

# 分子軌道認証:データ科学を 利用した簡便な反応電子論

### 常田貴夫 神戸大学科学技術イノベーション研究科 北海道大学理学部化学科

2021年11月17日 @神戸大学計算科学教育センター

### 生命とは何か?

[ポール・ナース著、『生命とは何か』(ダイヤモンド社)、2021年.]

生命は、高度に秩序化され、組織化された、並外れた形態 の化学である.

生命のほとんどの現象は、「酵素が触媒する化学反応」 の観点から理解するのが一番わかりやすい。

生命体で発生する膨大な化学反応のことを「代謝」とよぶ. それは生きているものが行なうすべての基礎であり、 プロセスを促進するのに必要なエネルギー源.

細胞周期は、非常に複雑な反応だが、**すべてが化学反応** に基づくという生命の重大な側面を明らかにする。

新しい細胞を作り出すのに必要な何百もの反応すべてが 連携し、ある目的を実行するシステムを築くとき、初めて 命が始まる.







Paul M. Nurse 細胞周期遺伝子cdc2 発見でノーベル生理 学・医学賞受賞(2001)

### 化学反応とは何か?



### 有機電子論

Lewis価電子理論:オクテット則で化学結合を解釈 [Lewis, J. Am. Chem. Soc. <u>38</u>, 762, 1916.] F H F H



電子親和性が化学結合を介して他原子に静電的に影響 [Robinson, J. Chem. Soc. <u>129</u>, , 1926.]

官能基が電子求引性、電子供与性を持ち、その間の電子移動で結合形成 [Ingold, Chem. Rev. <u>15</u>, 225, 1934 .]



その後、有機実験による検証を経て洗練化

#### 化学反応とは、電子を渡しやすい部位から電子を受け取りやすい部位 への電子遷移と経験的に推論

### フロンティア軌道論

[Fukui, Yonezawa, and Shingu, J. Chem. Phys. <u>20</u>, 722, 1952.]

フロンティア軌道論ダイアグラム



## ウッドワード・ホフマン則

[Woodward & Hoffmann, J. Am. Chem. Soc. <u>87</u>, 395, 1965; ibid <u>87</u>, 4389, 1965.]



#### **軌道の位相で反応方向が決まる**と発見的に推論

コンセプチュアルDFT

フロンティア軌道論の理論的基礎づけ [Nalewajski & Parr, J. Chem. Phys. <u>77</u>, 399, 1982] DFTの基礎定理を統計力学的に再解釈



化学反応の進行方向=核座標変化で反応エネルギーが最大 の方向=占有数一定で電子密度変化が最大の方向

福井関数:フロンティア軌道論の反応性指標

- 求電子反応:  $f^- = (\partial \rho_{\text{HOMO}} / \partial n)_V \approx \rho_{\text{HOMO}}(\mathbf{r})$
- 求核反応:  $f^+ = (\partial \rho_{\text{LUMO}} / \partial n)_V \approx \rho_{\text{LUMO}}(\mathbf{r})$

ラジカル反応:  $f^0 = (\partial \rho_{\text{SOMO}} / \partial n)_V$ ≈  $[\rho_{\text{HOMO}}(\mathbf{r}) + \rho_{\text{LUMO}}(\mathbf{r})]/2$ 

#### 反応性指標を統計力学的に提案する理論体系に

# 分子軌道画像化技術



図. 分子軌道トモグラフィによるCO<sub>2</sub> 分子 のHOMOの分子軌道画像の理論計算との 比較

[Vozzi et al., Nature Phys. <u>7</u>, 822, 2011] DFT計算 実験



### 分子軌道の実験的画像化が可能に

図. 多次元電子スペクトルによるN<sub>2</sub>分子の 分子軌道画像の理論計算との比較 [Yamazaki et al., Phys. Rev. A, <u>75</u>, 032721, 2007.]

図. PTCDAの占有・非占有分子軌道イメージ [Puschnig et al., Phys. Rev. B <u>7</u>, 235427, 2011]



### 外圏型電子移動反応の軌道論解釈

[Kellet et al., Nature Commun. 9, 4916, 2018.]

#### 外圏型電子移動反応はフロンティア軌道論的に制御可能



### ダイソン軌道:多電子波動関数の分子軌道

[Krylov, J. Chem. Phys. , <u>153</u>, 080901, 2020.]

多参照波動関数など多電子波動関数の軌道図を得ることが可能

ダイソン軌道:多電子波動関数の1電子関数

 $\phi^d(1) = \sqrt{N} \int \Psi_I^N(1,\ldots,n) \Psi_F^{N-1}(2,\ldots,n) d2 \ldots dn$ 

=  $\sum_{p} \langle \Psi_{I}^{N} | p^{\dagger} | \Psi_{F}^{N-1} \rangle \phi_{p}(1),$  イオン化状態の多電子波動関数 p番目の軌道の生成演算子 中性状態の多電子波動関数

図. ホルムアルデヒドの光イオン化収率



図. 開殻系SrOHのSOMOのダイソン軌道  $X^2\Sigma^+ A^2\Pi = B^2\Sigma^+ A'^2\Delta$ 



#### 軌道エネルギーを直接算出できないので、 観測可能量と関係づけて軌道を再定義

### 高精度軌道エネルギー計算の実現



### 交換汎関数の長距離補正

交換汎関数の長距離相互作用誤差の補正

[likura, Tsuneda, Yanai and Hirao, J. Chem. Phys. 115, 3540, 2001.]



$$E_{x}^{lr} = -\frac{1}{2} \sum_{\sigma} \sum_{i}^{occ} \sum_{j}^{occ} \int d^{3}\mathbf{r}_{1} d^{3}\mathbf{r}_{2} \phi_{i\sigma}^{*}(\mathbf{r}_{1}) \phi_{j\sigma}^{*}(\mathbf{r}_{2}) \frac{\operatorname{erf}(\mu r_{12})}{r_{12}} \phi_{j\sigma}(\mathbf{r}_{1}) \phi_{i\sigma}(\mathbf{r}_{2})$$

### 長距離補正をもとに、さまざまな汎関数が開発されてきた:

CAM-B3LYP [Yanai, Tew, & Handy, Chem. Phys. Lett. <u>91</u>, 551, 2004.] LC-ωPBE [Vydrov, Heyd, Krukau, & Scuseria, J. Chem. Phys. <u>125</u>, 074206, 2006.] BNL [Livshits & Baer, Phys. Chem. Chem. Phys. <u>9</u>, 2937, 2007.] ωB97-series [Chai & Head-Gordon, J. Chem. Phys. <u>128</u>, 084106, 2008.] Tuned RSH [Baer, Livshits, & Salzner, Ann. Rev. Phys. Chem. <u>61</u>, 85, 2010.]

# 長距離補正が解決・改善した問題



### 銀ナノクラスタの電子状態



軌道エネルギーの物理的意味

#### 分数占有数に対するエネルギー直線性定理

[Perdew et al., Phys. Rev. Lett. <u>49</u>, 1691, 1982; Yang et al., Phys. Rev. Lett. <u>84</u>, 5172, 2000.]

$$E\left(n+\frac{p}{q}\right) = \frac{p}{q}E(n+1) + \frac{q-p}{q}E(n)$$

全電子エネルギーは占有数の変化に ついて直線的に変化する

<u>ヤナクの定理</u> [Janak, Phys. Rev. B 103, 7165, 1978.]



### 化学反応と分子軌道との関係の再考





化学反応とは、軌道の安定化により駆動される構造変化

### 反応性軌道エネルギー論解析法

[Tsuneda & Singh, JCC, <u>35</u>, 1093, 2014; Tsuneda, Singh & Chattaraj, PCCP, <u>20</u>, 14211, 2018.] 長距離補正(LC)DFTの定量的軌道エネルギーにより可能になった反応解析法

ディールス・アルダー反応:  $C_2H_4 + C_4H_6 \rightarrow C_6H_{10}$ 

#### **反応経路上の軌道エネルギー** の自動追跡 → 反応性軌道の特定

規格化反応ダイアグラムの構築と 軌道エネルギーギャップ勾配の計算





# グローバル反応経路地図の反応性軌道

[Hasebe, Tsutsumi, Taketsugu, & Tsuneda, J. Chem. Theory Comput., <u>17</u>, 6901, 2021.]

図. ホルムアルデヒド分子の異性体間を結ぶ反応のグローバル反応経路地図



解離HCO + Hからの反 応はHOMOとLUMOが 駆動するものが多い が、HOCHを形成し、 最安定なH<sub>2</sub>COを生成 する反応経路は存在 しない

軌道論的に阻害?

単一の分子からの反応では、反応性軌道対が一対一対応

# 反応性軌道による電子論ダイアグラム

[Hasebe, Tsutsumi, Taketsugu, & Tsuneda, J. Chem. Theory Comput., <u>17</u>, 6901, 2021.]

図. ホルムアルデヒド分子の異性体間を結ぶ反応性軌道の変化とそれにより提案される 電子論ダイアグラム(赤点線内が電子移動駆動、青点線内がダイナミクス駆動).



# 反応性軌道と反応経路との一対一対応

[Hasebe, Tsutsumi, Taketsugu, & Tsuneda, J. Chem. Theory Comput., <u>17</u>, 6901, 2021.]

グローバル反応経路地図(GRRM)上の占有・非占有反応性軌道ペア

図. グリシン異性化反応経路の反応性軌道(口が電子移動駆動、口がダイナミクス駆動)



# ROETと有機電子論の反応解釈比較

[Hasebe, Tsutsumi, Taketsugu, & Tsuneda, J. Chem. Theory Comput., <u>17</u>, 6901, 2021.]



### 分子軌道認証

#### 反応性軌道エネルギー ダイアグラム

[Tsuneda, Singh & Chattaraj, Phys. Chem. Chem. Phys., <u>20</u>, 14211, 2018.]



反応性軌道エネルギー論

#### 最大の問題:高コストのIRC計算 が必要であること

### 反応性軌道エネルギー論の 解析からIRC計算を省く

反応経路IRC上の 自動軌道認証 →学習データに

分子軌道認証

自動電子論解析

機械学習で反応物・ 生成物のMOを認証



### データ科学の理論科学への応用

[Rupp, von Lilienfeld & Burke, J. Phys. Chem. <u>148</u>, 241743, 2018.]



#### DFT は 演繹 的 デ ー タ 科 学

[Tsuneda, The Chemical Record, <u>19</u>, 1, 2019.]

#### <u>データ科学における予測モデルの決定プロセス</u>

- 1. 予測に寄与する記述子を収集
- 2. パラメータを使ったデータ科学的モデルで予測関数を決定
- 3. 汎化性能を上げるために特徴量の種類を増やしてモデルを洗練化
   →過学習を避けるために特徴量を削減

<u>DFT汎関数の開発プロセス</u>

- 1. 系の電子状態に寄与する物理的効果を収集
- 2. 半経験的パラメータを使った物理的効果を結合したモデルで汎関数を決定
- 3. 適用可能性を改善するために物理的効果の種類を増やして汎関数を洗練化 →ポテンシャルの偽の極小点を避けるため、半経験的パラメータの数を削減

#### DFT汎関数開発は化学において最も成功したデータ科学

**帰納的** End to End



Origin to End

### 軌道認証法1: 予測関数の記述子

機械学習の記述子: 互いに独立でないと予測精度も汎化性能も低くなる 完全に独立な変数のみ採用

#### 各原子種の各AOのMO係数

 $\begin{array}{c} \text{ns, np}_{\text{x}}, \text{np}_{\text{y}}, \text{np}_{\text{z}}, \text{nd}_{\text{-2}}, \text{nd}_{\text{-1}}, \text{nd}_{0}, \text{nd}_{\text{+1}}, \text{nd}_{\text{+2}}, \\ \text{nf}_{\text{-3}}, \text{nf}_{\text{-2}}, \text{nf}_{\text{-1}}, \text{nf}_{0}, \text{nf}_{\text{+1}}, \text{nf}_{\text{+2}}, \text{nf}_{\text{+3}} \end{array}$ 

#### 主MO係数の原子の番号

各MOの1、2、3番目の大きさの MO係数を持つAOの原子番号





**軌道エネルギー**: 認証性能を高めるために 標準化した軌道エネルギー

### 軌道認証法2: 認証予測モデル



学習プロセス: プログラム言語: Python ライブラリ: scikit learn

5種の予測モデル: 重み (sample-weight) 対応 (True) → 10 非対応 (False) → 1



学習反応系: 43反応(広範な典型反応) +18反応(有機反応) IRC計算により特定したMOの組

## 軌道認証法3:予測による認証

機械学習予測モデルと 使った軌道認証

#### 1. 機械学習認証:予測モデル

を使い、各生成物MOに対応す る反応物MOを特定.2つ以上の MOが特定される場合は、3つま で記録.

2. 軌道エネルギー認証:異なる MOについて2つ以上のMOが特 定される場合,軌道エネルギー の最も近いMOを選択.

3. 最終認証:残りのMOは,特 定されてないMOのうち軌道エネ ルギーの近いMOに一旦特定.2 に戻る(3回繰り返し).

#### 図. 軌道認証のアウトプット例

Model in complete form:

RandomForestClassifier(bootstrap=True, class\_weight=None, criterion='gini',

 $max\_depth=None,\ max\_features='auto',\ max\_leaf\_nodes=None,$ 

min\_impurity\_decrease=0.0, min\_impurity\_split=None,

min\_samples\_leaf=1, min\_samples\_split=2,

- min\_weight\_fraction\_leaf=0.0, n\_estimators=100, n\_jobs=None,
- oob\_score=False, random\_state=None, verbose=0,

warm\_start=False)

<class 'numpy.ndarray'>

Orbital recognition evaluation:

001000000001000100101 Score of reaction 28: 0.9649122807017544 Orbital recognition in machine learning: [[0 1 2 -1 4 5 6 7 8 -1 -1 -1 12 -1 14 15 16 17] [1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1] First revision of orbital recognition: [0 1 2 -1 4 5 6 7 8 -1 -1 -1 12 -1 14 15 16 17] Second revision of orbital recognition: [0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17] Third revision of orbital recognition: [0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17] Fourth revision of orbital recognition: [0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17] Final orbital recognition result: [0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17] Reactant molecules Reactant complex Product complex

 MO
 (Mol.) Orb. Ene.
 MO
 Orb. Ene.
 MO
 Orb. Ene.

 Reac. occ. MO (forward):
 HOMO-3 (#1) -0.7207 -> HOMO-4 -0.6991 -> HOMO-4 -0.8326
 HOMO-3 (#1) 0.1662 -> LUMO+9 0.1754 -> LUMO+9 0.2210

 Reac. occ. MO (forward):
 LUMO+8 (#1) 0.1662 -> LUMO+9 0.1754 -> LUMO+9 0.2210
 HOMO (#2) -0.4498 <- HOMO -0.4503 <- HOMO -0.3517</td>

 Reac. unocc. MO (backward):
 HOMO (#1) 0.0461 <- LUMO 0.0558 <- LUMO 0.0238</td>

### 分子軌道認証結果

軌道認証率 反応錯体↔ 生成錯体 100%に近い認証:99.6%(42典型反応)、98.2%(61反応) ランダムフォレストがベスト 反応分子↔ 生成錯体 反応分子↔ 生成分子 実用上問題ないレベル:イオン反応の非占有軌道の認証率は低い 決定木がベスト

#### 分子軌道認証が実現すること 反応性軌道エネルギー論(ROET)ダイアグラムの自動生成











占有反応性軌道変化と電子移動方向





# 占有反応性軌道変化と電子移動方向



# 占有反応性軌道変化と電子移動方向







 $\stackrel{\leftarrow}{\longleftrightarrow}$ 

占有反応性軌道変化は 反応全体における 本来の電子遷移を明らかに

フロンティア軌道論的な ダイアグラム

有機電子論解析の矢印方向 と似ているが 必ずしも一致しない

> 有機電子論 解析と一致し ない反応の例

マリケンポピュレーション解析は正しいか?

図. 実際の電子移動方向とマリケンポピュレーション解析の結果の比較



# 反応性軌道



#### 反応が進行する条件

- 1. 反応を駆動する電子遷移を導くように、 反応性軌道間の交換相互作用が大きい
- 2. 結合間の電子遷移がスムーズ
   →結合のない原子間の電子遷移は進行しにくい
- 3. 結合を連続的に形成しながら原子が移動 →離れた原子間の原子移動は進行しにくい

#### <u>反応性軌道の特徴</u>

反応性軌道は結合性軌道でも
 反結合性軌道でもなく、
 電子遷移しやすく
 結合形態の異なる軌道対





- 水素移動は電子移動を伴う 1.
- **軌道位相**は電子移動後に維持される傾向に 2. ある
- 分子内電子移動と分子内水素移動は結合を 3. 通じて同じくらい頻繁に起こる
- 電子分布の局所化が解離反応を駆動し、非 4. 局所化が結合反応を進行させる
- 環状のみならず環状様構造でも、電子は結合 5. 交代をもたらしながら回転する

### より詳しく知りたい方へ



<u>英語版も</u> <u>"T. Tsuneda, Density Functional</u> <u>Theory in Quantum Chemistry</u> <u>(Springer, 2014)"</u>

### <u>"常田,『密度汎関数法の</u> <u>基礎』(講談社, 2012)"</u>

