

管状花目中多酚氧化酶的序列分析与功能预测

郭宏波, 孙奕玥, 舒志明, 梁宗锁, 张跃进

(旱区作物逆境生物学国家重点实验室, 西北农林科技大学 生命科学学院, 陕西 杨凌 712100)

摘要:以管状花目(Tubiflorae)中5个物种7个多酚氧化酶基因(PPO)序列为研究对象,采用序列对比分析和功能预测方法,研究了这些基因序列的碱基、酶蛋白氨基酸序列数量与比例,以及氨基酸序列中蛋白激酶C磷酸化、酪氨酸激酶II磷酸化、N-糖基化、N-豆蔻酰化、酰胺化、cAMP及cGMP蛋白激酶磷酸化6个功能位点结构域特征,以期揭示PPO酶调控的功能位点。结果表明:管状花目中丹参毛状根PPO基因含有G+C碱基最多,氨基酸中丙氨酸和亮氨酸含量也最高(16.9%),并含有最多的酪氨酸激酶II磷酸化位点(11个)。丹参毛状根PPO酶不具备信号肽结构,有1个卷曲螺旋,是无跨膜结构的亲水性蛋白。该试验为研究PPO酶调控酚酸类物质次生代谢的分子机制,特别是蛋白与蛋白、蛋白与基因之间的互作奠定了基础。

关键词:管状花目;丹参;多酚氧化酶;序列分析;功能预测

中图分类号:Q 946 **文献标识码:**A **文章编号:**1001-0009(2016)12-0085-07

多酚氧化酶(polyphenol oxidase, PPO)是由核基因编码,能与铜离子结合的金属蛋白酶,广泛存在于植物完整细胞的质体中,如丹参根质体、马铃薯块茎造粉体等。已有研究表明,PPO在苯丙烷代谢中发挥了重要作用^[1],并能响应多种生物(如病虫害)与非生物(如干旱)胁迫和信号分子(如茉莉酸)的调控^[2-5]。但其响应机制及在次生代谢调控中发挥的具体作用尚不清楚。

通过试验课题组前期发现,苯丙氨酸、肉桂酸和阿魏酸等酚酸类物质对丹参根中PPO活性具有低浓度促进高浓度抑制的作用^[6]。对PPO活性具有显著提升作用的酵母提取物能显著增加丹参毛状根PPO基因的表达;显著抑制酶活性的半胱氨酸、抗坏血酸和银离子则

显著抑制了PPO基因的表达,且PPO基因表达量上升,丹酚酸B积累增多,反之亦然^[7]。在干旱胁迫下,PPO基因的过表达可使植株中多酚含量增加^[8-9],酚类物质积累越多,抗胁迫能力就越强^[10]。抗性品系的总酚含量和PPO活性都比敏感品系高,在病虫害侵染后,马铃薯抗性品系的PPO活性升高速率比敏感品系快^[11]。这些证据表明,PPO基因和酶对酚酸类物质的积累具有调控作用。

作为PPO反应底物的酚酸类物质在药用植物中也广泛存在,如丹参、金银花和当归等。因PPO直接与酚酸类物质发生反应,是导致许多果蔬褐化的关键酶之一^[12],所以在园艺学中研究较多,主要集中在酶学性质分析、酶活性抑制剂的开发和基因克隆方面^[12],以期能在采后贮藏保鲜中有效抑制酶活性,但成效并不显著。

随着生物信息学的快速发展,对不同物种中PPO基因序列进行分析,并对酶功能位点进行预测,将会有力增强研究的针对性,使后续研究能效的设计试验,开展相关研究。该研究在课题组前期工作基础上,对丹参(*Salvia miltiorrhiza* Bunge)管状花目(Tubiflorae)的7个PPO基因序列进行分析,并对其蛋白序列及功能区域进行分析预测,该结果对阐明PPO的逆境机理和调控酚酸类物质代谢的分子机制,特别是蛋白与蛋白、蛋白与基因之间的互作具有重要意义。

第一作者简介:郭宏波(1977-),男,江苏泰州人,博士,副教授,现主要从事丹参雄性不育分子机制等研究工作。E-mail:hbguo@nwsuaf.edu.cn.

责任作者:张跃进(1960-),男,陕西乾县人,教授,现主要从事中药材规范化栽培的理论与技术等研究工作。E-mail:qxzhang@126.com.

基金项目:国家自然科学基金资助项目(81373908);陕西省自然科学基金资助项目(2015JM3092,2016JM8108);西北农林科技大学基本科研业务资助项目(2016)。

收稿日期:2016-02-26

acids. The theoretical molecular weight was 27.4 kDa and the isoelectric point was predicted as 5.06 and it was a kind of hydrophilic protein. Real-time quantitative PCR was used to detect the relative expression levels of *IDI* in different tissues of *Rabdosia rubescens*. And the expression of *IDI* was relatively high in the leaves, the lowest expression in callus.

Keywords: *Rabdosia rubescens* Henrs. Hara; diterpenoid; *IDI*

1 材料与方法

1.1 试验材料

管状花目中丹参 *PPO* 基因序列(登录号 KF712274)来源于课题组^[7],其余 6 个 *PPO* 基因序列均来源于 GenBank。它们分别是芝麻 *Sesamum indicum* L. (登录号 JQ658360),甘薯 *Ipomoea batatas* (L.) Lam. (登录号 AY822711),马铃薯 *Solanum tuberosum* L. (登录号 U22921),茄子叶绿体 *Solanum melongena* L. chloroplast (登录号 GQ246219),“Anupam”茄子叶绿体 *Solanum melongena* L. cv. Anupam chloroplast (登录号 JQ621952) 和茄子 *Solanum melongena* L. (登录号 GQ149349)。

1.2 试验方法

按国际互联网上提供的各类生物信息学软件进行在线分析。利用 DNAMAN 及 ORF Finder 对 7 个 *PPO* 基因的核酸和氨基酸序列进行在线分析和开放阅读框预测。利用 Prot Param (<http://web.expasy.org/prot-param/>)对 *PPO* 的氨基酸序列进行在线分析。利用 Blast (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>)对 *PPO* 基因的核苷酸序列进行多序列比对。利用 PROSITE (<http://prosite.expasy.org/>)数据库对不同 *PPO* 蛋白功能区域进行预测。利用 Signal P 3.0 Server (<http://www.cbs.dtu.dk/services/SignalP-3.0/>)对丹参毛状根中 *PPO* 的氨基酸序列进行信号肽预测。利用 COILS

Server (http://embnet.vital-it.ch/software/COILS_form.html)对丹参毛状根中 *PPO* 氨基酸序列进行螺旋卷曲预测。利用 TMHMM Server v. 2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>)对丹参毛状根 *PPO* 酶的跨膜序列进行跨膜结构域分析。利用 Prot Scale (<http://web.expasy.org/protscale/>)程序对丹参毛状根 *PPO* 氨基酸序列进行疏水性分析。利用 NCBI (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>)对丹参毛状根中 *PPO* 氨基酸序列进行功能结构域分析。利用 PORTER (<http://distill.ucd.ie/porter>)服务器对丹参毛状根 *PPO* 的氨基酸序列进行二级结构预测。利用 SWISS-MODEL (<http://swissmodel.expasy.org/>)同源建模对丹参毛状根多酚氧化酶基因的氨基酸序列进行蛋白质三维结构预测。

2 结果与分析

2.1 不同物种 *PPO* 基因和酶蛋白的序列分析

2.1.1 *PPO* 基因的序列分析 对管状花目 5 个物种 7 个 *PPO* 基因序列分析的结果显示,除茄属外其余 4 个 *PPO* 基因 5'端均含有一个 15~21 bp 的非翻译区。马铃薯 *PPO* 基因序列中含有最高的碱基 A+T 含量;丹参中碱基 G+C 最多(表 1)。丹参毛状根 *PPO* 基因序列与芝麻、甘薯、马铃薯、茄叶绿体、“Anupam”茄叶绿体和茄子的相似度分别达到 71%、64%、71%、76%、72%和 72%。

表 1 管状花目植物 *PPO* 基因序列分析

Table 1 Sequence analysis of *PPO* genes in Tubiflorae

条目 Item	管状花目 <i>PPO</i> 基因序列 <i>PPO</i> gene sequence in Tubiflorae						
	丹参毛状根	芝麻	甘薯	马铃薯	茄子叶绿体	“Anupam”茄子叶绿体	茄子
	<i>Salvia miltiorrhiza</i>	<i>Sesamum indicum</i> L.	<i>Ipomoea batatas</i> (L.) Lam.	<i>Solanum tuberosum</i> L.	<i>Solanum melongena</i> L. chloroplast	<i>Solanum melongena</i> L. cv. Anupam chloroplast	<i>Solanum melongena</i> L.
基因全长 Full-length/bp	1 930	1 947	1 984	1 958	1 848	1 788	1 782
起始密码子 Initiation codon	ATG	ATG	ATG	ATG	ATG	ATG	ATG
终止密码子 Termination codon	TGA	TAA	TAG	TAA	TAA	TAA	TAA
编码区位置 Position of coding region	19~1 788	19~1 773	16~1 782	22~1 815	1~1 803	1~1 787	1~1 781
碱基 A 数量 Number of base A	478	473	484	552	531	514	512
碱基 C 数量 Number of base C	633	545	565	389	413	395	391
碱基 G 数量 Number of base G	462	468	514	430	404	403	401
碱基 T 数量 Number of base T	357	461	421	587	500	476	478
碱基 G+C 含量							
Percentage of base G and base C / %	56.7	52.0	54.4	41.8	44.2	44.6	44.5
碱基 A+T 含量							
Percentage of base A and base T / %	43.3	48.0	45.6	58.2	55.8	55.4	55.5

2.1.2 *PPO* 酶蛋白氨基酸序列分析 利用 Prot Param 对 7 个 *PPO* 酶蛋白氨基酸序列进行分析后(表 2),发现其氨基酸残基数在 584(芝麻)至 600(茄子叶绿体)之间。分子量在 65.31 kDa(芝麻)至 67.67 kDa(茄子叶绿体)范围内。对 *PPO* 酶蛋白中含量最高的 4 种氨基酸分析后发现,丙氨酸(Ala)和亮氨酸(Leu)大量存在于 *PPO* 酶蛋白的氨基酸序列中,2 种氨基酸的含量在 14.1%(马铃

薯)至 16.9%(丹参)之间。在 7 种酶蛋白序列中,酸性氨基酸比例、碱性氨基酸比例、带正电荷氨基酸比例以及蛋白质不稳定性指数等 5 项指标的最大值均出现于丹参中,而最小值则均出现在马铃薯中。7 种多酚氧化酶蛋白的极性氨基酸比例均大于 50%,表明 *PPO* 酶蛋白属于亲水蛋白。而该推测在利用 Prot Scale 程序对丹参毛状根 *PPO* 酶蛋白的氨基酸序列进行亲水/疏水性

分析后得到证实。

2.1.3 PPO 酶蛋白功能区域预测 对不同物种 PPO 酶蛋白的蛋白激酶 C 磷酸化位点、酪氨酸激酶 II 磷酸化位点、N-糖基化位点、N-豆蔻酰化位点、酰胺化位点、cAMP 及 cGMP 蛋白激酶磷酸化位点 6 个功能区域进行了预测分析(表 3)。结果表明,茄子叶绿体 PPO 酶蛋白中含有最多的蛋白激酶 C 磷酸化位点(10 个),甘薯的蛋白激酶 C 磷酸化位点最少(4 个)。除丹参和芝麻外,其余 5 条 PPO 酶蛋白中蛋白激酶 C 磷酸化位点均包含

TPK 位点(16~26 位氨基酸)。除 TPK 位点外,所有茄属植物 PPO 酶蛋白均包含 SCK 及 SSR 位点。丹参含有最多的酪氨酸激酶 II 磷酸化位点(11 个),芝麻含有的酪氨酸激酶 II 磷酸化位点最少(3 个)。7 条 PPO 酶蛋白中 SN(G)V(I/E)D 结构域存在于第 341~376 位氨基酸中,其中 5 条是 SNVD 序列。第 383~396 位氨基酸的 T(S)D(H/Q)P(S/R/K)D 结构域以多种变化形式存在于 7 条 PPO 酶蛋白中。对 N-豆蔻酰化位点而言,芝麻含有的位点数最多(9 个),而茄子叶绿体和

表 2
Table 2
Analysis of PPO amino acid sequences in Tubiflorae plants

项目 Item	管状花目植物中多酚氧化酶基因氨基酸序列 PPO amino acid sequence of Tubiflorae plants						
	丹参毛状根 <i>Salvia miltiorrhiza</i>	芝麻 <i>Sesamum indicum</i> L.	甘薯 <i>Ipomoea batatas</i> (L.) Lam.	马铃薯 <i>Solanum tuberosum</i> L.	茄子叶绿体 <i>Solanum melongena</i> L. chloroplast	“Anupam”茄子叶绿体 <i>Solanum melongena</i> L. cv. Anupam chloroplast	茄子 <i>Solanum melongena</i> L.
氨基酸数目 Number of amino acids	589	584	588	597	600	595	593
分子量 Molecular weight/ kDa	65.88	65.31	65.66	66.89	67.67	66.69	66.91
理论等电点 Theoretical Pi	6.56	7.20	6.27	6.63	7.22	6.23	6.48
含量最丰富的氨基酸 Rich amino acids/ %	Ala 9.3	Leu 8.6	Leu 8.5	Ser 7.7	Leu 7.8	Leu 7.7	Leu 8.6
	Pro 9.2	Pro 8.0	Pro 8.0	Asn 7.5	Ala 6.8	Ser 7.2	Ser 7.8
Asp 8.7	Asp 8.7	Asp 8.0	Ala 7.8	Leu 7.2	Pro 6.8	Pro 6.9	Ala 6.6
	Leu 7.6	Ala 7.5	Lys 7.5	Ala 6.9	Thr 6.8	Ala/ Asn 6.7	Asn 6.4
酸性氨基酸比例 Acidic amino acid/ %	12.8	11.1	12.7	10.5	10.8	11.3	11.2
碱性氨基酸比例 Basic amino acid/ %	15.3	13.4	14.5	12.9	13.3	13.0	13.1
带负电荷氨基酸比例 Negative charged amino acid/ %	12.73	11.13	12.76	10.55	10.83	11.26	11.13
带正电荷氨基酸比例 Positive charged amino acid/ %	12.05	11.13	11.73	10.05	10.83	10.08	10.46
极性氨基酸比例 Polar amino acids/ %	54.7	54.6	54.6	57.4	57.3	57.0	58.0
疏水性氨基酸比例 Hydrophobic amino acids/ %	45.3	45.4	45.4	42.6	42.7	43.0	42.0
脂类结合蛋白质指数 Aliphatic index/ %	73.94	76.49	76.99	69.45	73.32	71.63	72.38
不稳定系数 Instability index	49.17	39.45	47.45	36.02	45.75	41.48	42.68

表 3
Table 3
The functional domains of PPO amino acid sequence of Tubiflorae plants

项目 Item	管状花目植物的多酚氧化酶基因氨基酸序列 PPO amino acid sequence of Tubiflorae plants						
	丹参毛状根 <i>Salvia miltiorrhiza</i>	芝麻 <i>Sesamum indicum</i> L.	甘薯 <i>Ipomoea batatas</i> (L.) Lam.	马铃薯 <i>Solanum tuberosum</i> L.	茄子叶绿体 <i>Solanum melongena</i> L. chloroplast	“Anupam”茄子叶绿体 <i>Solanum melongena</i> L. cv. Anupam chloroplast	茄子 <i>Solanum melongena</i> L.
酪氨酸酶 CuA 结合区 Tyrosinase CuA-binding region	176 HCAYCDGA YDQAGFPDLE LQVHNSWLF FPFH	179 HCAYCNGA YDQPGQGS DLQIHNSWLF FPFH	178 HCAYCNGG YVQTDYPDK EIQVHNSWLF FPFH	182 HCAYCNGA YRIGGKELQV HNSWLFFPH	184 HCAYCNGA YKIGGKELQV HNSWLFFPH	178 HCAYCNGA YKIGDKVLQV HNSWLFFPH	178 HCAYCNGA YKIGDKVLQV HNSWLFFPH
	328 HGPVHVWT GDRTQNTEN MGNFYSAARD PIFYAHH	327 HTAIHVVF GDPREPSGEDL GNFYSAGRDP LFYCHH	330 HTPVHRWV GDVQPRQTNG EDMGNFYSA RDILFYCHH	330 HGPVHIWS GTVRGSTLPN GAISNGENMG HFYSAGLDPV FFCHH	329 HIPVHIWVG TARGSKFPDGS TSYGEDMGNF YSAGLDPVFY CHH	325 HGPVHVWT GTVPGTTLPNG RTSHGENMGH FYSAGLDPVFF HH	325 TVLSTFGLV QCQVQPCNLGR TSHGENMGHF YSAGLDPVFFC HH
	65 DRRNLLIGL GGLYG	65 DRRNVLLGL GGGLYG	63 DRRDVLGL GGLAG	65 DRRNVLLGL GGLYG	66 DRRNVLLGL GGLYG	61 DRRNVLLGL GGMYG	61 DRRNVLLGL GGMYG
类囊体转运结构域 Thylakoid transfer domain	107 AGA	89 ASA	86 AFA	87 ASA	88 ASA	83 ASA	83 ASA
DWL 结构域 DWL motif	384 DPDFL	387 DPDFL	390 DSDWL	394 HKDWL	396 QRDWL	390 HKDWL	389 HKDWL
酪氨酸结构域 Tyrosine motif	417 YAY	419 YDF	423 YKY	427 YDY	430 YDY	423 YDY	422 YDY
KFDV 结构域 KFDV motif	505 KFDV	505 KFDV	504 KFDV	508 RFDV	511 RFDV	504 RFDV	503 RFDV
KFDV 结构域中谷氨酸 结构域 Glu enriched motif in KFDV domain	486 ELEELVI	479 ELLVLENI	484 EAEELIEI	491 EEMLTFSI	494 EEMLTFSI	487 EEMLTFSI	486 EEMLTFSI

表 3(续)

Table 3(Continued)

项目 Item	管状花目植物的多酚氧化酶基因氨基酸序列 PPO amino acid sequence of Tubiflorae plants						
	丹参毛状根 <i>Salvia miltiorrhiza</i>	芝麻 <i>Sesamum indicum</i> L.	甘薯 <i>Ipomoea batatas</i> (L.) Lam.	马铃薯 <i>Solanum tuberosum</i> L.	茄子叶绿体 <i>Solanum melongena</i> L. chloroplast	“Anupam”茄子叶绿体 <i>Solanum melongena</i> L. cv. Anupam chloroplast	茄子 <i>Solanum melongena</i> L.
第 3 个 Cu 结合结构域中 EFAGSF 结构域 EFAGSF motif in hypothesized third copper-binding domain	524 EFAGSF	513 AYAGTY	523 EYAGSF	528 EFAGSY	531 EFAGSY	524 EFAGSY	523 EFAGSY
蛋白激酶 C 磷酸化位点 Protein kinase C phosphorylation site	35 SSK 42 SYK	37 SNR	16 TPK	11 SLK	11 TLK	21 TPK	21 TPK
	46 SCK 168 SFK	123 TDR	40 TAK	26 TPK	26 TPK	36 TFK	36 TFK
	301 TPR 432 TPR	322 SER	440 TKK	45 SCK 263 SHR	45 SCK	40 SCK	40 SCK
	455 SPR	377 SNK	538 SKK	339 TVR	314 TLR	259 SHR	259 SHR
	480 SKK	408 TVK		393 THK	321 TLK	389 SHK	313 TLK
		438 TVK		422 TKK 480 SSR	342 TAR	418 TKK	388 SHK
		490 TTK		498 SIR	396 SQR	476 SSR	417 TKK
		534 SIR			450 SGK		475 SSR
					477 SIK		
					483 SSR		
酪氨酸激酶 II 磷酸化位点 Casein kinase II phosphorylation site	50 SNHD	341 SGED	275 TDTD	254 SLFD	105 TISD	102 TETE	102 TETE
	61 SKFD	386 TDPD	280 TDDE	351 SNGE	270 TVMD	115 TPDD	115 TPDD
	93 SGPD	408 TVKD	367 SNVD	373 SNVD	276 SFGD	218 SIID	218 SIID
	272 SGTD		389 TDSO	393 THKD	354 SYGE	270 SFGD	270 SFGD
	363 SNID		479 SKEE	483 TQQE	376 SNVD	347 SHGE	346 SHGE
	383 TDPD			513 SNVD	396 SQRD	369 SNVD	368 SNVD
	455 SPRE			593 SLAD	486 TGGE	389 SHKD	388 SHKD
	480 SKKE				546 SQT D	479 TQQE	478 TQQE
	518 TTAE				596 SLAD	590 SLAD	589 SLAD
	549 SISD						
N-豆蔻酰化位点 N-myristoylation site	578 TIHD						
	73 GGLYGA	47 GQNQSN	72 GGLAGA	74 GGLYGV	75 GGLYGV	70 GGMVMA	70 GGMVGA
	74 GLYGAT	72 GLGGGL	73 GLAGAA	78 GVANAI	79 GVANAI	71 GMYGAA	71 GMYGAA
	108 GATPTN	75 GGLYGA	323 GSIENI	323 GTIENI	341 GTARGS	334 GTVPGT	331 GLVQCQ
	273 GTDPAT	76 GLYGAA	351 GNFYSA	338 GTVRGS	360 GNFYSN	527 GSYTNL	526 GSYTNL
	321 GSIEGT	103 GTATNL	494 GIEVAI	348 GAISNG	534 GSYTSL	586 GATISL	585 GATISL
	347 GNFYSA	272 GMTGNT	574 GGDVTV	386 GKKRTD	592 GVEISL		
	527 GSFVNY	325 GSHTAI		531 GSYTSL			
		346 GNFYSA		589 GATISL			
		566 GVTIGG					
N-糖基化位点 N-glycosylation site	270 NYSG	49 NQSN		6 NSSS	268 NGTV	6 NSSI	6 NSSI
	287 NLTV	52 NSSQ		290 NLTL	292 NLTL	257 NQSH	257 NQSH
		249 NDTK			454 NTSA	286 NLTL	286 NLTL
		286 NLTI					
		392 NASF					
		485 NITV					
酰胺化位点 Amidation site	376 GGRR		382 KGRR	35 HGKR	35 HGKR	30 HGKR	30 HGKR
				386 GGKR	389 GGKR	382 GGKR	381 GGKR
cAMP 及 cGMP 蛋白 激酶磷酸化位点 cAMP- and cGMP-dependent protein kinase phosphorylation site	477 KKRS	34 KRRS 383 KRIT	441 KKKT		261 RRST	33 RKQT	33 RKQT

“Anupam”茄子叶绿体中仅含 5 个。7 条 PPO 酶蛋白序列中,第 70~75 位氨基酸都存在 GGL(M)Y(A)G(M)A(V)位点。芝麻也含有最多的 N-糖基化位点(7 个),但在甘薯中未发现 N-糖基化位点。在所有 PPO 酶蛋白中,第 286~292 位氨基酸都具有 NLTV(I/L)结构域,在茄属植物中均表现为 NLTL。除芝麻外的 6 条 PPO 酶蛋白中,第 376~389 位氨基酸都含有酰胺化作用位点 G(K)GR(K)R。在茄属植物中该位点表现为 GGKR。除马

铃薯以外的 6 条 PPO 酶蛋白中,均含有 cAMP 及 cGMP 蛋白激酶磷酸化位点 K(R)K(R)R(K/S/Q)T(S)。

2.2 丹参毛状根中 PPO 酶蛋白功能分析与预测

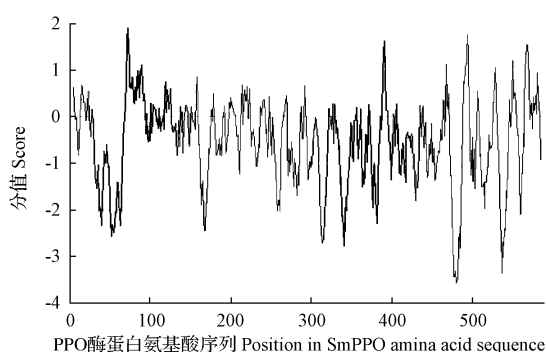
2.2.1 信号肽预测 利用 Signal P 3.0 server 对丹参毛状根中 PPO 酶蛋白的信号肽进行预测,发现多肽链第 19 位的丙氨酸(Ala)残基具有最高剪切位点分值 0.124,第 14 位的苏氨酸(Thr)残基具有最高信号肽分值 0.486。综合起来看,第 19 位的丙氨酸(Ala)残基具有最

高综合剪切点分值 0.141, 分值小于平均分 0.5。由此可知, 丹参毛状根 PPO 酶蛋白的氨基酸序列不具有信号肽。由此推断, 丹参 PPO 在细胞质中的核糖体上合成后, 直接存在于细胞质基质中。

2.2.2 卷曲螺旋预测 利用 COILS Server 对丹参毛状根中 PPO 酶蛋白的氨基酸序列进行螺旋卷曲预测。结果发现, 多肽链第 476~491 位的 16 个氨基酸残基分值区域为 [0.95, 1]。由此判断, 丹参毛状根中 PPO 酶蛋白存在一个卷曲螺旋结构。

2.2.3 跨膜区域预测 利用 TMHMM Server v. 2.0 对丹参毛状根 PPO 酶蛋白的跨膜序列预测后, 发现第 1~589 位氨基酸位于细胞膜表面。由此推断, 整条肽链位于细胞膜表面, 不存在跨膜螺旋结构。

2.2.4 氨基酸序列的亲水/疏水性分析 利用 Prot Scale 程序对丹参 PPO 酶蛋白氨基酸序列进行亲水/疏水性分析。结果表明, 第 480 位丝氨酸(Ser)具有最小值 -3.578, 亲水性最强。第 494 位赖氨酸(Lys)具有最大值 1.989, 疏水性最强。总体上, 亲水性氨基酸明显多于疏水性氨基酸(图 1)。由此可知, 丹参毛状根中 PPO 酶属于亲水蛋白。



注: 在分值 0 以上的为疏水性氨基酸, 以下为亲水性氨基酸。

Note: The amino acids above zero score are hydrophobic, while those below are hydrophilic ones.

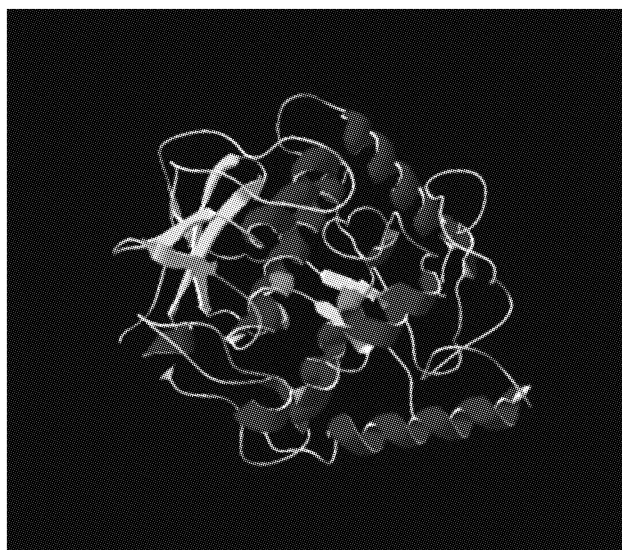
图 1 丹参毛状根中 PPO 酶蛋白氨基酸序列的亲水/疏水性分析

Fig. 1 The distribution of hydrophilic and hydrophobic amino acids in SmPPO protein, which PPO gene was cloned from *Salvia miltiorrhiza* hairy root

2.2.5 酶蛋白二级结构和三级结构的预测 利用 PORTER 服务器对丹参毛状根 PPO 酶蛋白的二级结构进行预测。结果显示, 丹参毛状根 PPO 酶蛋白中含 α 螺旋 27.16%, β -折叠 14.77% 以及无规则卷曲 58.06%。利用 SWISS-MODEL 同源建模对丹参毛状根 PPO 酶蛋白的三维结构进行预测。结果发现, 保守区三维结构骨架由 8 段 α 螺旋和 7 段 β -折叠串通无规则卷曲连接而成, α 螺旋和 β -折叠分散在整个肽链中, 其高度保守的 Cu 离子结合区嵌在蛋白活性腔中(图 2)。

3 讨论

PPO 酶蛋白是核基因编码的类囊体内腔蛋白, 由分



注: 红色为 α 螺旋; 黄色为 β 折叠; 绿色为保守的组氨酸位点。

Note: The α -helices are red; β -sheets are yellow; conserved histidines are green.

图 2 丹参毛状根中 PPO 酶蛋白的三级结构

Fig. 2 The three dimensional structure of SmPPO protein, which SmPPO gene was cloned from

Salvia miltiorrhiza hairy root

离的叶绿体经过 2 步反应后产生的^[13]。典型的植物 PPO 酶蛋白包含 3 个结构域: N 端叶绿体转运肽(cTP), Cu 结合区和 C 端伸展区。管状花目植物的 PPO 酶蛋白中, 前 35 个氨基酸含有大量丝氨酸残基, 而这正是 N 端叶绿体转运肽中基质肽段的典型特征, 这在陆生植物中已被证实^[14]。N 端叶绿体转运肽提供靶位点, 蛋白质靠此位点进入叶绿体中, 类囊体转运结构域(TTD)是 N 端叶绿体转运肽的重要组成部分^[15]。TTD 包含 3 个结构域: 带电的 N 端, 疏水性中心和亲水性 C 端^[16]。依赖类囊体转运的双精氨酸结构域具有普遍性, 即在疏水性中心前存在成对的精氨酸(RR)^[17]。该研究 7 个 PPO 酶蛋白类囊体转运结构域中, 均含 DRRxXXXXGGLY(M)G 结构域, 而该结构域除与转运有关外, 还与酶活性密切相关。管状目植物中 PPO 酶类囊体转运结构域的亲水性 C 端包含 2 个 N-豆蔻酰化位点: GGL(M)Y(A)G(M)A(V)和 GL(M)Y(A)GAT(A)。豆蔻酰化是通过由 N-肉豆蔻酰基转移酶催化的修饰蛋白, 将 G 蛋白附着在质膜内表面, 该情况通常出现在甘氨酸残基中^[18]。PPO 酶蛋白 CuA 结合区末端的 GGLY(M)G 结构域起豆蔻酰化作用。大约 75% 陆生植物的 PPO 酶蛋白中都存在质体定位结构域^[14]。

铜结合区包含 2 个保守的铜结合位点(CuA 和 CuB), 每个都包含 3 个组氨酸残基。组氨酸残基不但可以调控 Cu 离子, 而且形成了活性区域^[19]。CuA 结合区由大约 30 个氨基酸组成, CuB 结合区由大约 40 个氨基酸组成, 2 个结合区之间有一个由 120 个氨基酸残基组

成的连接区域。尽管 2 个结构域都是保守的,且都是 PPO 酶蛋白的特征区域,但相对于 CuA 结合区而言,CuB 结合区的可变性更强,而且这种可变性也会影响基质的偏向^[4]。在活性位点处,组氨酸与精氨酸残基中都有基质反应结构域,用乙酸溴盐和乙酰丙酮修饰其咪唑基团与胍基团后,苹果 PPO 酶活性降低 20%~30%^[20]。该研究中的 7 条 PPO 酶蛋白的 CuA 结合区中,都存在 3 个组氨酸残基。此外,在 CuA 结合区中,并未发现有蛋白激酶 C 磷酸化位点、酪氨酸激酶 II 磷酸化位点、N-糖基化位点、N-豆蔻酰化位点、cAMP 及 cGMP 蛋白激酶磷酸化位点。以上预测表明,CuA 结构域的主要功能是与基质发生反应。位于 CuA 结合区 N 端的 HCAYC 结构域中有 2 个半胱氨酸,进而形成二硫键,有增强蛋白结构稳定性的功能^[21]。

其它的豆蔻酰化位点、蛋白激酶 C 磷酸化位点、酪氨酸激酶 II 磷酸化位点、N-糖基化位点都被发现靠近或位于 CuB 结合区中。在所有 PPO 酶蛋白的 CuA 和 CuB 结合区连接处,距离 CuB 结合区不远的糖基化位点都有 NxT(S)x 结构域。现已发现,HxxxH 结构域的保守组氨酸残基位于 CuB 结合区的起始处,该研究也发现管状花目 PPO 存在 HxxV(I)HxxxGD(T)结构域。从以上结果可知,这 2 个组氨酸残基,CuB 结合区中的第 4 个组氨酸残基,在结合 Cu 离子后共同构成酶活性区域。现已证实,蛋白激酶 C 可以通过丝氨酸和苏氨酸残基中羟基的磷酸化作用控制其它蛋白质的功能,并可以通过增强钙离子或甘油二酯浓度等信号,使之被激活^[22]。与该研究相对应的,靠近或位于 CuB 结合区的蛋白激酶 C 磷酸化位点发现有 S(T)xR 结构域。

研究发现,当色氨酸残基(W)被 $0.14 \text{ mmol} \cdot \text{L}^{-1}$ 的 N-溴代琥珀酰亚胺(NBS)修饰后,苹果多酚氧化酶活性完全丧失^[20]。在该研究所有 PPO 酶蛋白中,色氨酸残基以 NSWLFFPFH 模体形式存在于 CuA 结合区的 C 端,但也有少数位于 CuB 结合区和/或 DWL 结构域中。这表明 CuA 结合区 C 端有维持酶催化反应活性的功能^[4]。有趣的是,在 PPO1_DWL 结构域中,酪氨酸激酶 II 磷酸化位点与 DWL 结构域紧密连接在一起。酪氨酸激酶 II 是一种丝氨酸/苏氨酸蛋白激酶,它可以使酪蛋白等酸性蛋白质磷酸化。它的 α 亚基具有催化活性,而 β 亚基可以自身磷酸化。自身磷酸化是一种转录后的蛋白修饰过程。通常定义为激酶自身的磷酸化作用。在真核细胞中,这一过程通过使蛋白激酶中的丝氨酸、苏氨酸或酪氨酸残基磷酸化,达到调控催化活性的目的^[23]。在该研究 7 条 PPO 酶蛋白的酪氨酸激酶 II 磷酸化位点处,都存在 S(T)xxD(E)模体。天冬氨酸(D)和谷氨酸(E)都属于酸性氨基酸。所以,DWL 结构域具有调节 PPO 酶活性的作用。

在 PPO1_DWL 结构域中,最接近 C 端的酪氨酸模体 YD(A/D)Y 是蛋白水解过程中的分裂位点^[24]。该反

应之后,1 个大约 16~18 kDa、包含 PPO1_KFDV 结构域的多肽片段消失^[25]。在管状花目 PPO1_KFDV 结构域中,存在富含谷氨酸的 EL(A/E)xxILV(N/E)I、EF(Y)AGSF(Y)以及 K(R)FDV 模体。根据酪氨酸激酶 II 磷酸化位点位于富含谷氨酸的 EL(A/E)xxILV(N/E)I 结构域之前这一结果,推测该结构域可能与酪氨酸激酶 II 磷酸化有关,并形成细胞外基质的成分,通过蛋白多糖中大量多糖复合物之间的相互作用,使一个细胞黏附于其它细胞。pH 为 5 时,谷氨酸的质子化作用使紫黄素去环氧酶的催化活性最强^[26]。在该研究中有 5 条 PPO 酶蛋白的 EF(Y)AGSF(Y)结构域是一个豆蔻酰化位点,其 C 端 GSF(Y)是 GSF(Y)xxx 结构域的 N 端。这表明该结构域与豆蔻酰化相关,即可以将活化的 G-蛋白固定在质膜内表面。体外试验证明,在低 pH 条件下,C 端的特殊分裂存在蛋白酶^[27]。研究发现,当 pH 从 3 升至 10 后, β -转角结构和无规则卷曲所占百分比有所增加,而 α -螺旋和 β -折叠所占百分比却有所下降。这说明 pH 影响 PPO 的二级结构及其活性。另一方面,在某些情况下的酶活化可能需要对 PPO 酶蛋白的 C 端进行加工^[27]。此外,在 EFAGSF 结构域末端,最靠近 C 端的组氨酸以及 4 个另外的组氨酸残基被预测共同组成第 3 个铜结合区^[28]。第 3 个铜结合区推测可能与氧分子发生反应,而 NO 可以降低它与氧分子结合的活性^[29]。

该研究结果表明,丹参 PPO 酶蛋白包含 27.16% α -螺旋、14.77% β -折叠和 58.06% 无规则卷曲。这与莲藕和蘑菇中 PPO 酶蛋白一致^[12,30]。当 PPO 与莲藕中酚类物质发生反应时,它的 α -螺旋从 39.3% 降低至 12.3%,而 β -折叠与无规则卷曲则有所增加^[31]。 α -螺旋的降低预示 PPO 酶活性降低,而这种情况在蘑菇、烟草及莲藕中皆有发现^[12,30,32]。因此, α -螺旋中可能包含活性位点,而 β -折叠和 β -转角则是催化基质或失活的延长结构^[12,31]。位于所有 PPO 酶蛋白 CuA 结合区 N 端和 C 端的 HCAYC 和 FFPFH 结构域都位于 α -螺旋中,意味着它们可能与 PPO 酶活性相关。相比之下,CuB 结合区的 33 个氨基酸中有一半都位于 α -螺旋中,这表明该结构域是 PPO 酶维持催化底物活性的关键结构域^[31]。CuA 和 CuB 结合区都只有 α -螺旋和无规则卷曲,这就暗示当 α -螺旋显著降低时,无规则卷曲可能会有明显增加。同时, β -折叠可能位于丹参以及莲藕 PPO 酶蛋白的 C 端,这说明 β -折叠需要保持延长结构,而 PPO 酶在低 pH 条件下失活或与基质发生反应后, β -折叠会有不明显的增加^[31-32]。

参考文献

- [1] KOJIMA M, TAKEUCHI W. Detection and characterization of *p*-coumaric acid hydroxylase in mung bean, *Vigna mungo*, seedlings[J]. J Biochem, 1989, 105: 265-270.
- [2] CONSTABEL C P, BERGEY D R, RYAN C A. Systemin activates synthesis of wound-inducible tomato leaf polyphenol oxidase via the octadecanoid defense signaling pathway[J]. Proc Natl Acad Sci USA, 1995, 92: 407-411.

- [3] THIPYAPONG P, STEFFENS J C. Tomato polyphenol oxidase; differential response of the polyphenol oxidase F promoter to injuries and wound signals[J]. Plant Physiol, 1997, 115: 409-418.
- [4] 王曼玲, 胡中立, 周成全, 等. 植物多酚氧化酶的研究进展[J]. 植物学通报, 2005, 22(2): 215-222.
- [5] MAHANIL S, ATTAJARUSIT J, STOUT M J, et al. Overexpression of tomato polyphenol oxidase increase resistance to common cutworm[J]. Plant Sci, 2008, 174: 456-466.
- [6] 宛国伟, 董娟娥, 梁宗锁, 等. 培养条件对离体丹参根苯丙氨酸解氨酶和多酚氧化酶活性的影响[J]. 西北植物学报, 2007, 27(12): 2471-2477.
- [7] 孙奕珏. 丹参多酚氧化酶基因的克隆及其表达分析[D]. 杨凌: 西北农林科技大学, 2014.
- [8] THIPYAPONG P, HUNT M D, STEFFENS J C. Antisense downregulation of polyphenol oxidase results in enhanced disease susceptibility[J]. Planta, 2004, 220: 105-107.
- [9] THIPYAPONG P, MELKONIAN J, WOLFE D W, et al. Suppression of polyphenol oxidase increases stress tolerance in tomato[J]. Plant Sci, 2004, 167: 693-703.
- [10] KUC J. Induced immunity to plant disease[J]. Bioscience, 1982, 32(11): 854-860.
- [11] THYGESSEN P W, DRY I B, ROBINSON S P. Polyphenol oxidase in potato. A multigene family that exhibits differential expression patterns[J]. Plant Physiol, 1995, 109: 525-531.
- [12] 张跃进, 郝晓燕, 梁宗锁, 等. 莲藕多酚氧化酶基因(PPO)的克隆与表达分析[J]. 农业生物技术学报, 2011, 19(4): 634-641.
- [13] SOMMER A, NE'EMAN E, STEFFENS J C, et al. Import, targeting, and processing of a plant polyphenol oxidase[J]. Plant Physiol, 1994, 105: 1301-1311.
- [14] TRAN L T, TAYLOR J S, CONSTABEL C P. The polyphenol oxidase gene family in land plants: lineage-specific duplication and expansion[J]. BMC Genomics, 2012, 13: 395.
- [15] de BOER A D, WEISBEEK P J. Chloroplast protein topogenesis: import, sorting and assembly[J]. Biochim Biophys Acta, 1991, 1071: 221-253.
- [16] VON HEIJNE G, STEPPUHN J, HERRMANN R G. Domain structure of mitochondrial and chloroplast targeting peptides[J]. Eur J Biochem, 1989, 180: 535-545.
- [17] ROBINSON C, BOOLHUIS A. Protein targeting by the twin-arginine translocation pathway[J]. Nat Rev Mol Cell Biol, 2001, 2: 350-356.
- [18] ZHA J, WEILER S, OH K J, et al. Posttranslational N-myristoylation of BID as a molecular switch for targeting mitochondria and apoptosis[J]. Science, 2000, 290: 1761-1765.
- [19] KLABUNDE T, EICKEN C, SACCHETTINI J C, et al. Crystal structure of a plant catechol oxidase containing a dicopper center[J]. Nat Struct Mol Biol, 1998, 5: 1084-1090.
- [20] 彭益强, 邓峰, 刘宇, 等. 富士苹果中多酚氧化酶活性的中心必需基团与抑制动力学[J]. 华侨大学学报(自然科学版), 2012, 33(1): 51-54.
- [21] 韦朝领, 江昌俊, 陶汉之, 等. 植物紫黄素脱环氧化酶结构特征的生物信息学分析[J]. 安徽农业大学学报, 2002, 29(3): 250-255.
- [22] NEWTON A C. Protein kinase C: structure, function, and regulation[J]. J Biol Chem, 1995, 270: 28495-28498.
- [23] GAO Y, WANG H Y. Casein kinase 2 is activated and essential for Wnt/beta-catenin signaling[J]. J Biol Chem, 2006, 281(27): 18394-18400.
- [24] MARUSEK C M, TROBOUGH N M, FLURKEY W H, et al. Comparative analysis of polyphenol oxidase from plant and fungal species[J]. J Inorg Chem, 2006, 100: 108-123.
- [25] ROBINSON S P, DRY I B. Broad bean leaf polyphenol oxidase is a 60-kilodalton protein susceptible to proteolytic cleavage[J]. Plant Physiol, 1992, 99: 317-323.
- [26] 潘宇, 王坤波, 徐仲溪, 等. 茶叶多酚氧化酶的序列分析与结构预测[J]. 茶叶科学, 2008, 28(3): 157-165.
- [27] RATHJEN A H, ROBINSON S P. Aberrant processing of polyphenol oxidase in a variegated grapevine mutant[J]. Plant Physiol, 1992, 99: 1619-1625.
- [28] STEFFENS J C, HAREL E, HUNT M D. Polyphenol oxidase[M]// ELLIS B E. Genetic engineering of plant secondary metabolism. New York: Plenum Press, 1994: 275-312.
- [29] TORRES J, SVISTUNENKO D, KARLSSON B, et al. Fast reduction of a copper centre in laccase by nitric oxide and formation of a peroxide intermediate[J]. J Amer Chem Soc, 2002, 124: 963-967.
- [30] 易建勇, 董鹏, 王永涛, 等. 应用 SRCD 和 FTIR 分析超高压处理对蘑菇多酚氧化酶二级结构的影响[J]. 光谱学与光谱分析, 2012, 32(2): 318-322.
- [31] 李春美, 胡婉峰, 杨立, 等. 莲藕多酚氧化酶与底物或抑制剂相互作用的光谱学研究[J]. 分析实验室, 2010, 29(8): 14-17.
- [32] 肖厚荣, 徐小龙, 解永树, 等. pH 诱导烟草多酚氧化酶二级结构变化的光谱学研究[J]. 化学物理学报, 2004, 17(2): 196-200.

Sequence Analysis of Polyphenol Oxidase in Tubiflorae Plants and Function Prediction

GUO Hongbo, SUN Yiyue, SHU Zhiming, LIANG Zongsuo, ZHANG Yuejin

(State Key Laboratory of Crop Stress Biology for Arid Areas/College of Life Sciences, Northwest Agricultural and Forestry University, Yangling, Shaanxi 712100)

Abstract: The sequences of seven polyphenol oxidase genes(PPO) in five species of Tubiflorae was analyzed by using the methods of sequence comparison and functional domain prediction. The number of bases and amino acids, and six functional sites including protein kinase C phosphorylation site, casein kinase II phosphorylation site, N-myristoylation, N-glycosylation, amidation site and cAMP- and cGMP-dependent protein kinase phosphorylation site, were analyzed to discover those regulation sites of PPO. The results showed that among Tubiflorae plants, PPO gene cloned from hair roots of *Salvia miltiorrhiza* Bunge (*SmPPO*) had the most G+C bases. The highest content of alanine and leucine (16.9%) was found in its amino acid sequences (*SmPPO*), as well as the most casein kinase II phosphorylation sites (11). No signal peptide was found in *SmPPO*, but it existed a coiled coil. The *SmPPO* was a hydrophilic protein but no transmembrane domain. This work provided foundation to investigate the molecular mechanism how PPO regulated the accumulation of phenolics, especially for interaction between protein and protein, protein and gene.

Keywords: Tubiflorae; *Salvia miltiorrhiza*; polyphenol oxidase; sequence analysis; function prediction