



STAT6识别N4位点DNA的结构基础被揭示

王璞玥*, 杨正宗, 杜生明

国家自然科学基金委员会生命科学部, 北京 100085

* 联系人, E-mail: wangpy@nsfc.gov.cn

收稿日期: 2016-11-23; 接受日期: 2016-12-16; 网络版发表日期: 2017-01-22

天然免疫又称为固有免疫, 存在于所有多细胞生物中, 从1883年发现细胞的吞噬现象至今, 天然免疫的研究经历了130多年, 但是直到20世纪90年代, 法国科学家Jules Hoffmann和美国科学家Bruce Beutler根据Charles Janeway Jr.在1989年提出的“模式识别受体”(pattern-recognition receptors, PRRs)假说, 分别在果蝇(*Drosophila melanogaster*)和小鼠(*Mus musculus*)首次发现了模式识别受体分子, 才开启了对天然免疫分子机制的研究. 天然免疫系统通过PRRs对入侵微生物中高度保守的、不存在于健康的宿主细胞中的病原相关分子模式(pathogen-associated molecular patterns, PAMPs)和损伤相关分子模式(damage-associated molecular patterns, DAMPs)进行特异性识别, 从而感知细菌、病毒等外来微生物入侵并激活多种胞内信号通路和基因表达, 启动机体免疫反应. PRRs识别PAMPs导致多信号级联反应最终引起干扰素、促炎症因子和趋化因子的分泌. 此外, PRRs还可以通过识别DAMPs参与感受内源性的“危险”信号.

在高等动物中迄今已经发现了至少5个家族的PRRs, 即Toll样受体(Toll-like receptor, TLR)、视黄酸诱导基因样受体(RIG-I-like receptor, RLR)、NOD-样受体(nucleotide-binding domain-like receptor, NLR)、C型凝集素受体(C-type lectin receptor, CLR)、胞质DNA感受因子(cytosolic DNA sensors). 这几类PRRs通过识别

病原体生存必需的特异性保守成分和机体应激或损伤时释放的结构成分, 诱导I型IFN(type I interferon)、炎性细胞因子、趋化因子和共刺激分子等的释放和表达发挥天然免疫防御功能, 同时诱导获得性免疫的建立, 对成功清除病原体至关重要.

近年来, 随着人们对哺乳动物模式识别受体的发现和研究的不断深入, PRRs对病原体的PAMPs, 尤其是对病原微生物遗传物质核酸(DNA和RNA)的识别已成为天然免疫学研究的热点和重点领域. 核酸是病毒感染过程中最主要的PAMPs. 如果自身核酸的不当识别可能导致自身免疫疾病, 如系统性红斑狼疮(systemic lupus erythematosus, SLE). 了解机体免疫系统对核酸识别和检测的分子基础是天然免疫研究的重要内容. 对天然免疫机制的深入研究已经为阐明机体如何识别并抵抗微生物的入侵提供了重要线索, 但是许多问题和困扰依然存在, 病原体引发机体的多效抗感染机制远比想象复杂, 人们对免疫细胞内与细胞间信号级联网络交叉响应引发的免疫系统动态变化和调控机制研究还处于相当初级的阶段.

病毒RNA可以被RIG-I样受体识别, 即RNA感受因子(RNA sensors), 包括RIG-I, MDA5以及LGP2. 另一类识别病毒RNA的分子是TLR3. TLR3是最早被报道具有病毒核酸识别功能的TLR, 能够识别长度大于40 bp的dsRNA和合成的RNA, 如Poly(I:C)和Poly(U)等.

引用格式: 王璞玥, 杨正宗, 杜生明. STAT6识别N4位点DNA的结构基础被揭示. 中国科学: 生命科学, 2017, 47: 240–243
英文版见: Wang P Y, Yang Z Z, Du S M. The structural basis of DNA recognition by STAT6 was revealed. Sci Sin Vitae, 2017, 47: 240–243, doi: 10.1360/N052016-00324

TLR3不仅可以识别病毒颗粒内的dsRNA,也能识别感染后被吞噬的、已经发生凋亡的细胞内部的病毒核酸。此外,主要分布于巨噬细胞和树突状细胞的TLR7和TLR8也能识别RNA配体。

由于DNA病毒和致病菌感染,以及受损的宿主细胞或逃逸凋亡的细胞进入胞质内并释放dsDNA可以激发天然免疫反应,促使I型干扰素产生以消灭病原体,因此,近年来许多胞内DNA感受因子(DNA sensors)被发现。首先被发现的DNA感受因子是I型膜定位糖蛋白Toll样受体9(Toll-like receptor 9, TLR9),识别未甲基化CpG DNA,在浆细胞样树突状细胞(pDC)中介导I型干扰素产生^[1]。由于人源TLR9的表达仅限于B细胞^[2]和浆细胞样树突状细胞(pDC)^[3]中,不能解释其他类型的细胞如巨噬细胞中DNA诱导的免疫反应。因此,机体内可能还存在其他不依赖于TLR9的胞质DNA识别受体。随后发现细胞内双链DNA感受因子DAI^[4-6]。然而,在几种免疫系统原代细胞中敲除掉DAI基因后,由胞质内B型DNA激发的I型干扰素的产生并没有受到显著影响,说明还存在其他的DNA感受因子^[4,5]。近年来,胞质内DNA的识别机制和信号通路被不断完善,AIM2, RNA聚合酶III(RNA polymerase III), LRRFIP1, IFI16, DDX41, DNA-PKC, Ku70, DHX36/DHX9, DHX36/DDX1/DDX21, cGAS等DNA感受因子相继被发现,其中有些感受因子尚需体内实验证实。值得一提的是,2012年底到2013年Zhijian J. Chen(陈志坚)及其同事^[7-10]在*Science*上连续发表了4篇论文,发现了环化鸟腺苷酸(cGAMP)分子作为宿主内源性的第二信使,在细胞质DNA诱导的内在免疫信号传导中起关键的作用,并分离确定了cGAMP合成酶(cGAS)。证明cGAS作为DNA受体,能与细胞内异源或病原DNA分子(如通过转染或DNA病毒感染)结合,催化生成cGAMP, cGAMP又与接头蛋白STING的C-端结构域(C-terminal domain, CTD)结合,从而激活下游信号通路。

武汉大学舒红兵课题组^[11]在2008年首次发现天然免疫细胞识别DNA病毒信号通路中的关键信号分子MITA(也被命名为STING或ERIS^[12]),当它接受上游DNA信号时,形成寡聚体并从内质网上转移,招募激酶TBK1/IKK ϵ ,激活下游通路引发天然免疫反应。2013年, Zhijian J. Chen(陈志坚)课题组^[13]发现了环化鸟腺苷酸(cGAMP)分子作为宿主内源性的第二信使,与接

头蛋白STING的C-端结构域(C-terminal domain, CTD)结合,导致其构象变化,从而招募并激活下游蛋白激酶TBK1与IKK ϵ ,进而活化干扰素调控因子IRF3及转录因子NF κ B,从而诱导I型干扰素的表达,启动的一系列细胞免疫功能。这些工作阐明了病源DNA引发的免疫应答反应的感应机制^[7-10]。

2011年, Burdette等人^[14]发现STING能通过其CTD直接结合c-di-GMP。2012年, Ouyang等人^[15]在国际上最早解析了STING CTD以及STING CTD与c-di-GMP二元复合物的晶体结构。发现STING CTD具有一种独特的构架(unique architecture),阐明了STING CTD形成功能性二聚体的分子机制,文献报道的最后一个跨膜区(153~173 aa)其实并不是跨膜区而是STING CTD形成同源二聚体的疏水相互作用界面。STING作为天然免疫DNA识别信号通路中的关键分子,其研究工作日新月异,进展迅速。随后STING与cGAMP^[13], DMXAA^[16], CMA^[17]等的复合物结构解析工作相继完成,使人们对STING信号通路的认识变得更加深入和全面。北京大学蒋争凡课题组^[18]继而发现细胞可通过MITA/STING/ERIS-TBK1活化一个转录因子STAT6(signal transducer and activator of transcription 6),从而调控一类趋化因子的表达,招募更多免疫细胞引发抗感染反应。而此前的研究认为, STAT6作为信号转导及转录激活因子家族中的一员,主要在Th2细胞的分化过程中起关键作用:受IL-4/IL-13信号调控,参与JAK-STAT信号通路,将信号从细胞膜传递到细胞核并激活下游一系列基因的转录表达。

信号转导与转录激活因子家族蛋白STAT共有7个成员,在进化上高度保守,含有750~850 aa。所有的STAT蛋白都具有相同的结构域组成^[19]: N端结构域NH2(N-terminal domain)、卷曲螺旋结构域CCD(coiled coil domain)、DBD结构域(DNA binding domain)、LD结构域(linker domain)、SH2结构域(src homology 2 domain)和C端的TAD结构域(trans-activation domain)(图1)。作为转录因子, STAT6首先在胞质中在Y641位点磷酸化后形成二聚体,然后进入细胞核并识别和启动一系列靶基因表达^[20]。哺乳动物STAT家族蛋白识别基因启动区相似的回文DNA序列(基本结构为5'-TTCN_{3/4}GAA-3'),然而STAT蛋白不同成员识别的回文DNA序列的中间间隔区(spacer)长度不同,其中STAT6是唯一一个以识别N4位点DNA(5'-TTCN₄GAA-3')为主同时也能识别

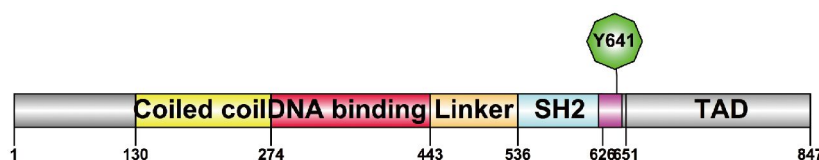


图1 全长STAT6蛋白结构域分布(网络版彩图)

N3位点DNA的家族成员,也有文献报道STAT5能识别少量N4位点DNA,而其他STAT蛋白成员则仅识别N3位点启动子DNA(5'-TTCN₃GAA-3')。对于这种DNA识别的差异一直缺乏比较合理的阐释,也未见文献报道过STAT蛋白哪些区域或氨基酸对于不同DNA类型识别起到关键作用,尤其STAT6对N4位点DNA的识别机制是该领域内过去一直没能解决的重要科学问题。

近日,中国科学院生物物理研究所欧阳松应研究员率领团队^[21]在STAT6蛋白对DNA识别的结构生物学研究取得重要进展,研究论文发表在《美国科学院院报》(*Proceedings of the National Academy of Sciences of the United States of America*)。这项研究采用结构生物学方法,通过一个特殊的原核表达菌获得Y641位点磷酸化蛋白,首次成功解析了人源STAT6核心区(core fragment, STAT6^{CF})的磷酸化同源二聚体晶体结构,同时解析了其N4位点DNA和N3位点DNA两个复合物的晶体结构,这是在以前的STAT蛋白结构生物学研究中没有过的,因此他们有更多的三维结构可以进行分析。与其他结构已解析的STAT蛋白相似,STAT6^{CF}的磷酸化二聚体形成“V”字形构架(architecture),在各结构域组织形式、二聚体形成机制等方面均相似。该研究通过几个结构比对发现,STAT6二聚体结合核酸前后发生一个明显的构象变化。这是首次针对同一STAT蛋白观察到结合DNA前后的构象变化。分子动力学模拟(molecular dynamics simulation)和小角散射实验(small angle X ray scattering, SAXS)进一步表明,STAT6的构象在DNA结合后趋于更稳定的状态。这种构象变化可以形象地比喻为:结合DNA前的STAT6^{CF}的磷酸化二聚体“V”字形看成是含羞草的两片叶子,触碰到DNA底物后,两片叶子自身发生状态变化并同时收紧直到与DNA稳定结合(图2)。

通过与其他已报道STAT蛋白三维结构比对发现,STAT6磷酸化二聚体中“V”形的夹角更大。此外,最主要的差异在于SH2结构域的C端loop(609~620 aa),从

序列和三维结构都能看出STAT1和STAT3在该区域比STAT6短了10个氨基酸。以上2个因素导致二聚体结构中的两个分子之间的动态变化范围比其他STAT蛋白更广,为其能识别更长的DNA奠定了结构基础。有意思的是,他们发现STAT6蛋白的DNA结合结构域上H415氨基酸是唯一与DNA碱基直接相互作用的氨基酸,序列分析发现除STAT5以外,在其他STAT蛋白上H415所对应的氨基酸是N(Asn)。有趣的是,通过结构比较发现在DNA双螺旋轴方向的位置变化距离与DNA双螺旋中相邻两个碱基之间的垂直距离(普遍认为B型双链DNA相邻每碱基对之间沿轴向约上升3.4 Å)相符。进一步的体外实验发现,将STAT6的关键氨基酸H415位点突变成N后,其识别N4位点DNA的能力明显下降而与此同时识别N3位点DNA的能力显著上升;然而,若将STAT1对应的氨基酸N460位点突变成H后,其识



图2 STAT蛋白二聚体(两片叶子)识别N4位点DNA并改变构象直至与DNA稳定结合(网络版彩图)

别N4位点DNA的能力上升而识别N3位点DNA的能力明显下降. 细胞体内实验也进一步证明, STAT6蛋白的H415位点对其特异性结合N4位点DNA具有关键作用. 此外, 该研究还阐释了STAT5等其他STAT蛋白不能广泛识别N4位点DNA可能的原因, 他们认为一方面STAT5二聚体夹角稍小, 另一方面H471的侧链伸展方向与STAT6的H415不同, 不利于与碱基的相互作用.

STAT6分子作为转录因子, 结合核酸后启动下游蛋白表达, 在MITA/STING/ERIS天然免疫信号转导通路发挥重要作用, 同时作为Th2细胞分化相关重要蛋白, 与炎症和多种肿瘤也有紧密联系^[22], 针对STAT6与DNA识别的结构生物学研究, 将为STAT6通路异常所致疾病的潜在预防和治疗药物提供理论基础和可能靶点.

参考文献

- Ishii K J, Coban C, Kato H, et al. A Toll-like receptor-independent antiviral response induced by double-stranded B-form DNA. *Nat Immunol*, 2006, 7: 40–48
- Bourke E, Bosisio D, Golay J, et al. The toll-like receptor repertoire of human B lymphocytes: inducible and selective expression of TLR9 and TLR10 in normal and transformed cells. *Blood*, 2003, 102: 956–963
- Kadowaki N, Ho S, Antonenko S, et al. Subsets of human dendritic cell precursors express different Toll-like receptors and respond to different microbial antigens. *J Exp Med*, 2001, 194: 863–870
- Ishii K J, Kawagoe T, Koyama S, et al. TANK-binding kinase-1 delineates innate and adaptive immune responses to DNA vaccines. *Nature*, 2008, 451: 725–729
- Takaoka A, Wang Z C, Choi M K, et al. DAI (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response. *Nature*, 2007, 448: 501–505
- Wang Z, Choi M K, Ban T, et al. Regulation of innate immune responses by DAI (DLM-1/ZBP1) and other DNA-sensing molecules. *Proc Natl Acad Sci USA*, 2008, 105: 5477–5482
- Gao D, Wu J, Wu Y T, et al. Cyclic GMP-AMP synthase is an innate immune sensor of HIV and other retroviruses. *Science*, 2013, 341: 903–906
- Li X D, Wu J, Gao D, et al. Pivotal roles of cGAS-cGAMP signaling in antiviral defense and immune adjuvant effects. *Science*, 2013, 341: 1390–1394
- Sun L, Wu J, Du F, et al. Cyclic GMP-AMP synthase is a cytosolic dna sensor that activates the type I interferon pathway. *Science*, 2013, 339: 786–791
- Wu J, Sun L, Chen X, et al. Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA. *Science*, 2013, 339: 826–830
- Zhong B, Yang Y, Li S, et al. The adaptor protein MITA links virus-sensing receptors to IRF3 transcription factor activation. *Immunity*, 2008, 29: 538–550
- Sun W, Li Y, Chen L, et al. ERIS, an endoplasmic reticulum IFN stimulator, activates innate immune signaling through dimerization. *Proc Natl Acad Sci USA*, 2009, 106: 8653–8658
- Zhang X, Shi H, Wu J, et al. Cyclic GMP-AMP containing mixed phosphodiester linkages is an endogenous high-affinity ligand for STING. *Mol Cell*, 2013, 51: 226–235
- Burdette D L, Monroe K M, Sotelo-Troha K, et al. STING is a direct innate immune sensor of cyclic di-GMP. *Nature*, 2011, 478: 515–518
- Ouyang S, Song X, Wang Y, et al. Structural analysis of the STING adaptor protein reveals a hydrophobic dimer interface and mode of cyclic di-GMP binding. *Immunity*, 2012, 36: 1073–1086
- Gao P, Ascano M, Zillinger T, et al. Structure-function analysis of STING activation by c[G(2',5')pA(3',5')p] and targeting by antiviral DMXAA. *Cell*, 2013, 154: 748–762
- Cavlar T, Deimling T, Ablasser A, et al. Species-specific detection of the antiviral small-molecule compound CMA by STING. *EMBO J*, 2013, 32: 1440–1450
- Chen H, Sun H, You F, et al. Activation of STAT6 by STING is critical for antiviral innate immunity. *Cell*, 2011, 147: 436–446
- Hebenstreit D, Wirnsberger G, Horejs-Hoeck J, et al. Signaling mechanisms, interaction partners, and target genes of STAT6. *Cytokine Growth Factor Rev*, 2006, 17: 173–188
- Yildiz M, Li H, Bernard D, et al. Activating STAT6 mutations in follicular lymphoma. *Blood*, 2015, 125: 668–679
- Li J, Rodriguez J P, Niu F, et al. Structural basis for DNA recognition by STAT6. *Proc Natl Acad Sci USA*, 2016, 113: 13015–13020
- Wurster A L, Tanaka T, Grusby M J. The biology of Stat4 and Stat6. *Oncogene*, 2000, 19: 2577–2584