

DOI:10.13718/j.cnki.xsxb.2021.10.013

3 种 RGS 蛋白社区网络的研究^①

白艺昕¹, 李 勇², 刘 勇³

1. 西南大学 数学与统计学院, 重庆 400715; 2. 创伤骨科 桂阳县第一人民医院, 湖南 郴州 424400;
3. 中科院广州能源研究所, 广州 510640

摘要: RGS 蛋白家族(regulator of G protein signaling protein family)通过与 G 蛋白的 $G\alpha$ 链的相互作用以及水解 GTP, 在调控细胞信号传导中起着重要作用. RGS 蛋白家族中各蛋白结构上非常相似, 具有一个保留约 120 个氨基酸的 RGS 盒子. 这种同源性的特征使得药物具有很低的选择性以及高度的药理相似性. 在本文中, 我们使用分子动力学模拟研究了 RGS4、RGS8 以及 RGS17 这 3 种 RGS 蛋白, 并使用动力学数据将 RGS 蛋白划分为不同的社区网络, 进而深入理解了药理相似性以及给出了潜在的药物选择性.

关键词: RGS 蛋白; 社区网络; 分子动力学; 变构调节

中图分类号: Q615 **文献标志码:** A **文章编号:** 1000-5471(2021)10-0085-06

On Community Network of Three RGS Proteins

BAI Yixin¹, LI Yong², LIU Yong³

1. School of Mathematics and Statistics, Southwest University, Chongqing 400715, China;
2. Orthopaedic Trauma, The First People's Hospital of Guiyang, Chenzhou, Hunan 424400, China;
3. Guangzhou Institute of Energy Conversion, Chinese Academy of Sciences, Guangzhou 510640, China

Abstract: Regulator of G protein signaling (RGS) protein family plays an important role in regulation of signal transduction through interacting with $G\alpha$ subunit of G protein and hydrolysis of GTP. RGS proteins are structurally similar with a conserved RGS-box containing approximately 120 amino acids. The homologous feature leads to a low selectivity of drugs and high pharmacology. In this study, molecular dynamics (MD) simulations have been employed to study the dynamics of RGS4, RGS8, and RGS17, partition structure of RGS proteins into community network utilizing the MD data, and to provide insights into pharmacology and potential selectivity.

Key words: RGS protein; community network; MD simulations; allosteric regulation

信号传导是基本的生物途径, 可将外源信息的变化转化为细胞体内功能的改变^[1]. 它控制着包括神经介质和荷尔蒙的释放、调节膜离子通道的开启与关闭、基因转录等重要的细胞活动^[2]. 由 $G\alpha$ 、 $G\beta$ 以及 $G\gamma$ 这 3 条不同链构成的 G 蛋白控制着生物体内重要的信号传导^[1]. 当外源信息发生变化时, 与 GDP 结合的

① 收稿日期: 2020-08-25

基金项目: 重庆自然科学基金项目(cstc2019jcyj-msxm2393); 中央高校基本科研业务费(XDJK2015C159).

作者简介: 白艺昕, 讲师, 硕士, 主要从事非线性控制与理论计算研究.

G 蛋白分解为 $G\beta\gamma$ 以及与 GTP 结合的 $G\alpha(\text{GTP-G}\alpha)$, 具有传导活性的 $G\beta\gamma$ 以及 $\text{GTP-G}\alpha$ 可实现信号传导, 进而实现细胞体内功能的转变^[3]. 具有传导活性的物质失活较慢, 而 RGS 蛋白可大大加速这个过程: RGS 蛋白与 $\text{GTP-G}\alpha$ 相结合, 加速 GTP 水解为 GDP, 进而使得 $G\alpha$ 与 $G\beta\gamma$ 相结合形成没有信号传导活性的与 GDP 结合的 G 蛋白^[3].

过快的失去信号传导活性会导致一系列的疾病, 包括中枢神经系统紊乱、心血管疾病以及癌症^[4]. 因此, 阻止 RGS 蛋白与 $G\alpha$ 相结合成为药物研发的一个重要方向. 药物可直接作用于 RGS 蛋白与 $G\alpha$ 的 PPI 界面(protein-protein interface), 但是该 PPI 界面过于平坦以至于并没有可供药物结合的位点^[5]. 药物也可以作用于远离 PPI 界面的 RGS 蛋白中的氨基酸, 通过变构调节实现阻止 RGS 蛋白与 $G\alpha$ 相结合. TDZD 抑制剂可与远离 PPI 界面的半胱氨酸残基通过共价键相结合, 通过变构调节来抑制 RGS 蛋白与 $G\alpha$ 相结合^[6].

最近的研究表明 TDZD 抑制剂对 RGS 蛋白具有一定的选择性, 而这种选择性与 RGS 蛋白中的半胱氨酸残基的数目有关^[7]. 然而, 研究也表明 TDZD 抑制剂对具有相同数目半胱氨酸残基的 RGS 蛋白也有不同的抑制作用. 尽管 TDZD 抑制剂具有选择性, 但选择性较低, 进而可能导致药物产生副作用^[7]. 因此, 需要进一步了解 RGS 蛋白的氨基酸残基动力学, 深入理解 RGS 蛋白中的药物低选择性以及药理相似性. 本研究使用分子动力学模拟方法, 集中深入分析了 3 种不同 RGS 蛋白(RGS4, RGS8, RGS17)的结构动力学, 在一定程度上解释了药物选择性以及药理相似性的原因.

1 模拟及分析方法

1.1 模型以及模拟方法

在本研究中, 我们使用了 VMD (Visual Molecular Dynamics)^[8] 来建立模型以及分析数据等, NAMD (Nanoscale Molecular Dynamics)^[9] 来进行分子动力学模拟并使用了 CHARMM 力场^[10]. RGS 蛋白的初始坐标来源于 PDB (protein data bank): RGS4 (PDB code: 1EZT)、RGS8 (PDB code: 2IHD)以及 RGS17 (PDB code: 1ZV4). 我们使用 TIP3P 水分子对体系溶剂化以及 NaCl 进行体系的电荷中和. 最终每个体系的原子数目约为 30 000.

模拟主要分为 3 个阶段: 首先是 500 步的体系能量最小化; 其次为使用 NPT 系综进行 100 ps 的体系体积优化; 最后使用 NVT 系综进行 $2\mu\text{s}$ 的分子动力学模拟. 在模拟中, 时间步长为 2 fs, 并使用了周期性条件. 通过使用朗之万恒温器, 温度控制在 310 K; 使用 Nosé-Hoover 恒压器, 压力控制在 1 个标准大气压.

1.2 分析方法

为了考察氨基酸残基的柔性, 我们计算了均方根涨落(RMSF: root mean squared fluctuation). 较大的均方根涨落表明该氨基酸残基具有较大的柔性, 例如处于蛋白质端基的氨基酸, 一般来说具有较高的亲水性. 当药物作用于多肽或蛋白质时, 也会使得某些氨基酸残基的柔性增大, 这表明该药物对这些氨基酸残基的影响较大. 柔性增加的程度越大, 表明了该药物对作用的蛋白质影响越大. 我们使用 Carma 进行了氨基酸残基对的动态互相关(DCC: Dynamic cross correlation)分析^[11]. 氨基酸残基对 IJ 的互相关值(C_{IJ})定义如下:

$$C_{IJ} = \frac{\langle (\Delta \vec{r}_I(t) \cdot \Delta \vec{r}_J(t)) \rangle}{(\langle \Delta \vec{r}_I(t)^2 \rangle \langle \Delta \vec{r}_J(t)^2 \rangle)^{1/2}}$$

$\Delta \vec{r}_I(t) = \vec{r}_I(t) - \langle \vec{r}_I(t) \rangle$, $\vec{r}_I(t)$ 是氨基酸残基 I 的 α 碳坐标, $\langle \vec{r}_I(t) \rangle$ 为氨基酸残基 I 的 α 碳平均坐标. 当 2 个残基对有 75% 以上的模拟时间处在 $4.5\text{\AA}$ 距离以内时, 会形成一个 edge (用一条线段表示). DCC 分析是变构调节常用的分析方法. 当氨基酸残基对有较高的 DCC 值时, 表明该氨基酸残基对具有较强的(远程)相关作用.

我们进一步使用了 VMD 中的 NetworkView 插件进行了社区网络分析, 在 DCC 分析的基础上将 RGS 蛋白划分成了多个社区^[12]. 划分方法采用的是 Girvan-Newman 算法^[13]. 最优的社区划分在最大的模块化值取得. 一般而言, 模块化值在 0.4~0.7 之间, 本文的模块化值约为 0.6.

2 结果与讨论

2.1 结构对比

RGS 蛋白家族结构上非常相似, 具有一个同源的 RGS 盒子. 该盒子通常含有约 120 个氨基酸残基, 包含了 9 个 α 螺旋结构($\alpha 1 \sim \alpha 9$), 如图 1 所示. RGS4 与 RGS8 在氨基酸序列上有 56.3% 的相似性, 而 RGS4, RGS8 与 RGS17 的相似性分别为 39.1% 以及 38.8% (图 1b). RGS4 与 RGS8 高度相似是因为它们同属于 R4 亚家族, 而 RGS17 属于 RZ 亚家族^[1]. RGS4 有 4 个半胱氨酸残基可供 TDZD 抑制剂作用, 而 RGS8 具有 2 个半胱氨酸残基以及 RGS17 只有 1 个半胱氨酸残基(图 1a). 3 个 RGS 蛋白具有一个相同位置的半胱氨酸残基, 位于 $\alpha 4$ 螺旋上. RGS4 和 RGS8 还具有另一个相同位置的半胱氨酸残基, 位于 $\alpha 6 \sim \alpha 7$ 之间的无规卷曲(图 1a). RGS4 的另外 2 个半胱氨酸残基分别位于 $\alpha 3$ 螺旋和 $\alpha 6$ 螺旋上.

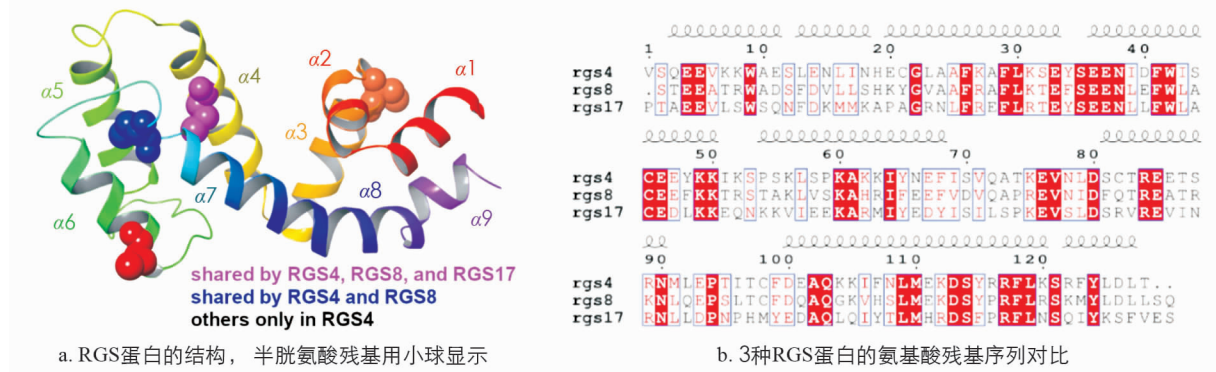


图 1 RGS 蛋白的结构及氨基酸残基序列

2.2 均方根涨落 RMSF

图 2 描述了 3 种 RGS 蛋白的氨基酸残基 RMSF 值. 一个大的 RMSF 数值表明该残基具有强的柔性. 图 2 的结果表明端基上的氨基酸残基具有较高的 RMSF 值, 这是由于与端基上的氨基酸残基相互作用的氨基酸较少以及所受到的其他氨基酸的牵制较弱. 同样, 我们也可以发现无规卷曲上的氨基酸残基也会有相对较高的 RMSF 值.

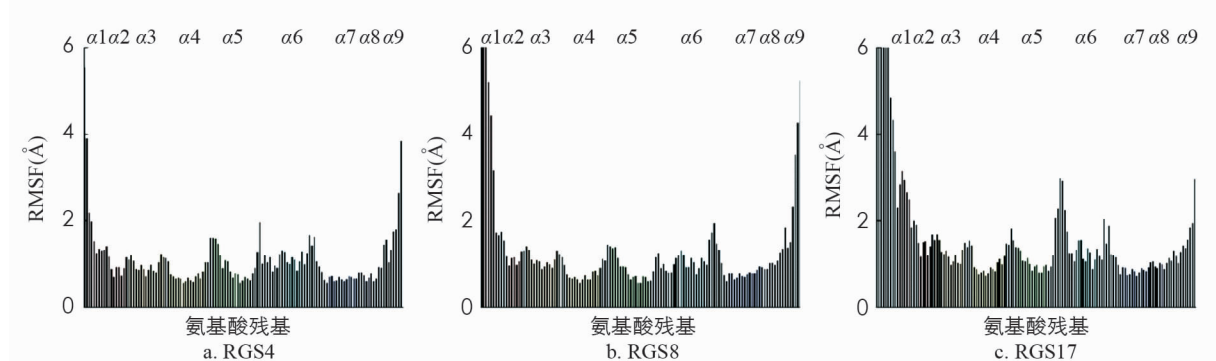


图 2 3 种 RGS 蛋白的均方根涨落 RMSF

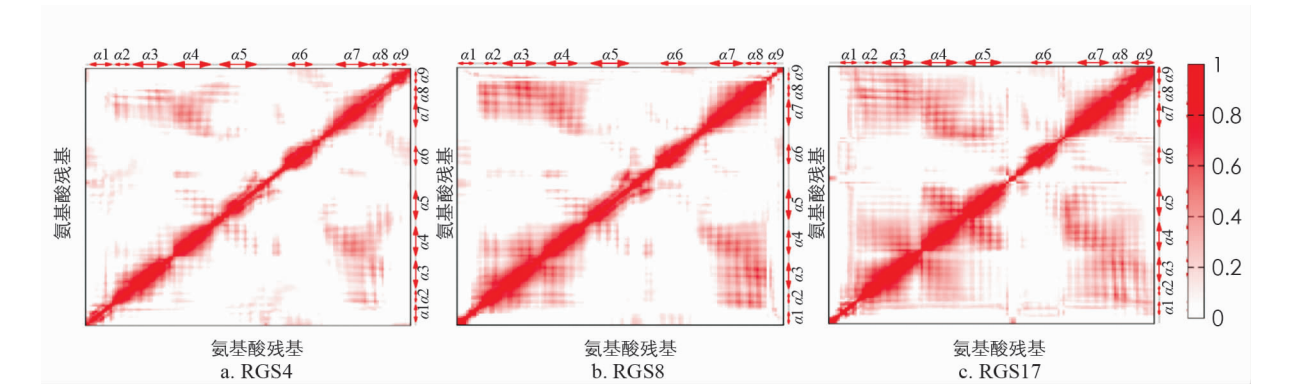
RGS 蛋白中与 $G\alpha$ 链直接作用的氨基酸残基主要位于 $\alpha 3 \sim \alpha 4$ 之间以及 $\alpha 5 \sim \alpha 6$ 之间的无规卷曲. 因此, 该部分氨基酸残基若具有较大的 RMSF 值会导致与 $G\alpha$ 链较弱地结合, 进而使得 RGS 蛋白与 $G\alpha$ 链的相互作用下降, 从而表明 TDZD 抑制剂具有较高的抑制作用. 在图 2 中, 我们可以发现 RGS17 的 $\alpha 5 \sim \alpha 6$ 之间的无规卷曲具有明显高于 RGS4 以及 RGS8 的 RMSF 值. 如果 TDZD 抑制剂可以变构调节该无规卷曲, 我们可以预测在 TDZD 抑制剂只结合 3 种 RGS 蛋白共同的半胱氨酸残基时, TDZD 抑制剂会对 RGS17 具有相对较高的抑制作用. 而图 2 表明 3 种 RGS 蛋白在 $\alpha 3 \sim \alpha 4$ 之间的无规卷曲的 RMSF 值差异较小.

同时我们也注意到有 3 个半胱氨酸残基位于无规卷曲(图 1a 中的蓝色小球)或者靠近无规卷曲(图 1a 中的橙色和红色小球). 若该无规卷曲具有较大的 RMSF 值, 会使得半胱氨酸残基容易暴露在外, 从而易

于与 TDZD 结合, 进一步易于抑制 RGS 蛋白与 $G\alpha$ 链的结合. 这种机制也会导致 TDZD 抑制剂对不同 RGS 蛋白产生不同的作用机理.

2.3 氨基酸残基对动态互相关 DCC

当氨基酸残基对在所有的 MD 运动轨迹中, 运动方向一致时, 会给出一个较大的正互相关 DCC 值, 否则给出一个较小的负互相关 DCC 值. DCC 值越大表明该氨基酸残基对相互作用越强, 变构调节时二者的相互影响越强. 图 3 描述了氨基酸残基对之间的动态正互相关 DCC 值. 在 RGS 蛋白中, 我们观察到 α 螺旋之间会具有较大的 DCC 值, 表明该 α 螺旋对具有较强的(远程)相互影响作用. 当 2 个靠近的 α 螺旋具有较大的 DCC 值时, 处于蛋白质内部的残基会较难暴露, 进而难以成为药物的作用靶点. 从图 3 我们可以看出, 3 种 RGS 蛋白的 $\alpha 4$ 螺旋与 $\alpha 7$ 螺旋之间的氨基酸残基对具有较大的 DCC 值. 这表明位于 $\alpha 4$ 螺旋上的半胱氨酸残基(图 1a 中粉色小球)暴露在溶液中的窗口较小, TDZD 抑制剂较难作用于该结合位点. 同时我们也发现 $\alpha 4$ 螺旋与 $\alpha 7$ 螺旋之间的氨基酸残基对 DCC 值有如下关系: $RGS17 > RGS8 > RGS4$. 这表明 TDZD 抑制剂相对更容易作用于 RGS4 而更难作用于 RGS17.



图中标注了同一 RGS 蛋白不同的 α 螺旋. DCC 值从 0~1 表明了氨基酸残基对的动态互相关从弱到强.

图 3 氨基酸残基对动态正互相关 DCC 图

RGS4 和 RGS8 还共同拥有一个位于 $\alpha 6 \sim \alpha 7$ 之间的无规卷曲上的半胱氨酸残基(图 1a 蓝色小球). 该半胱氨酸残基的暴露涉及 $\alpha 4$ 螺旋与 $\alpha 7$ 螺旋之间以及 $\alpha 4 \sim \alpha 5$ 无规卷曲与 $\alpha 6 \sim \alpha 7$ 无规卷曲之间的氨基酸残基对 DCC 值. 从图 3a 和图 3b 中, 我们发现这些 DCC 值都较小(淡红色)或呈负相关状态(白色). 这表明该半胱氨酸残基容易暴露成为 TDZD 抑制剂的作用位点.

RGS4 中位于 $\alpha 3$ 螺旋始端的半胱氨酸残基也具有较大暴露窗口成为抑制剂的作用位点, 这是由于该处与相邻的氨基酸残基对 DCC 值较小或为负相关(图 3a). RGS4 中位于 $\alpha 6$ 螺旋始端的半胱氨酸残基也具有较大暴露窗口. 但该半胱氨酸残基位于 RGS4 与 $G\alpha$ 的 PPI 界面处, 因此该残基较难成为 TDZD 抑制剂的作用位点.

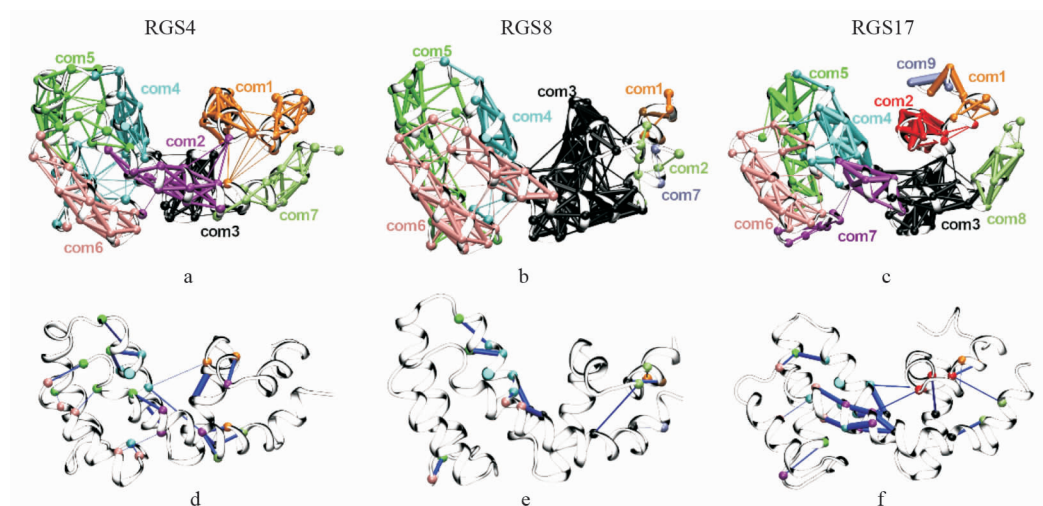
2.4 社区网络分析

使用 Girvan-Newman 算法^[13]可以将蛋白质划分为多个社区(图 4a,b 和 c). 一个社区内的氨基酸残基联系紧密且一般运动方向一致(图 4 中的粗线段), 社区之间的氨基酸残基则关联较弱(图 4 中的细线段). 图 4d,e 和 f 描述了各个社区之间的联系.

从图 4 中我们发现 RGS4 和 RGS8 被划分为 7 个社区, 而 RGS17 被划分为 9 个社区. RGS4 中 $\alpha 1$ 螺旋与 $\alpha 2$ 螺旋属于同一个社区(图 4a, com1), 而 RGS8 和 RGS17 的 $\alpha 1$ 螺旋则被划分为多个社区. RGS17 的 $\alpha 2$ 螺旋独立为一个社区(图 4c, com2). RGS4 中 $\alpha 3$ 螺旋成为一个单独社区(图 4a, com3), 而 RGS8 中 $\alpha 2$ 螺旋、 $\alpha 3$ 螺旋和 $\alpha 8$ 螺旋成立一个社区(图 4b, com3), RGS17 中 $\alpha 3$ 螺旋和 $\alpha 8$ 螺旋成立一个社区(图 4c, com3). 这表明 RGS8 和 RGS17 中的 $\alpha 3$ 螺旋和 $\alpha 8$ 螺旋联系较为紧密, 而 RGS4 中的 $\alpha 3$ 螺旋和 $\alpha 8$ 螺旋联系较少. 这个现象和 DCC 图中一致(图 3): RGS8 和 RGS17 的 $\alpha 3$ 螺旋和 $\alpha 8$ 螺旋 DCC 值较大, 而 RGS4 的 $\alpha 3$ 螺旋和 $\alpha 8$ 螺旋 DCC 值较小. RGS4 和 RGS8 中 $\alpha 4$ 螺旋没有和 $\alpha 7$ 螺旋中氨基酸残基形成同一个社区(图 4, com4), 而 RGS17 中 $\alpha 4$ 螺旋和 $\alpha 7$ 螺旋中部分氨基酸残基形成了同一个社区. 这个现象也和 DCC 图中

一致. RGS4 和 RGS8 中 $\alpha 4$ 螺旋和 $\alpha 5$ 螺旋中部分氨基酸残基形成同一个社区. RGS8 和 RGS17 中的 $\alpha 5$ 螺旋独立成为了一个社区, 而 RGS4 中的 $\alpha 5$ 螺旋部分氨基酸残基与 $\alpha 6 \sim \alpha 7$ 无规卷曲形成了一个社区(图 4, com5). RGS4 和 RGS17 中的 $\alpha 6$ 螺旋独立成为一个社区, 而 RGS8 中 $\alpha 6$ 螺旋和 $\alpha 7$ 螺旋成为一个社区(图 4, com6). RGS4 和 RGS17 中的 $\alpha 7$ 螺旋部分氨基酸残基成为一个社区(图 4a, com2; 4c, com7). RGS4 和 RGS17 中的 $\alpha 9$ 螺旋独立成为一个社区, 而 RGS8 中 $\alpha 9$ 螺旋部分残基会和 $\alpha 1$ 螺旋部分残基成为一个社区(图 4a, com7; 4b, com2; 4c, com8).

从上面分析可以看出, 不同 RGS 蛋白的社区网络存在着相似性, 如 RGS4 和 RGS8 中(图 4a, 4b, com4)、RGS4 和 RGS17 中的(图 4a, 4c, com6)和(图 4a, com7; 4c, com8)以及 RGS8 和 RGS17 中的(图 4b, 4c, com3)和(图 4b, 4c, com5). 同样不同 RGS 蛋白的社区网络存在着不同之处, 如 RGS4 的 com1 和 com3、RGS8 的 com6 以及 RGS17 的 com4.



RGS4 (a 和 d); RGS8 (b 和 e); RGS17 (c 和 f). 图中相同颜色的小球表明这些氨基酸残基属于同一个社区, 社区的标记为 com#.

图 4 3 种 RGS 蛋白的社区网络图以及各个社区之间的联系图

3 结论与展望

药物广谱作用于一个蛋白质家族是药物研发的一个重要方面, 提高药物的选择性可以减少副作用的发生^[14]. 在本研究中, 我们使用了 3 种 RGS 蛋白为模型, 从氨基酸序列、均方根涨落 RMSF、氨基酸残基对互相关 DCC 以及社区网络分析 4 个方面给出了 3 种 RGS 蛋白的类似性和不同之处, 这对于设计新的具有高选择性的药物提供了一定的理论基础, 具有一定的指导意义. 本文主要集中讨论了半胱氨酸作为常用的作用位点在不同的 RGS 蛋白中表现出的相同与不同之处. 研究新的抑制作用位点可能会提高药物的选择性. 同时, 本文中的研究方法还可适用于其他的蛋白质家族, 如蛋白激酶.

由于有限的计算资源, 虽然总的计算时长达到了 $6 \mu\text{s}$, 但本文中的蛋白仅限于 RGS 蛋白, 未能考虑 RGS 蛋白和 $G\alpha$ 的相互作用, 这是本研究的局限之处. 在未来有足够计算资源的情况下, 应对更大的体系进行更长时间的分子动力学模拟^[15].

参考文献:

- [1] NEUBIG R R, SIDEROVSKI D P. Regulators of G-Protein Signalling as New Central Nervous System Drug Targets [J]. Nature Reviews Drug Discovery, 2002, 1(3): 187-197.
- [2] STOCK A M, ROBINSON V L, GOUDREAU P N. Two-Component Signal Transduction [J]. Annual Review of Biochemistry, 2000, 69(1): 183-215.
- [3] MCCUDDEN C R, HAINS M D, KIMPLE R J, et al. G-Protein Signaling: Back to the Future [J]. Cellular and Molecular Life Sciences, 2005, 62(5): 551-577.
- [4] CHEN J G, WILLARD F S, HUANG J, et al. A Seven-Transmembrane RGS Protein that Modulates Plant Cell Prolif-

- eration [J]. Science, 2003, 301(5640): 1728-1731.
- [5] BOSCH D E, ZIELINSKI T, LOWERY R G, et al. Evaluating Modulators of “Regulator of G-Protein Signaling” (RGS) Proteins [J]. Current Protocols in Pharmacology, 2012, 56(1): 2-8.
- [6] XIE Z H, CHAN E C, DRUEY K M. R4 Regulator of G Protein Signaling (RGS) Proteins in Inflammation and Immunity [J]. The AAPS Journal, 2016, 18(2): 294-304.
- [7] NIESSEN K A, XU M Y, DENG Y T, et al. Importance of Protein Vibration Directionality on Function [J]. Biophysical Journal, 2017, 112(3): 353a.
- [8] HUMPHREY W, DALKE A, SCHULTEN K. VMD: Visual Molecular Dynamics [J]. Journal of Molecular Graphics, 1996, 14(1): 33-38.
- [9] NELSON M T, HUMPHREY W, GURSOY A, et al. NAMD: a Parallel, Object-Oriented Molecular Dynamics Program [J]. The International Journal of Supercomputer Applications and High Performance Computing, 1996, 10(4): 251-268.
- [10] HUANG J, RAUSCHER S, NAWROCKI G, et al. CHARMM36m: an Improved Force Field for Folded and Intrinsically Disordered Proteins [J]. Nature Methods, 2017, 14(1): 71-73.
- [11] GLYKOS N M. Software News and Updates Carma: a Molecular Dynamics Analysis Program [J]. Journal of Computational Chemistry, 2006, 27(14): 1765-1768.
- [12] SETHI A, EARGLE J, BLACK A A, et al. Dynamical Networks in tRNA: protein Complexes [J]. Proceedings of the National Academy of Sciences, 2009, 106(16): 6620-6625.
- [13] GIRVAN M, NEWMAN M E. Community Structure in Social and Biological Networks [J]. Proceedings of the National Academy of Sciences, 2002, 99(12): 7821-7826.
- [14] 郑登峰, 文 静, 喻会平, 等. 烟蚜茧蜂对桃蚜和萝卜蚜的寄生选择性 [J]. 西南大学学报(自然科学版), 2016, 38(1): 60-66.
- [15] 简文杰, 甘纯玑, 庞 杰. 水溶性红曲红素的分子动力学模拟研究 [J]. 西南大学学报(自然科学版), 2012, 34(11): 112-116.

责任编辑 崔玉洁