究により距離行列を用いたアポ構造からホロ構造を予測する機械学習モデルの可能性が示されたが、さらなる研究が必要である。

キーワード：タンパク質構造予測, holo 構造予測, 機械学習, 距離行列

Development of prediction method of holo-structure from apo-structure of protein by machine learning

Mingyeh Lee[1,a)]　Nobuaki Yasuo[2]　Masakazu Sekijima[1]

Abstract: Although the holo-structure of a protein is necessary for drug design based on its 3D structure, it is difficult to determine the structure of the holo-structure. In this study, we developed a method to predict the distance matrix of the holo protein from the distance matrix of the apo protein using the distance matrix of the alpha carbon of the protein. As a result, we were able to generate the distance matrix of the holo structure for the training data, but not for the test data, suggesting that the model may be overfitting. This study shows the possibility of a machine learning model to predict the holo-structure from the apo structure using the distance matrix, but further research is needed.

Keywords: protein structure prediction, holo structure prediction, machine learning, distance matrix

1. 序論

1.1 背景

生体機能を理解し構造に基づいて新薬を設計するとき、ドッキングシミュレーションにはタンパク質の立体構造が必要である。タンパク質は熱揺らぎにより、結合候補部位を変化させるため、ドッキングシミュレーションの結果はタンパク質の初期構造の影響を強く受ける。標的タンパク質の結合候補部位が狭い場合、多種多様な化合物がドッキングシミュレーションによって得られない可能性が高く、本来結合する化合物が正しく評価されないという問題が存在する。よって、低分子化合物が結合したタンパク質の構造、すなわちホロタンパク質の構造に関する詳細な知識が必要になる。しかし、ホロタンパク質の構造を決定するにはとても難しく、時間とコストがかかる。

しかし、ホロタンパク質の構造を実験的に決定することは難しく、単独のタンパク質、すなわちアポタンパク質の結晶化条件が確立されていても、化合物の結合に伴って受容体の構造が変化することなどによってホロタンパク質に

[1] 東京工業大学情報理工学院
[2] 東京工業大学物質・情報卓越教育院
[a)] lee.m.ad@m.titech.ac.jp

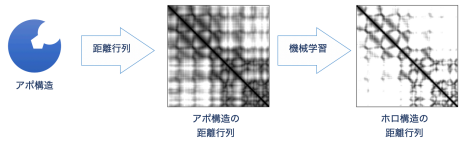


図 1 研究のフローチャート

は適用できない場合がある。分子動力学法によってアポ構造からホロ構造は理論上求められる [1] が、構造探索に膨大な時間がかかる上に、シミュレーションの時系列の中でどれがホロ構造なのかもわからないという問題も存在する。

近年、創薬を加速する研究がたくさん存在している。先行研究において、機械学習を用いたアポ構造のタンパク質から隠れポケットがあるかどうかを予測する機械学習モデル [2] は存在する。また、タンパク質の立体構造からポケットを探索する手法 [3][4] も存在する。しかし、隠れポケットがあるかどうかを予測できたとしても、薬設計において重要であるホロ構造の立体構造がわからないままである。

ただし、アポタンパク質の構造からホロタンパク質の構造を予測する機械学習手法は存在していない。最終的には、タンパク質のアポ構造からホロ構造を直接予測する機械学習手法を開発することを目的とするが、本研究ではタンパク質のアルファカーボンの距離行列を用いて、アポタンパク質の距離行列からホロタンパク質の距離行列を予測する手法の開発を行った。

2. 手法

本章では，本研究で提案する距離行列を用いたアポ構造からホロ構造の予測手法の詳細について提案する。

2.1 概要

今回の研究において、アポ構造やホロ構造はもともと三次元の構造であるが、これを C_α に関する二次元の距離行列に変換し、アポ構造の距離行列からホロ構造の距離行列に予測する問題として扱う。距離行列を使用する理由として、距離行列はタンパク質の回転や平行移動に対して不変であり、また距離行列の次元数も相対的に少ないからである。二次元の距離行列をある種の画像と見なし、画像変換の問題として考える。図 1 に本研究のフローチャートを示している。

2.2 損失関数

ここではモデルの損失関数について説明する。以下入力画像（アポタンパク質の距離行列）を x、Generator の出力画像を $G(x)$、目標画像（ホロタンパク質の距離行列）を y とする。

2.2.1 理論

敵対損失 モデルが真のホロ構造に類似した距離行列が生成できるようにするため、敵対損失を式 1 と定義する。

$$L_{adv} = E_{x,y}[logD(x,y)] + E_x[log(1 - D(x, G(x)))] \quad (1)$$

Generator はこの式 1 をできるだけ小さくしたいが、逆に Discriminator はできるだけこの式 1 を最大化しようとする。

L1 損失 上記の本物らしさだけでなく、出力の距離行列と目標の距離行列の一致具合を測る指標も必要である。そこで、pix2pix[5] の L1 損失の式 2 も追加する。

$$L_1(G) = E_{x,y}[\|y - G(x)\|_1] \quad (2)$$

L1 損失によって出力画像と目標画像の全体像が一致するように学習する。

Generator の対称損失 敵対損失と L1 損失を最小化することによって、Generator は本物みたいな距離行列を生成することができた。しかし、式 1 と式 2 を最小化したとしても、距離行列の対称性を満たしたとは言えない。ここでの距離行列の対称性とは N 点が与えらている場合、距離行列は $N \times N$ の対称行列であり、すなわち自分自身の転置行列と一致するような行列である。この対称性は図 2 に示す。これを満たせるため、自分自身と転置行列が一致するような対称損失を追加する。この項は式 3 によって表記している。

$$L_{G_s}(G) = E[\|G(x) - G(x)^{\mathrm{T}}\|_1] \quad (3)$$

ここでは L1 ノルムを使って生成した距離行列 $G(x)$ とその転置行列である $G(x)^{\mathrm{T}}$ の差を測っている。

Discriminator の対称損失 目標の距離行列が対称行列であり、生成した距離行列も対称行列になるべきであれば、PatchGAN を通した後得られた真偽を判断するロジットの行列（便宜上真偽行列と称す）も対称行列になるはずである。すなわち図 2 において、右上の赤枠が真と判断する場合、左下の赤枠も同じぐらい真になるはずである。しかし、式 1 から式 3 を最小化したとしても前の条件を達成することができない。前述の条件を達成するために、真偽行列もその転置行列が一致するような対称損失を追加する。この項は式 4 によって表記している。

$$L_{D_{sReal}}(D) = E[\|D(x,y) - D(x,y)^{\mathrm{T}}\|_1] \quad (4)$$

ここで式 4 は L1 ノルムを使って本物の距離行列 y の真偽行列 $D(x,y)$ とその転置行列である $D(x,y)^{\mathrm{T}}$ の差を測っている。

損失関数 よって、最終的な損失関数は以上の式を全部足し合わせたものであり、Generator の損失関数は式 5 によって表せ、Discriminator の損失関数は式 6 によって表せる。

$$L_G = L_{adv} + \lambda_1 L_1(G) + \lambda_2 L_{G_s}(G) \quad (5)$$

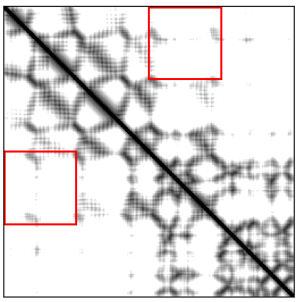


図 2 距離行列

$$L_D = -L_{adv} + \lambda_3 L_{D_{sReal}}(D) \tag{6}$$

ここで λ_1 は L1 損失の強さを調整する項であり、λ_2 は Generator の対称性損失の強さを調整する項であり、λ_3 は Discriminator の本物の距離行列に対する対称性損失の強さを調整する項である。

2.2.2 実装

実際の学習では学習を安定させつつ、より良い結果を導くため、式 1 の代わりに WGAN-GP[6] のように、

$$\begin{aligned} L'_{adv} = E_x[D(x, G(x))] - E_{x,y}[D(x, y)] + \\ \lambda_{gp} E_{\hat{x}\sim p_{\hat{x}}}[(\|\nabla_{\hat{x}}\|_2 - 1)^2] \end{aligned} \tag{7}$$

敵対損失 L_{adv} は式 7 の L'_{adv} の形になる。

よって、最終の Generator と Discriminator の損失関数はそれぞれ

$$L_G = L'_{adv} + \lambda_1 L_1(G) + \lambda_2 L_{G_s}(G) \tag{8}$$

と

$$L_D = -L'_{adv} + \lambda_3 L_{D_{sReal}}(D) \tag{9}$$

になる。$\lambda_1 \sim \lambda_3$ は前述の通りである。

2.3 モデル構造

2.3.1 Generator

Generator はダウンサンプリング用の畳み込み層 2 つ、RseNet[7] の残差ブロック 6 つ、アップサンプリング用のストライドサイズ 2 の逆畳み込み層 2 つからなる。Generator にスペクトル正規化 [8][9] を用いる。スペクトル正規化は元々 Discriminator の方に使用している GAN の学習を安定化させる技術ですが、Self-attention GAN[10] や BigGAN[11] に示しているように、Generator にも使用でき、学習の安定とより良い画像の生成に貢献できる。また、前述のモデルの出力に元々の入力を足した結果が最終の出力になる。すなわち ResNet[7] の残渣ブロックみたいに入力（アポタンパク質の距離行列）と目標（ホロタンパク質の距離行列）の残差を学習している。

2.3.2 Discriminator

Discriminator は pix2pix[5] の Discriminator である PatchGAN を利用している。PatchGAN は出力画像と目標画像の詳細な部分について注目している。2.2.1 で説明した L1 損失によって出力画像と目標画像の全体像が一致するように学習しており、PatchGAN と L1 損失はそれぞれ詳細と全体を分担している。また、PatchGAN の出力が値ではなく行列になっているため、異なるサイズの入力でも異なるサイズの出力行列で対応できる。Discriminator の方もスペクトル正規化 [8][9] を使用している。

2.3.3 全体

ここでは Generator の方も Discriminator の方も全連結層を使用しておらず、全部コンボリューションだけによって構成されている。これにより、入力のサイズを一意に決める必要がなく、ほぼ任意のサイズの画像を取り入れることができる。これは元々取り入れたいタンパク質のサイズも一意に決まっていないため、任意のタンパク質でも対応できるようにこのような構造を取っている。

2.4 推論

距離行列が対称行列であるため、推論するときこの条件を満たせるためモデルから出力した行列に転置行列をを足した上で平均した結果を最後の出力結果とする。$G(x)$ をモデルからの出力とし、Y を最終の出力とすると、

$$Y = \frac{G(x) + G(x)^{\mathrm{T}}}{2} \tag{10}$$

を推論時の最終出力とする。

3. 実験および考察

本章では、モデルを用いてホロ構造予測の評価実験の手法およびその結果について詳説する。

3.1 データセット

3.1.1 データ

データセットは、先行研究 [2] の Supporting information の Table S1 と Table S5 を参照し、それぞれ学習データとテストデータのアポ構造とホロ構造のタンパク質を Protein Data Bank(PDB) からダウンロードして編集し、タンパク質構造データセットを構築した。

具体的な編集内容として、3D 構造が比較できる DALI Server[12][13] のペアワイズ構造比較（Pairwise structure comparison）を用いて、アポタンパク質とホロタンパク質の 3D 構造を比較する。チェーンの長さが揃っていない構造対に対し、C 末端または N 末端だけ揃っていない場合は多めになった部分だけを削除し、途中で長さが揃っていない場合はその構造対を使用しないようにする。また、チェーンの長さが 500 以上の構造対も使用しない。

以上の編集を実行した結果、学習に用いるタンパク質が

36 個、テストタンパク質が 6 個になる。

3.1.2 モデルの入出力

以上のタンパク質に対して、C_α を用いて距離行列を作る。目標のホロ構造の距離行列に対し AlphaFold 1[14] を参考して 22Å までの距離を使用し、22Å 以上の距離を 22Å とする。入力に対しできるだけ多くの情報を与えるため、入力のアポ構造の距離行列に対して 2 倍の 44Å までの距離を使用し、44Å 以上の距離は 44Å とする。これらを実際に使用している距離行列とする。

距離行列の i 行 j 列の数字の意味とは i 番目の C_α と j 番目の C_α の間の距離を意味する。また、j 行 i 列も j 番目の C_α と i 番目の C_α の間の距離を意味するため、距離行列は対称行列である。

距離行列の例を図 2 に示している。図 2 において、i 行 j 列が黒ければ黒いほど、その値が小さくなり、両者間の距離が小さく、コンタクトしていると考えられる。また、逆に i 行 j 列が白ければ白いほど、その値が大きくなり、両者間の距離も大きく、コンタクトしていないと考えられる。

3.1.3 データ拡張（data augmentation）

データ拡張（data augmentation）として、2 種類の手法を使用している。一つはアポタンパク質に対して分子動力学法（Molecular Dynamics, MD）を用いシミュレーションでアポタンパク質の数を拡張する。具体的な MD 条件は表 1 に示す。MD 実行によって得られたトラジェクトリー中 500ps ずつの構造を使用し、これにより 1 つのアポ構造から 1000 個のアポ構造を近似的に作り出す。これにより、学習データは $36 \times 1000 = 36000$ 個になる。学習のときは 1epoch 中に 10 フレームずつ、5000ps ずつの構造を使って距離行列を生成しモデルに学習を行う。すなわち 1000 個の構造の中で等間隔で構造を取り出し、これによって 1epoch の学習データは $36 \times 100 = 3600$ 個になる。

表 1 MD 条件

計算条件	詳細
programe	Amber
Temperature	300 K
Pressure	1.01325 bar
Force field	FF19SB
time	500 ns

もう一つは学習時生成した距離行列をランダムで 180° 回転する [15]。これはすなわちタンパク質の C 末端と N 末端を逆転してから距離行列を生成することと同じである。

3.2 学習で用いたパラメータ

学習に用いるパラメータは表 2 にまとめている。Generator および Discriminator の学習率は、Two Time-scale Update Rule[16] を使用し、それぞれ異なる学習率を設定し、Discriminator が 1 回学習するたびに Generator も 1 回学習する。

表 2 パラメータ

パラメータ	詳細
最適化手法	Adam[17]
Generator の学習率	0.0001
Discriminator の学習率	0.0002
バッチサイズ	1
学習回数	100

また $\lambda_1 \sim \lambda_3$ と λ_{gp} は以下の通りである。

- $\lambda_1 = 50$
- $\lambda_2 = 50$
- $\lambda_3 = 50$
- $\lambda_{gp} = 10$

3.3 評価手法

生成した距離行列、目標のホロ構造の距離行列およびアポ構造の距離行列間の平均絶対誤差（MAE）をそれぞれ測る。C 末端または N 末端にゆらぎがあるかもしれないので、MAE を計算するとき最初の C_α および最後の C_α は計測に入らないようにする。すなわち距離行列の最初行と最初列、最後行および最後列は MAE の計測に入れないようにする。

3.4 全体の実験結果

ここからは評価実験の結果を詳説する。

本論文中に出現している 1BP5A や 1NUWA などの前 4 文字が PDB id であり、最後の 1 文字が Chain になっている。すなわち 1BP5A に対し、PDB id が 1BP5 のタンパク質の chain A を意味している。

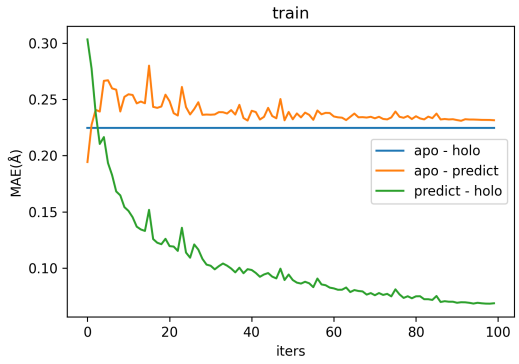
実験に対し、目標として以下の 2 つがあげられます。

(1) アポ構造の距離行列（以下アポ距離行列と称する）とホロ構造の距離行列（以下ホロ距離行列と称する）間の平均絶対誤差（MAE）はできるだけアポ距離行列と予測の距離行列（以下予測距離行列と称する）間の MAE に近づけたい
(2) 予測距離行列とホロ距離行列間の MAE はできるだけ 0 に近づけたい

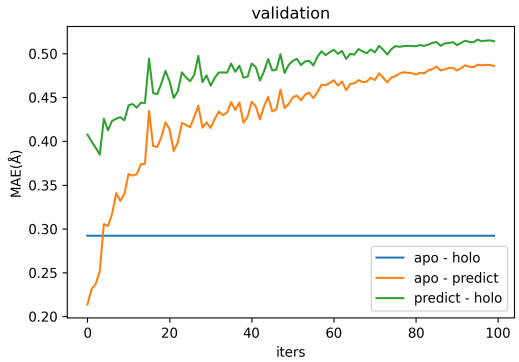
学習の結果は図 3(a) に示し、予測の結果は図 3(b) に示す。図 3(a) と図 3(b) の中で、縦軸は MAE（Å）、横軸は学習回数である。

図 3(a) の学習結果について見てみると、学習回数が増えるとともに予測距離行列とホロ距離行列間の MAE は 0 に近づけ、アポ距離行列と予測距離行列間の MAE はアポ距離行列とホロ距離行列間の MAE に近づけている。よって、学習データについては目標に達していると言える。

図 4(a) と図 4(b) は実際に使用しているアポ距離行列とホロ距離行列および予測距離行列の一部である。左から順

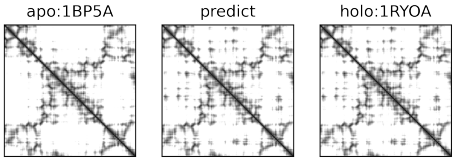


(a) 学習結果

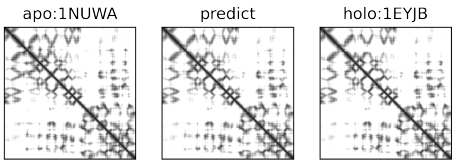


(b) テスト結果

図 3 (a)：学習結果、(b)：テスト結果



(a) コンタクトしていない → している



(b) コンタクトしている → していない

図 4 (a)：コンタクトしていない → している、(b)：コンタクトしている → していない

にアポ距離行列、予測距離行列、ホロ距離行列になる。まず図 4(a) を見ると、アポ構造の中でコンタクトしていない部分（白の部分）がホロ距離行列の中でコンタクトしている（黒の部分）ようになっている。実際に予測距離行列の中でも黒くなっている（コンタクトしている）。図 4(b) では逆に元々アポ距離行列の中でコンタクトしている部分（黒の部分）がホロ距離行列の場合コンタクトしていないようになっている（白になる）。この場合予測距離行列の中でも白になっている（コンタクトしていない）。

しかし、予測データについての結果（図 3(b)）を見ると、学習回数が増えるとともに予測距離行列とホロ距離行列間の MAE もアポ距離行列と予測距離行列間の MAE も段々増えるようになり、これは明らかにモデルが過学習を起きていると考えられる。

3.5 過学習の原因

図 3(b) よりモデルが過学習が起きていることが確認できた。過学習の原因の 1 つとして、学習データ量の不足があげられる。本研究において元々学習用のデータは 39 対しかなく、ディープラーニングとしてのデータがとても少ないため、データの水増しとして 3.1.3 で説明した分子動力学法を実行した。その結果、全学習データの量が 36000 になり、ディープラーニングのデータ量としては成り立つと考えられる。しかし、図 3(b) などからも確認できるように過学習が起きている。そこで、考えられる 1 つの原因として、データの水増しによって得られた新しいデータが元のデータ間からの変化が不十分と考えられる。実際に PDB からダウンロードしたアポ構造の距離行列と MD によって得られた構造の距離行列間の平均 MAE をプロットした図 5 を見ると、アミノ酸の数が増えるにつれて、平均 MAE が段々減少していることが確認できる。また、一番高い MAE の値でも 1.2 未満になっている。よって、MD によって得られた構造は元のアポ構造から十分に揺らいでいないと考えられ、学習データの中で似たデータが多く存在し、過学習が起きてしまい、汎化性能の向上に貢献できなかった。

4. 結論

本研究では、タンパク質のアポ構造とホロ構造の距離行列を画像として見なし、GAN を用いた画像変換手法を用いて学習することにより、アポ構造からホロ構造を予測する手法の開発を目指した。具体的に言うとタンパク質のアルファカーボンの距離行列を二次元の画像と見なし、GAN を用いた画像変換手法を用いて構造間の距離行列の変換手法を提案した。GAN の Generator の部分に ResNet の残差ブロックを用い、Discriminator に PatchGAN を用いた手法を開発した。また、GAN の敵対損失と pix2pix[5] な

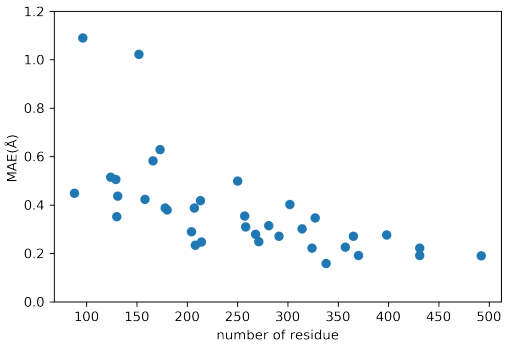


図 5 元のアポ距離行列と MD 結果のアポ距離行列間の平均 MAE とアミノ酸数の関係

どで使っている L1 損失以外に、距離行列の対称性を考慮し対称損失を提案した。

先行研究 [2] で使っているデータから編集し、総計 36 個の学習データを得られた。ディープラーニングのデータ量としては少ないため、分子動力学法を用いてデータの水増しを試みた。これにより、全部で 36000 個のデータが得られ、これらのデータを用いて学習を行った。また、6 個のテストデータに対して評価を行った。

その結果、学習データについての予測は目標に達しているが、テストデータの結果を見るとモデルが過学習を起きていることが分かった。過学習の理由について考察し、分子動力学法によって得られた構造は元の構造との差が小さく、十分揺らいでいないから、モデルの入力データに似ているデータが大量に存在し、過学習が起きたことが考えられる。以上から、この手法の可能性を示唆しているが、さらなる研究が必要とされていることが分かった。

また、今回の研究ではアポ構造から距離行列を生成し、そのアポ距離行列からホロ距離行列を予測するまでになっているが、将来的にはそのホロ距離行列を用いてタンパク質を構築することが考えられる。距離行列からタンパク質の 3D 構造を構築する先行研究として GDFuzz3D[18] が存在するため、ホロ距離行列からホロタンパク質を構築する可能性が存在すると考えられる。将来的には距離行列からタンパク質に戻したのち、既存のタンパク質構造予測手法との比較も検討している。

参考文献

[1] Daniel Seeliger and Bert L De Groot. Conformational transitions upon ligand binding: holo-structure prediction from apo conformations. *PLoS computational biology*, Vol. 6, No. 1, p. e1000634, 2010.

[2] Peter Cimermancic, Patrick Weinkam, T Justin Rettenmaier, Leon Bichmann, Daniel A Keedy, Rahel A Woldeyes, Dina Schneidman-Duhovny, Omar N Demerdash, Julie C Mitchell, James A Wells, et al. Cryptosite: expanding the druggable proteome by characterization and prediction of cryptic binding sites. *Journal of molecular biology*, Vol. 428, No. 4, pp. 709–719, 2016.

[3] Manfred Hendlich, Friedrich Rippmann, and Gerhard Barnickel. Ligsite: automatic and efficient detection of potential small molecule-binding sites in proteins. *Journal of Molecular Graphics and Modelling*, Vol. 15, No. 6, pp. 359–363, 1997.

[4] Takeshi Kawabata. Detection of multiscale pockets on protein surfaces using mathematical morphology. *Proteins: Structure, Function, and Bioinformatics*, Vol. 78, No. 5, pp. 1195–1211, 2010.

[5] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1125–1134, 2017.

[6] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron Courville. c. *arXiv preprint arXiv:1704.00028*, 2017.

[7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[8] Yuichi Yoshida and Takeru Miyato. Spectral norm regularization for improving the generalizability of deep learning. *arXiv preprint arXiv:1705.10941*, 2017.

[9] Takeru Miyato, Toshiki Kataoka, Masanori Koyama, and Yuichi Yoshida. Spectral normalization for generative adversarial networks. *arXiv preprint arXiv:1802.05957*, 2018.

[10] Han Zhang, Ian Goodfellow, Dimitris Metaxas, and Augustus Odena. Self-attention generative adversarial networks. In *International conference on machine learning*, pp. 7354–7363. PMLR, 2019.

[11] Andrew Brock, Jeff Donahue, and Karen Simonyan. Large scale gan training for high fidelity natural image synthesis. *arXiv preprint arXiv:1809.11096*, 2018.

[12] Liisa Holm. Using dali for protein structure comparison. In *Structural Bioinformatics*, pp. 29–42. Springer, 2020.

[13] DALI PROTEIN STRUCTURE COMPARISON SERVER. `http://ekhidna.biocenter.helsinki.fi/dali/`.

[14] Andrew W Senior, Richard Evans, John Jumper, James Kirkpatrick, Laurent Sifre, Tim Green, Chongli Qin, Augustin Žídek, Alexander WR Nelson, Alex Bridgland, et al. Improved protein structure prediction using potentials from deep learning. *Nature*, Vol. 577, No. 7792, pp. 706–710, 2020.

[15] Shaun M Kandathil, Joe G Greener, and David T Jones. Prediction of interresidue contacts with deepmetapsicov in casp13. *Proteins: Structure, Function, and Bioinformatics*, Vol. 87, No. 12, pp. 1092–1099, 2019.

[16] Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems*, Vol. 30, , 2017.

[17] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[18] Michal J Pietal, Janusz M Bujnicki, and Lukasz P Kozlowski. Gdfuzz3d: a method for protein 3d structure reconstruction from contact maps, based on a non-euclidean distance function. *Bioinformatics*, Vol. 31, No. 21, pp. 3499–3505, 2015.