

doi: 10.3969/j.issn.1674-1242.2022.04.005

MEK1 及其突变体的曲美替尼结合轨迹的马尔可夫模型分析

刘明皓, 李明昊, 陈雪, 韩葳葳

(吉林大学分子酶学工程教育部重点实验室, 吉林大学生命科学学院, 吉林长春 130012)

【摘要】 MEK1 是 Ras-Raf-MEK-ERK 通路中重要的信号分子, 在肿瘤发生中发挥着重要的作用。通路中的 Raf、MEK1 等基因发生突变导致通路被持续活化被认为是导致肿瘤无限增殖的重要原因之一。曲美替尼能有效地抑制 MEK1 的活性, MEK1 结构中 α 螺旋的消失会降低曲美替尼的治疗效果。该研究通过 PCCA+ 聚类算法、马尔可夫模型和通量分析等方法, 对 E203K 突变体和曲美替尼的作用轨迹进行分析, 研究发现 E203K、P124S 和野生型 MEK1 在与曲美替尼作用时, C166-S181 部分靠近曲美替尼的一端会发生解旋变为氢键转角, 而远端螺旋则不会发生解旋。

【关键词】 丝裂原活化蛋白激酶激酶 1; 曲美替尼; E203K 突变体; 马尔可夫模型; 通量分析

【中图分类号】 Q6

【文献标志码】 A

文章编号: 1674-1242(2022)04-0223-07

Markov Model Analysis of Trametinib Binding Trajectory of MEK1 and Its Mutants

LIU Minghao, LI Minghao, CHEN Xue, HAN Weiwei

(Key Laboratory for Molecular Enzymology and Engineering of Ministry of Education, School of Life Science, Jilin University, Changchun, Jilin 130012, China)

【Abstract】 MEK1 is an important signal molecule in the Ras-Raf-MEK-ERK pathway. It plays an important role in tumorigenesis. Mutation of Raf and MEK1 genes of the pathway, resulting in the pathway being continuously activated is believed to be one of the important causes of unlimited tumor proliferation. Trametinib can effectively inhibit the activity of MEK1, and the loss of alpha helix in MEK1 structure will reduce the therapeutic effect of trametinib. In this experiment, the trajectories of action of E203K mutant and trametinib were analyzed by PCCA+ clustering algorithm, Markov model and flux analysis, and it was found that E203K, P124S and wild type all undergo deconvolution into hydrogen bonding corner at the end of C166-S181 part near trametinib when interacting with trametinib, while there are differences in structural changes at the distal end.

【Key words】 MAP Kinase Kinase 1; Trametinib; E203K Mutant; Markov Model; Flux Analysis

1 引言

MEK1 (MAP kinase kinase 1), 即丝裂原活化蛋白激酶激酶 1, 是 Ras-Raf-MEK-ERK 通路中

重要的信号分子。通过抑制 MEK1 的激酶活性, 使 Ras-Raf-MEK-ERK 通路中下游的 ERK 磷酸化现象减少, 从而能够调节细胞的增殖等活动。在 ERK 信号通路的研究中, 常常以 MEK1 作为肿瘤

收稿日期: 2022-11-21

基金项目: 吉林省科技发展规划项目 (3D5204381465)

作者简介: 刘明皓 (2001—), 女, 辽宁省鞍山市人, 本科生, 从事生物物理学研究

通信作者: 韩葳葳, 女, 教授, 博士生导师, 电话 (Tel.): 13756550860; E-mail: weiwei@jlu.edu.cn

治疗的靶标。目前常用的抑制剂曲美替尼可以对 MEK1 的活性产生有效的抑制。然而, MEK1 不同突变体之间有着活性差异, 和曲美替尼的作用效果差于野生型 MEK1。本研究通过分子动力学模拟来研究曲美替尼与突变型 MEK1 之间的相互作用过程。为了弥补分子动力学模拟无法识别模拟轨迹中的具体状态, 甚至无法确定不同状态之间的动力学关系的缺陷, 本研究将创新性地采用马尔可夫模型分析微状态之间的动力学关系。本研究将结合 TICA 降维法、k-means 聚类算法、PCCA+ 聚类算法和通量分析等, 对比 E203K 突变体、P124S 突变体、野生型 MEK1 与曲美替尼作用过程的区别, 解释其作用效果差异产生的分子机制。本研究所采用的方法也适用于探究其他药靶作用中突变型效果差异机理的研究。

2 MEK1 蛋白简介

2.1 MEK1 蛋白与癌症的关系

MEK1 是 Ras-Raf-MEK-ERK 通路中重要的信号分子, 该通路和癌症肿瘤的形成息息相关^[1]。当癌症发生时, 原癌基因产物 Ras 发生活化, 引发相关器官处的薄膜组织局部化, 引发 Raf 的二聚作用^[2]。Raf 蛋白激酶在二聚作用活化后, 激活通路下游的 MEK1 蛋白^[3]。MEK 中的一个亚型 MEK1, 可以激活通路下游底物 ERK 中的苏氨酸及酪氨酸, 使 ERK 发生活化, 从而进入细胞核, 对 Ets、Jun、RSK、Fos、Elk-1 等转录因子进行磷酸化修饰, 进而改变基因的表达, 参与到细胞增殖、分化、代谢、凋亡等生物过程中。研究证明, 在肿瘤发生过程中, 该通路的 Raf、MEK1、ERK 等基因会发生突变, 导致通路被持续活化, 这一现象被认为是导致肿瘤无限增殖的重要原因之一。因此, 在 ERK 信号通路的研究中, 常以这 4 种蛋白激酶为靶标进行药物研究^[4]。

虽然目前尚无 ERK1/2 抑制剂被正式批准上市, 但是已经有一些小分子 ERK 抑制剂处于临床或临床前研究阶段。有研究发现, 选择性敲除 ERK1/2 能够有效杀死 A375 黑色素瘤细胞, 同时增强 A375 细胞对 BRAF 抑制剂维罗替尼的敏感性。此外, MEK 抑制剂耐药的细胞持续保持对 MAPK 信号通路的依赖, 其选择性地对 ERK1/2 抑制剂敏感, 应用 ERK 抑制剂能够有效阻断细胞

增殖。

MEK1 蛋白激酶中包含着两个空间位置接近的结合位点, 分别是别构结合位点和 ATP 结合位点。抑制剂对别构结合位点的非竞争可逆抑制作用对 MEK1 激酶的活性进行抑制。ATP 结合位点被抑制后, 导致 MEK1 无法获得足够的 ATP 而效果降低^[5]。

2.2 MEK1 不同突变体与曲美替尼作用时的区别

曲美替尼是一种 MEK1 别构抑制剂, 能有效地抑制 MEK1 的活性, 使通路中下游的 ERK 磷酸化现象减少, 从而能够调节细胞的增殖、扩散、凋亡。然而, MEK1 的不同突变体与曲美替尼的作用过程也会有差异。MEK1 的常见突变形式包括 K57N 型、E203K 型和 A52V 型。在曲美替尼与 K57N 型、E203K 型突变体结合的过程中, 两种突变体 MEK1 蛋白的 C166-S181 部分残基对应的 α 螺旋会出现解旋的现象, 而 A52V 型突变体和野生型 MEK1 对应的 α 螺旋不会完全解旋。因此, 对此处 α 螺旋解旋进行研究可以为新型 MEK1 抑制剂药物的设计提供一些帮助^[6]。

曲美替尼与突变型 MEK1 相互作用的过程可以通过分子动力学模拟来研究。然而, 模拟轨迹中的状态无法确定, E203K 突变体和曲美替尼相互作用的过程中, α 螺旋解旋过程的具体构象变化无法被识别。因此, 本研究通过 PCCA+ 聚类算法、马尔可夫模型和通量分析等方法, 对 E203K 突变体、P124S 突变体和曲美替尼的作用轨迹进行分析, 同时将作用轨迹进行对比, 期望研究出 E203K 突变体中 α 螺旋解旋的具体原因。

3 马尔可夫模型构建过程所需算法

3.1 实验流程总述

本研究首先获取曲美替尼和两种类型的 E203K 的作用轨迹, 将轨迹信息用数组数据表示。使用 TICA 降维方法对数组进行降维处理, 之后进行 k-means 聚类。获得准确的聚类结果后, 确定 lag time 的数值, 在确定 lag time 后即可结合聚类结果构建马尔可夫模型。再将分子轨迹进行宏观描述, 即 PCCA 聚类, 上述获得的状态结果和转移概率矩阵便可用于通量分析。实验流程图见图 1。

3.2 TICA 降维方法

TICA (Time-lagged Independent Component Analysis, 时滞独立成分分析), 是马尔可夫模型

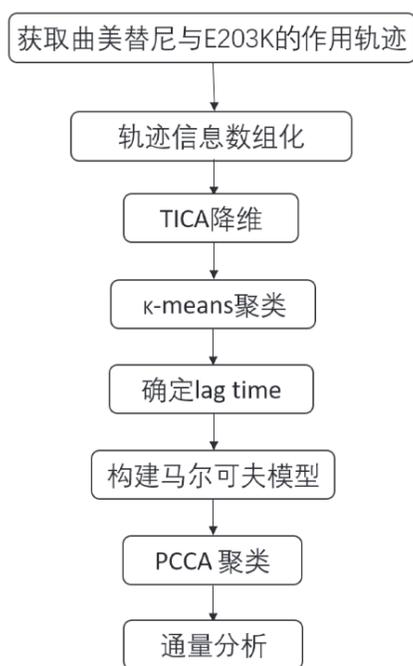


图1 实验流程图

Fig.1 Flowchart of the experiment

构建中常用的降维方法。TICA 算法是 ICA 算法中的一种，它可以使变换坐标的自相关最大化，特别适用于分子动力学输入。TICA 可以在给定的滞后时间上找到最大自相关的坐标^[7]，并有效提取分子动力学数据中的慢序参数，因此是在 k-means 聚类之前对分子模拟数据进行处理的最佳选择^[8]。

3.3 k-means 聚类算法

在马尔可夫模型中，使用 k-means 聚类算法的主要原因在于将分子模拟轨迹中的构象状态进行聚类，将聚得的每一类都定义为马尔可夫模型中的一个微状态，即得到微状态离散轨迹。k-means 聚类算法是机器学习中一种常用的算法，是非监督学习方法中的一种。本研究在进行 k-means 聚类步骤时，对每一个 k 值都重复进行了 10 次聚类来计算得分。另外，还利用肘部法则简化了计算步骤，仅对较合理的多个 k 值进行网格搜索，通过对这一 k 值聚类结果进行 VAMP 打分，获得最适 k 值^[9]。

3.4 lag time 的确定

lag time 是指在离散轨迹之间进行跳转时每次跳转所使用的时间。lag time 是马尔可夫模型通过离散轨迹计算转移概率矩阵中极其重要的一个参数，它直接决定了转移概率矩阵的数值。为了检验 lag time 的选择是否合适，常常使用 CK 检测

(Chapman-Kolmogorov Test) 方法进行检验^[10]。

3.5 PCCA + 聚类算法

PCCA+ 聚类算法专门用于马尔可夫模型中微状态的处理。使用 PCCA+ 聚类算法，将轨迹构象整体划分为多个宏观状态，方便后续的通量分析^[11]。

3.6 通量分析

获得转移概率矩阵后，便可以利用这一结果来获取轨迹的热力学和动力学信息。由于这一模型具有高复杂度，因此要使用更为粗粒化的模型且以更紧凑的方式提供相同的定量信息^[12]。转移概率的存在导致中间体序列并不是唯一的，每一种序列情况都会依据其路径上的通量确定一个概率，在计算得到各个路径的概率后，便可利用每条路径的构象变化和该路径出现的概率进行通量分析了。

4 用马尔可夫模型研究 MEK1 与曲美替尼的相互作用

4.1 三个体系的 TICA 降维

本研究的全部计算均用 PyEMMA 软件包进行。首先，计算曲美替尼与 MEK1 及其两种突变体的相互作用轨迹，数据来自 NCBI。平行运行 10 份，每一份的时长均为 500 ns，步长均为 0.1 ns。利用 featurizer 函数和 load 函数，将轨迹及对接结构的拓扑信息导入 PyEMMA 软件包。导入数据时，选取 MEK1 及其两种突变体的 CYS166-SER181 残基作为研究对象。导入数据时选择的轨迹信息为此处残基骨架的 phi 角、psi 角信息，对 10 条分子轨迹的每一帧都进行量化，将抽象的分子轨迹用具象化的数组数据来表示。经过上述操作，10 条分子轨迹被转化成 10 个维度为 5 000×32 的二维数组，5 000 表示每一条轨迹有 5 000 帧，32 表示每一帧用于量化残基骨架的 phi 角、psi 角信息为 32 维。

本研究使用 TICA 降维方法对数组进行降维处理，保留 95% 的方差信息，使用的是 PyEMMA 软件包中的 coordinates.tica 方法，之后对降维结果进行 k-means 聚类。k-means 聚类使用 PyEMMA 软件包中自带的 coordinates.cluster_kmeans 方法进行计算，聚类过程中， k 值选择 75。野生型 MEK1 的轨迹前后变化较小，仅有一小部分与整体分离。其原因可能是野生型 MEK1 与曲美替尼作用过程中 α 螺旋基本不发生变化（图 2）。

4.2 三个体系 lag time 的确定

采用 `msm.its` 方法进行系统弛豫时间的计算，之后利用 `plots.plot_implied_timescales` 方法将系统弛豫时间标度与 lag time 进行绘图。对于 E203K 突变体和野生型 MEK1，整个系统的弛豫时间标度

在 lag time 为 1 ns 时趋于收敛，因此，本研究选择 lag time 为 1 ns。因计算分子模拟轨迹时所采用的步长为 0.1 ns，所以定义，lag time 为 10 steps。而 P124S 突变体系统弛豫时间尺度在 3 ns，即 30 steps 处发生收敛，因此定义 lag time 为 30 steps（图 3）。

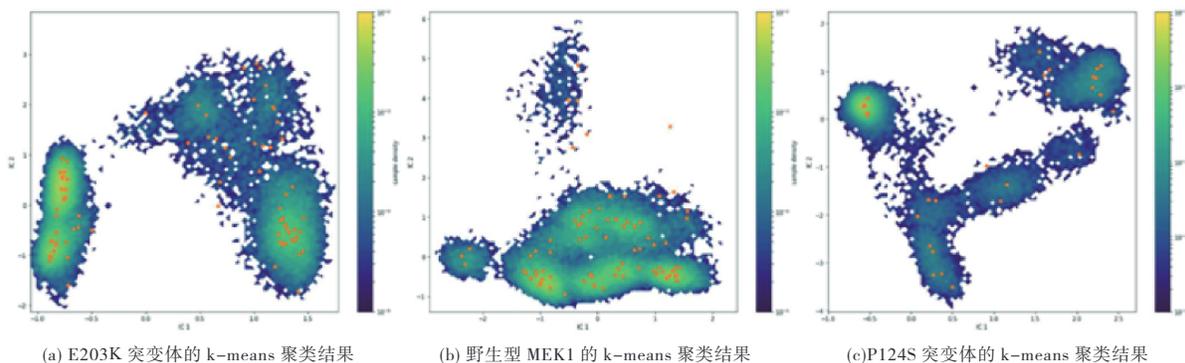


图 2 三个体系的 k-means 聚类结果

Fig.2 k-means clustering results of the three systems

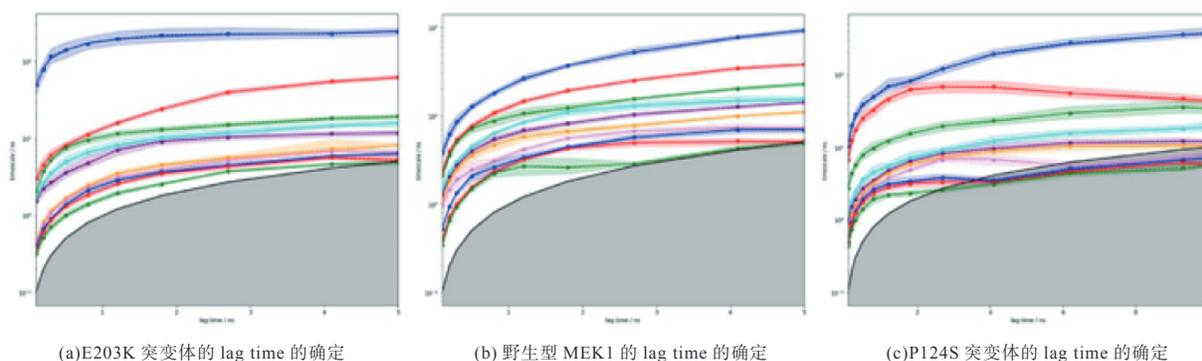


图 3 三个体系 lag time 的确定

Fig.3 Determination of the lag time of three systems

4.3 马尔可夫模型构建和结果分析

4.3.1 马尔可夫模型构建和 CK 检测

利用得到的聚类结果和 lag time 进行马尔可夫模型构建并计算转移概率矩阵。进一步通过 CK 检测对所构建模型的马尔可夫性进行检测。CK 检测采用 PyEMMA 软件包中的 `msm.cktest` 方法，参数中的 `mlags`，即 CK 检测中的倍数 K ，选择为 4；`nstates`，即后期 PCCA 聚类的宏观状态数，选择为 5。CK 检测结果表明，所构建的马尔可夫模型具有良好的马尔可夫性。

4.3.2 曲美替尼与 E203K 突变体 MEK1 的相互作用

通过 PCCA 聚类，将分子轨迹系统用宏观状态进行描述。采用 `msm.pcca` 方法，使用 PCCA 聚

类来描述的优点之一便是可以对轨迹结果进行很好的可视化分析，聚类结果中的 SA 为模拟轨迹中构象的起始状态，SB 为模拟轨迹中构象最终收敛的状态。S1、S2、S3 是按时间顺序出现的中间状态，是由多个微状态聚类得到的持续周期更长的亚稳态。各个聚类宏观状态如图 4 所示。紫色标记区域为主要研究的 α 螺旋区域。

利用体系的状态结果和马尔可夫模型的转移概率矩阵可以进行通量分析。最有代表性的构象转变路径为 SA→S2→SB，占总通量的 45.15%（表 1）。为了对构象转变过程进行更为具体的分析，本研究利用 VMD 中的 DSSP 功能进一步对蛋白质的二级结构进行计算，以了解转换路径上各个

残基的二级结构改变。可知在最有代表性的 SA->S2->SB 构象转变中,在曲美替尼与 E203K 突变形式的 MEK1 相互作用的过程中, E203K 从原本呈螺旋状态的 174~175 号残基转变为无规卷曲的状态, 176~179 号残基螺旋程度降低, 由四转角的 α 螺旋转变为三转角的 310 螺旋, 180 号残基转变为无规卷曲。之后, 螺旋程度进一步降低, 175~181 号残基失去螺旋结构, 转变为氢键转角的二级结构, 其余部分仍为无规卷曲。

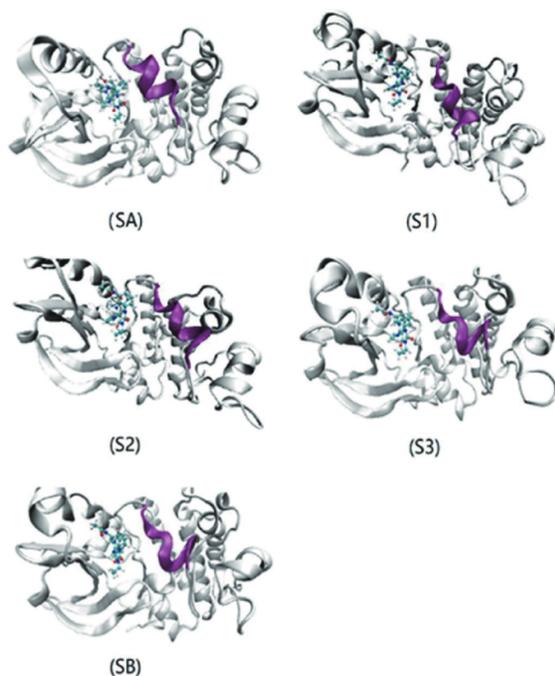


图4 E203K 突变体的 PCCA 聚类宏观状态图

Fig.4 Macro-state diagram of PCCA clustering of E203K mutant

表1 E203K 突变体通量分析中各条路径的通量

Tab.1 Fluxes of each pathway in the flux analysis of E203K mutant

通路	通路通量 (S^{-1})	占总通量的百分比 (%)
SA->S2->SB	4.00E-04	45.15
SA->S2->S3->SB	2.20E-04	24.83
SA->S1->S2->SB	2.18E-04	24.60
SA->S1->S2->S3->SB	3.80E-05	4.29
SA->SB	1.00E-05	1.13
	8.86E-04	100.00

SA->S2->S3->SB 与 SA->S1->S2->SB 这两种构象转变途径大致的二级结构转变与 SA->S2->SB 基本类似, 都是 α 螺旋逐渐解旋的过程, 并且螺旋整体从更为靠近曲美替尼的 N 端开始解旋, 逐渐向 C 端延伸, 最终整体由 α 螺旋变为 H 转角结构。

4.3.3 曲美替尼与野生型 MEK1 的相互作用

各个 PCCA 聚类宏观状态的特征分布极为接近。5 个宏观状态之间的构象差距较小, 并且紫色强调处都保持着 α 螺旋的结构 (图 5)。

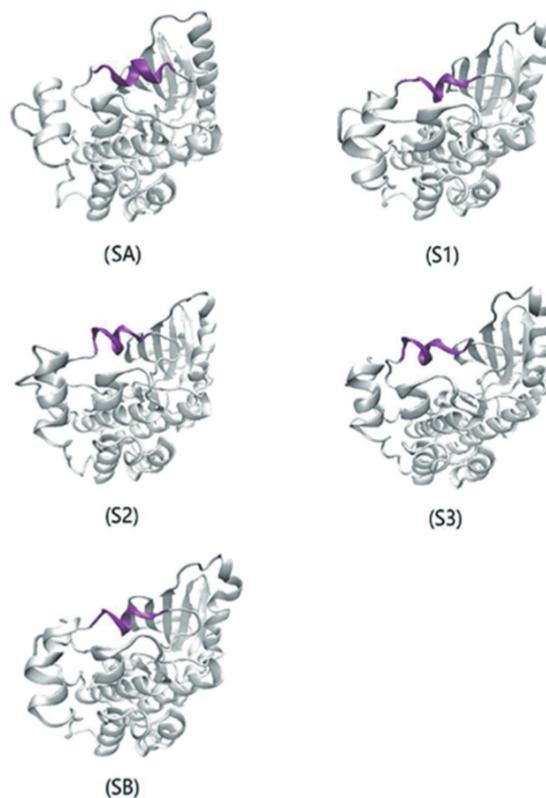


图5 野生型 MEK1 的 PCCA 聚类宏观状态图

Fig.5 Macro-state diagram of PCCA clustering of original type MEK1

通量分析结果显示, 最具有代表性的构象变化途径为 SA->S3->SB 及 SA->SB, 并且两者的通量占比相同 (表 2)。观察宏观状态的构象及二级结构的计算结果可知, 宏观状态之间的差距极小且 S1~SB 的二级结构基本没有变化, 都是 174~175 号残基为无规卷曲, 176~178 号残基为 310 螺旋, 179~181 号残基为 H 转角, 182 号残基为无规卷曲。这与初始状态 SA 有着一定差别, 但 S1~SB 都保持着螺旋的结构, 说明螺旋只有一部分发生了解旋, 未彻底解旋。发生解旋的部分是这一螺旋部分残基的 N 端, 而 C 端仍保持着螺旋的结构。

4.3.4 曲美替尼与 P124S 突变体 MEK1 的相互作用

经过研究, P124S 突变体的 MEK1 与 E203K 突变形式类似, 在与曲美替尼相互作用的过程中, 位于 C166-S181 部分的 α 螺旋会发生解旋^[10]。本研究对 P124S 突变体也构建了一个马尔可夫模

表2 野生型 MEK1 通量分析中各条路径的通量

Tab. 2 Fluxes of each pathway in the flux analysis of origin type MEK1

通路	通路通量 (S^{-1})	占总通量的百分比 (%)
SA→S3→SB	7.00E-03	27.34
SA→SB	7.00E-03	27.34
SA→S2→S3→SB	5.50E-03	21.48
SA→S3→S1→SB	3.22E-03	12.58
SA→S2→S3→S1→SB	1.72E-03	6.72
SA→S2→SB	8.60E-04	3.36
SA→S1→SB	3.00E-04	1.17
	2.56E-02	99.99

型进行分析,若结果与 E203K 突变形式接近,则可认为构建马尔可夫模型进行解释有一定意义。因轨迹整体结构变化较为简单,因此将 PCCA 聚类的状态数设置为 4 个(图 6),以便后续的通量

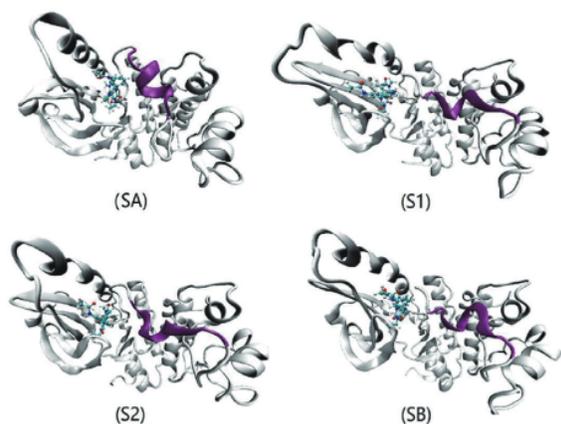


图6 P124S 突变体的 PCCA 聚类宏观状态图

Fig.6 Macro-state diagram of PCCA clustering of P124S mutant

分析。

最具有代表性的状态构象转变为 SA→S2→S1→SB, 占总通量的 58.59% (表 3)。剩余两个状态构象转变过程与这一通路相比均少其中一个中间微状态, 并且通量占比类似。故忽略这两条通路, 主要对第一条通路进行分析。使用 VMD 中的 DSSP 算法对各个状态的二级结构进行分析。

构象转变途径为 SA→S2→S1→SB, 根据所得的二级结构可知, 174~180 号残基由最初 SA 状态

表3 P124S 突变体通量分析中各条路径的通量

Tab. 3 Fluxes of each pathway in the flux analysis of P124S mutant

通路	通路通量 (S^{-1})	占总通量的百分比 (%)
SA→S2→S1→SB	3.00E-04	58.59
SA→S2→SB	1.07E-04	20.90
SA→S1→SB	1.05E-04	20.51
	5.12E-04	100.00

的螺旋结构变为 S2 状态的氢键转角和部分三转角螺旋, 之后这部分残基都变为 S1 状态的氢键转角, 最后到 SB 状态时, 二级结构未发生变化, 仍为氢键转角。蛋白的 181~182 号残基一直为无规卷曲, 基本未发生变化。

因此, 综合这一结果, 我们可知 P124S 突变体的 MEK1 在与曲美替尼的相互作用过程中, 174~180 号残基的 α 螺旋是逐步解旋的, 首先发生部分解旋变为三转角螺旋, 之后三转角螺旋也发生解旋, 整体变为氢键转角。

5 结论

曲美替尼与 MEK1 蛋白发生作用时, 若 MEK1 蛋白发生了 E203K、P124S 等形式的突变, 则 MEK1 蛋白 C166-S181 部分靠近曲美替尼的 α 螺旋会先发生解旋变为氢键转角的构象形式。随着与曲美替尼的作用不断增强, MEK1 蛋白 C166-S181 部分会完全发生解旋整体变为氢键转角的构象。若 MEK1 蛋白是野生型, 则在与曲美替尼相互作用的过程中, C166-S181 部分也会发生小程度的解旋, 但残基整体仍保持着螺旋的结构, 变为螺旋程度稍低的三转角螺旋。MEK1 结构中 α 螺旋的消失会导致这两种突变体的关键结构域发生结构紊乱, 进而降低曲美替尼对肿瘤的治疗效果。

参考文献

- [1] KOLCH W. Meaningful relationships: the regulation of the Ras/Raf/MEK/ERK pathway by Protein Interactions [J]. *Biochemical Journal*, 2000, 351(2): 289-305.
- [2] 郑宇静, 封宇飞. MEK1/2 抑制剂曲美替尼的药理作用与临床评价 [J]. *中国新药杂志*, 2014, 23(15): 1723-1725, 1733. ZHENG Yujing, FENG Yufei. Pharmacology and clinical evaluation of trametinib, a reversible inhibitor of MEK1/2 [J]. *Chinese Journal of New Drugs*, 2014, 23(15): 1723-1725, 1733.
- [3] HILGER R, SCHEULEN M, STRUMBERG D. The Ras-Raf-MEK-ERK pathway in the treatment of cancer [J]. *Oncology Research and Treatment*, 2002, 25(6): 511-518.
- [4] MA, QUIRION. The ERK/MAPK pathway, as a target for the treatment of neuropathic pain [J]. *Expert Opinion on Therapeutic Targets*, 2005, 9(4): 699-713.
- [5] OHREN J F, CHEN H, PAVLOVSKY A, et al. Structures of human MAP kinase kinase 1 (MEK1) and MEK2 describe novel noncompetitive kinase inhibition [J]. *Nature Structural & Molecular Biology*, 2004, 11(12): 1192-1197.
- [6] ZHU J, LI C, YANG H, et al. Computational Study on the Effect of Inactivating/Activating Mutations on the Inhibition of MEK1

- by Trametinib [J]. **International Journal of Molecular Sciences**, 2020, 21(6): 1-15.
- [7] BEYERLE E, GUENZA M. A Comparison of the Slow Dynamics in the Protein Ubiquitin Predicted by the LE4PD, PCA, and tICA from a Long Equilibrium Molecular Dynamics Simulation [J]. **Bulletin of the American Physical Society**, 2021: 1-28.
- [8] GUILLERMO P-H, FABIAN P, TONI G, *et al.* Identification of slow molecular order parameters for Markov model construction [J]. **The Journal of Chemical Physics**, 2013, 139(1): 1-28.
- [9] WU H, NOÉ F. Variational Approach for Learning Markov Processes from Time Series Data [J]. **Journal of Nonlinear Science**, 2020, 30(1): 1-27.
- [10] R B G. An overview and practical guide to building Markov state models [J]. **Advances in Experimental Medicine and Biology**, 2014, 797: 7-22.
- [11] NOÉ F, WU H, PRINZ J-H, *et al.* Projected and hidden Markov models for calculating kinetics and metastable states of complex molecules [J]. **The Journal of Chemical Physics**, 2013, 139(18): 11B609_1.
- [12] YILIN M, DIWAKAR S, S P V, *et al.* Transition path theory analysis of c-Src kinase activation [J]. **Proceedings of the National Academy of Sciences of the United States of America**, 2016, 113(33): 9193-9198.