

Исследование первичной структуры каталазы гриба *Penicillium vitale*.

5. Реконструкция полипептидной цепи и сравнение ее с другими каталазами

Э. А. Козлов, Т. Л. Левитина, М. Т. Бобровская, Л. В. Гудкова¹, Н. Ф. Радомский

Институт молекулярной биологии и генетики НАН Украины
252143, Киев, ул. Академика Заболотного, 150

¹ Институт биохимии им. А. В. Палладина НАН Украины
252030, Киев, ул. Леонтовича, 30

Из пептидов T6, T12, T29, T30, T32, T44, Tm2, Tm11, Tm28, Tm29, Tm35, Tm36, Tm39, Sp1, Sp2, Sp5, Sp11, Sp14, Sp26, Sp27, BrCN3—BrCN10, строение которых опубликовано ранее, реконструировано три фрагмента каталазы P. vitale. В результате сравнения их строения с первичными структурами других каталаз эти три фрагмента выстроены в непрерывную полипептидную цепь, содержащую 696 остатков аминокислот. Сопоставляются первичные структуры 26 каталаз из 23 организмов шести таксономических групп. Построена матрица степеней родства первичных структур сравниваемых каталаз.

Настоящее сообщение цикла статей, посвященных выяснению полной аминокислотной последовательности каталазы гриба *P. vitale*, включает описание реконструкции полипептидной цепи фрагментов каталазы, проведенной на основании данных построению пептидов, опубликованных в предыдущих статьях [1—4], а также сравнение первичной структуры и выяснение степени гомологии 26 каталаз различного происхождения.

Из набора пептидов [1—4] можно реконструировать фрагменты полипептидной цепи каталазы различными способами. Мы провели реконструкцию, отбирая наиболее крупные пептиды: T6, T12, T29, T30, T32, T44 [1], Tm11, Tm28, Tm29, Tm36, Tm39 [2], Tm2, Tm35 [4], Sp1, Sp2, Sp5, Sp11, Sp14, Sp26, Sp27 [3], BrCN3—BrCN10 [4]. Эти пептиды позволили реконструировать с оптимальным перекрытием три фрагмента полипептидной цепи каталазы: 1—301, 302—468, 469—696. В первом фрагменте только в двух местах (99—101 и 114—116) достигается перекрытие тремя остат-

ками, а в положении 19—20 — двумя. В остальных местах перекрытие достигается большим числом остатков. Во втором фрагменте в положении 456 пептиды перекрываются одним остатком, а в положениях 419—420 и 423—424 — двумя. В третьем фрагменте только в положении 672—673 получено перекрытие двумя остатками. Надежность реконструкции трех фрагментов достигается сравнением их аминокислотных последовательностей с первичными структурами других каталаз. При этом три реконструируемых фрагмента соединяются встык, образуя непрерывную полипептидную цепь, насчитывающую 696 остатков аминокислот. На рис. 1 приведена реконструированная таким способом полная аминокислотная последовательность каталазы *P. vitale*. В эту последовательность вписываются все исследованные нами ранее [1—4] пептиды. На N-конце каталазы определен остаток треонина, что указано стрелкой. Вертикальная стрелка в положении 395—396 указывает наиболее уязвимую связь Phe-Ala, которая, как мы предполагаем, расщепляется протеазами при выделении или хранении каталазы. Это предположение воз-

230 Arg
 Asn-Gly-Lys-Gln-Phe-Gln-Cys-Ser-Asn-His-Trp-Gln-Pro-Leu-Gln-Gly-Phe-

 250 Ser Val Lys
 Gly-Val-Glu-Glu-Ala-Trp-Arg-Phe-Pro-Glu-Glu-Gly-Glu-Gly-Tyr-Val-Ala-

 270 Glu
 Phe-Glu-Ser-Ile-Glu-Leu-Leu-Thr-Val-Gly-Asp-Glu-Glu-Leu-Glu-Ile-Gln-

 /-----
 Gln 290
 Phe-Asn-Asn-Asp-Leu-Arg-Glu-Arg-Phe-Asn-Ser-Ser-Glu-Val-Thr-Lys-Ser-

 /-----Tml1-----
 -----Sp11-----/
 310 Asp
 Arg-Leu-Val-Pro-Leu-Ile-Thr-Gln-Gly-Lys-Leu-Val-Phe-Asn-Lys-Asn-Ile-

 /-----
 330
 Phe-Asn-Glu-Val-Ile-Gly-Ala-Met-Phe-Gln-Pro-Gly-His-Ile-Val-Arg-Gly-

 -----BrCN9-----/
 -----/
 350 Ala Pro
 Thr-Glu-Asp-Pro-Leu-Leu-Gln-Gly-Arg-Leu-Phe-Ser-Tyr-Leu-Asp-Thr-Gln-

 Phe Arg 370 Arg
 His-Gly-Pro-Asn-Ile-Gln-Gln-Leu-Gly-Phe-Asn-Arg-Pro-Pro-Arg-Ala-Pro-

 390 ↓
 Asn-Asn-Arg-Asp-Gly-Ala-Gly-Glu-Met-Ile-Asp-Leu-Pro-Pro-Phe-Ala-Ser-

 /-----T6-----
 Tyr 410 Ser
 Thr-Gln-Glu-Trp-Gly-Ala-Lys-Asp-Ile-Lys-Gln-Thr-Ala-Val-Gly-Gln-Asn-

 /-----Sp14-----
 -----/
 430 Lys
 Gln-Glu-His-Arg-Phe-Ser-His-Trp-Lys-Phe-Gly-Val-Asn-Gly-Phe-Val-His-

 ---T44-----/

```

                                     450                                     460
Asp-Asp-Asn-Val-Thr-His-Ala-Arg-Gly-Phe-Phe-Thr-Ala-Pro-Glu-Arg-Gly-Gln-Gln-Lys-
-----/-----T30-----/
                                     470                                     480
Lys-Arg-Val-Ala-Ala-Phe-Asp-Arg-Met-Phe-Thr-Val-Val-Gly-Leu-Ser-Val-Asp-Gly-Gln-
-----Sp2-----/ /-----
      /-----Tm29-----/ /-----
      Asp          Arg      490                                     500
Gln-Ala-Asn-Ser-Asp-Gln-Tyr-Ala-Asp-Phe-Asp-Ala-Ala-Ala-Gly-Lys-Lys-Val-Ala-Lys-
-----
                                     510                                     520
Ala-Ile-Gly-Val-Glu-Ala-Pro-Lys-Pro-Asn-Ser-Asn-Tyr-Phe-His-Pro-Thr-Asp-Val-Phe-
-----
                                     530                                     540
Gly-Glu-His-Ile-Ala-Ala-Ser-Gly-Thr-Lys-Tyr-Gly-Val-Pro-Glu-Gly-Asn-Thr-Lys-Gly-
-----Tm2-----
-----BrCN7-----
                                     550                                     560
Val-Leu-Leu-Ala-Ser-Val-Asn-Lys-Pro-Ala-Ser-Ile-Ala-Gln-Gly-Ala-Lys-Leu-Gln-Val-
-----
                                     570                                     Pro 580
Val-Ala-Ser-Ser-Gly-Asp-Phe-Ala-Glu-Phe-Phe-Ile-Ser-Ala-Lys-Gln-Leu-Asn-Met-Arg-
-----/-----
                                     590                                     600
Glu-Val-Thr-Gln-Gly-Ile-Ile-Pro-Leu-Val-Pro-Val-Leu-Lys-Leu-Ala-Lys-Leu-Asp-Leu-
-----/-----
                                     610                                     620
Gly-Lys-Thr-Phe-Arg-Phe-Gln-Leu-Met-Gln-Val-Gly-Asn-Ile-Glu-Glu-Leu-Glu-Arg-Phe-
-----/ /-----
-----Sp27-----/
                                     630                                     640
Gly-Phe-Asp-Leu-Pro-Asp-Leu-Thr-Asp-Lys-Gln-Val-Leu-Asp-Ser-Ala-Met-Gly-Met-Phe-
-----BrCN8-----/-----
                                     /-----T32-----

```

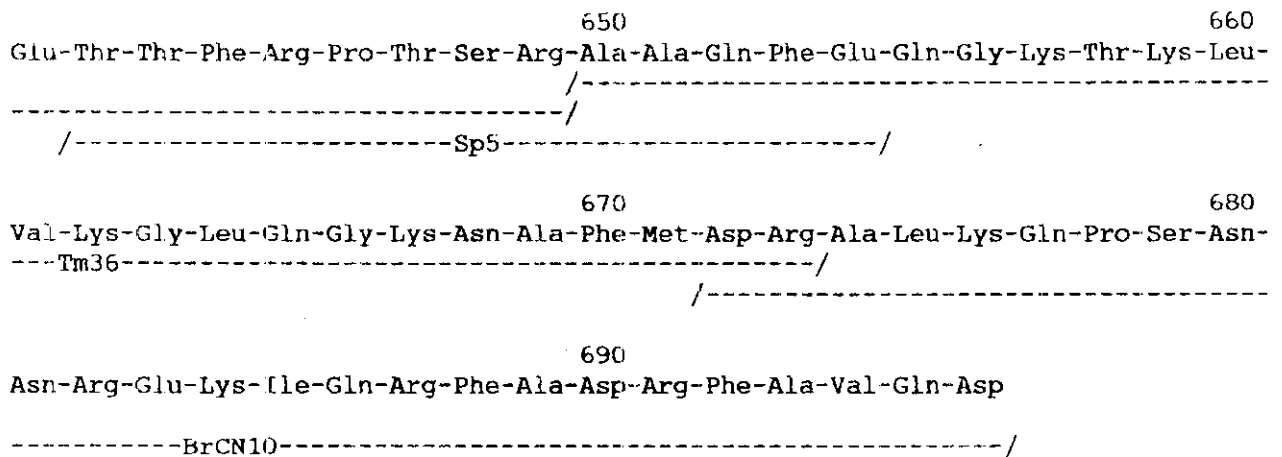


Рис. 1. Схема реконструкции полипептидной цепи каталазы *P. vitale*

никло в результате исследования строения пептидов, полученных расщеплением трипсином немодифицированной [1] и модифицированной [2] каталаз и стафилококковой протеиназой V8 [3]. Остатки аминокислот, записанные в 36 положениях над основной аминокислотной последовательностью, отражают микрогетерогенность исследованных препаратов каталазы, продуцируемой грибом в условиях промышленного производства фермента. Именно из культуральной среды промышленного производства выделяли и очищали для структурных исследований [5] образцы каталазы. Наличие микрогетерогенности было установлено при исследовании строения пептидов, полученных при расщеплении каталазы различными способами [1—3].

Выясненная нами аминокислотная последовательность каталазы *P. vitale* содержит на 26 остатков аминокислот больше, чем «наиболее вероятная» последовательность (670 остатков) по данным рентгеноструктурного анализа с разрешением 2 Å [6]. Из сравнения аминокислотной последовательности, представленной на рис. 1, с рентгеноструктурной последовательностью видно, что в последней пропущено 38 остатков аминокислот и вставлено 11. В 40 положениях рентгеноструктурной последовательности пары аминокислот располагаются в обратном порядке («парные инверсии»). С учетом таких «парных инверсий» 33 % аминокислотных остатков рентгеноструктурной последовательности идентифицированы правильно, в то время как авторы [6] предполагали 50 %. Довольно большое различие объясняется трудностью иденти-

фикации остатков при разрешении 2 Å (около 30 % остатков, обозначенных буквой А, неидентифицировано авторами [6]). По их мнению, при таком разрешении трудно отличить такие пары остатков, как Asp и Asn, Glu и Gln, Val и Thr, Leu и Asp, Phe и His и др. [6]. Условно эти пары можно назвать «гомологичными по объему». При сравнении рентгеноструктурной последовательности с последовательностью, выясненной нами, видно, что наиболее частые ошибки встречаются в следующих группах: Glu, Asp, Gln, Asn; Phe, His, Gln, Glu; Val, Thr, Ser; Leu, Asp, Asn; Leu, Ile, Val; Tyr, Trp, Phe; Arg, Lys. Очевидно, эти группы можно рассматривать как группы «гомологичных по объему» аминокислотных остатков при рентгеноструктурном анализе аминокислотной последовательности с разрешением 2 Å. Суммируя число правильно идентифицированных остатков (33 %) и «гомологичных», получим, действительно, как и предполагали авторы [6], 50 % сходства рентгеноструктурной аминокислотной последовательности каталазы *P. vitale* с «химической» аминокислотной последовательностью.

Сравнение первичных структур 26 каталаз из 23 организмов шести таксономических групп приведено на рис. 2. Выясненная нами первичная структура каталазы *P. vitale* занимает на рисунке нижнюю строку. Аминокислотные последовательности каталаз Mn, Rn, Hs, Bp, Dm, Ps, Gh, A1, Ib, Os, Le, Zm, Bs, Lsa, Lsc, Ec, Ct, Hp, Sca, Sct взяты из работы [7], каталаз Zm1, Zm2 и Zm3 — из работ [8, 9], каталаз Hi и Ma — из работ [10] и [11]

		10	20	30	40
Ec		MSQHNEKNPHQHQSPLHDSSEAKPGMDSLAPEDGSHRPAAEPTPPGAQ			
	50	60	70	80	90
Mn	MSDSRDPASDQMKQWKEQRASQRPDVLTGGGNP-IGDKLNIIMTAGSRGPLLVDVVFTEMAHFDRERI				
Rn	MADSRDPASDQMKQWKEQRAPQKPDVLTGGGNP-IGDKLNIIMTAGPRGPLLVDVVFTEMAHFDRERI				
Hs	MADSRDPASDQMKQWKEQRAAQKADVLTGGAGNP-VGDKLNVITVGRGPLLVDVVFTEMAHFDRERI				
Bp	ADNRDPASDQMKHWKEQRAAQKPDVLTGGGNP-VGDKLNSLTVGPRGPLLVDVVFTEMAHFDRERI				
Dm	MA-GRDAASNQLIDYKNSQTV-SPGAITTGNGAP-IGIKDASQTVGPRGPILLQDVNFDLDEMSHFDRERI				
Ps	MDPYKH-RPSSAFNSP-FWTTNSGAP-VWNNNSLTVGSRGPILLEDYHLVEKLAQFDRERI				
Gh	MDPYKH-RPSSAFNSP-FWTTNSGAP-VWNNNSLTVGSRGPILLEDYHLVEKLANFDRERI				
At	MDPYKY-RPASSYNP-FWTTNSGAP-VWNNNSMTVGRGPIILLEDYHLVEKLANFDRERI				
Io	MDPSKY-RPSSSFNTP-FCTTNSGAP-VWNTCALTVGSRGPILLEDYHLVEKIQNFTRERI				
Os	MDPCKF-RPSSSFDTK-TTNTNAGAP-VWNDNEALTVGPRGPILLEDYHLIEKVAHFARERI				
Le	MDPSKY-RPSSAYDTP-FLTTNAGGP-VYNNVSSLTVGPRGPILLEDYHLIEKLAFFDRERI				
Zm	MDPYKH-RPSSAFNAP-YWTTNSGAP-VWNNSSLTVGARGPILLEDYHC-EKLAFFDRERI				
Zm1	MDPYKH-RPSSGSNS-SFWTTNSGAP-VWNNNSALTVGQRGPILLEDYHLIEKLAQFDRERI				
Zm2	VRRSGSSSSLATAMDYPKHAR-RAPSTPV-LDH-NS-APRVEQRQL-LTVGARGPILLEDL-VVEKLANFDRERI				
Zm3	MTMDPTKF-RPSSSHDTP-VTNTNAGAP-VWNDNEALTVGPRGPILLEDYHLIEKVAHFDRERI				
Bs	MSSNK-LTTSWGAP-VGDNQNSMTAGSRGPTLIQDVHLLKLAHFNRERV				
Lsa	MTNQ--LTTNEGQP-WADNQHSQTAANAAPSLIQDYQLLEKLAHFNRERI				
Lse	MTDRRLTTNQGV-IGDNQNSMTAGLKGPTLLEDYVLLIEKLAHFDRERV				
Hi	MSSQCPFSLAATNLTMGNGAP-VADNQNSLTVGPRGPILLAQDLWLNEKLAFFVREVI				
Ma	MATDHTSDAPDPKQRDLES-ARFRDTG-Y-LTTQGV-VDHTDDALTVGERGPTLLEDHFHAREKI THFDHERI				
Ec	PTAPGSLKAPDTRNEKLNLEDVRKGSSENYALTINQGV-IADDQNSLRAGSRGPTLLEDLREKI THFDHERI				
Ct	MAP-TFTNSNGQP-IPFPATQRVGQHGPLLQDFNLIDSLAHFDRERI				
Hp	MSNPPVFTTSQGCPSDFPTT-QRIP-L-DSTGYKYAPPIGPLLQDFKLIIDSLAHFDRERI				
ScA	MSKLGQEKNEVNSDVRVVTNSTGNP-INEPFTQRI GEHGPLLQDYNLIDSLAHFNRENI				
ScT	MNVFGKKEEKQEKVYSLQNGFPYSHHPYASQYSRDPGPIILQDFHLLNIASFDRERV				
Pv	TASGKLQRKFLDRFAISMGRGVA-LGKTYGTLGAASRGATLLQDLLETFEITFAFDRERV				

	120	*	130	140	150	160	*170	180	190
Mn	PEFVVHAKGAGAFGYFEVTH----DITRYSKAKVFEHIGKRTPIAVRFSTVAGESGS----ADTVRDPRGFAVKF								
Rn	PEFVVHAKGAGAFGYFEVTH----DITRYSKAKVFEHIGKRTPIAVRFSTVAGESGS----ADTVRDPRGFAVKF								
Hs	PEFVVHAKGAGAFGYFEVTH----DITRYSKAKVFEHIGKRTPIAVRFSTVAGESGS----ADTVRDPRGFAVKF								
Bp	PEFVVHAKGAGAFGYFEVTH----DITRYSKAKVFEHIGKRTPIAVRFSTVAGESGS----ADTVRDPRGFAVKF								
Dm	PEFVVHAKGAGAFGYFEVTH----DITRYCAAKIFDKVKKRTPLAVRFSTVGGESGS----ADTARDPRGFVKF								
Ps	PEFVVHARGASAKGFFEVTN----DISLTLTCAFLRAPGVQTPVIVRFSTVIHERGS----PETLRDPRGFVKF								
Gh	PEFVVHARGASAKGFFEVTN----DISLTLTCAFLRAPGVQTPVIVRFSTVIHERGS----PETLRDPRGFVKF								
At	PEFVVHARGASAKGFFEVTN----DISLTLTCAFLRAPGVQTPVIVRFSTVIHERGS----PETLRDPRGFVKF								
Io	PEFVVHARGASAKGFFEVTN----DITHTLTCADFLRAPGVQTPVIVRFSTVIHERGS----PETLRDPRGFVKM								
Os	PERVVHARGASAKGFFEVTN----DVTDLTCAFLRSPGAQTPVIVRFSTVIHERGS----QETLRDPRGFVKF								
Le	PERVVHARGASAKGFFEVTN----DISLTLTCAFLRAPGAQTPVIVRFSTVIHERGS----PESIRDIRGFVKF								
Zm	PERVVHARGASAKGFFEVTN----DITHTLTCADFLRAPGVQTPVIVRFSTVIHERGS----PETLRDPRGFVKF								
Zm1	PERVVHARGASAKGFFEVTN----DVSHLTLTCAFLRAPGVQTPVIVRFSTVIHERGS----PETLRDPRGFVKF								
Zm2	PERVVHRCGASAKGFFEVTN----DITTDVADSAGRTAGMQAVIAVSLNCVALALRYHDGSPETL--PRVASV--								
Zm3	PERVVHARGASAKGFFESTH----DVTSLTLTCAFLRAPGVRTPVIVRFSTVIHERGS----GRTLRDARGFAVKF								
Bs	PERVVHAKGAGAHGYFEVTN----DVTKYTKAAFLSEVGKRTPLFIRFSTVAGELGS----ADTVRDPRGFAVKF								
Lsa	PERVVHAKGAGRKGYFKVTK----DMSAYTKAAVFSGVGKKTPLITRFSTVQVAGEAGY----PDTVRDVRGFVKF								
Lse	PERVVHARGAGAHGKF-VTKKSMK---KYTKAQLQEEGTETEVFARFSTVIHQHS----PETLRDPRGFVKF								
Hi	PERRMHAKGSGAFGTFTVTH----DITKYTRAKIFSEVGKKTETEMFARFTTVAGERGA----ADAERDIRGFALKF								
Ma	PERVVHARGAGAYGYFEPYDDRLAQYT---AAQFLTSPGTRTPVVFVRFSTVAGSRGS----ADTVRDPRGFAVKF								
Ec	PERIVHARGASAAHGYFQPYK-SLSDIT---KADFLSDPNKITPVVFVRFSTVQGGAGS----ADTVRDIRGFATKF								
Ct	PERVVHAKGSGAYGVFEVTD----DITDVCAAKFLDTVGGKTRIFTRFSTVGGELGS----ADTARDPRGFATKF								
Hp	PERVVHAKGAGAYGVFEVTD----DITDVCSAKFLDTVGGKTRIFTRFSTVGGEEKGS----ADTARDPRGFATKF								
ScA	PQRNPHAHGSGAFGYFEVTD----DITDICGSAMFSKIGKRTKCLTRFSTVGGDKGS----ADTVRDPRGFAVKF								
ScT	PERVVHAKGGCRLEFELTD-SLSDIT-Y--AAPYQNVGYKCPGLVRFSTVGGESGT----PDTVRDPRGVSKF								
Pv	PERAVHARGTAGHGTFLSYE----DWSNLTAASFLSAEGKFTPEMTRFSTVSGARGS----ADTARDVHGFAVRF								

	200	* 210	220	230	240	250
Mn	YTEDGNWDLVGNNTPIFFIRDAILFPSF--IHSQKRNPQTHL-K-----DPDMVWDFWSLRPES					
Rn	YTEDGNWDLVGNNTPIFFIRDAMLFPSF--IHSQKRNPQTHL-K-----DPDMVWDFWSLCPES					
Hs	YTEDGNWDLVGNNTPIFFIRDPIILFPSF--IHSQKRNPQTHL-K-----DPDMVWDFWSLRPES					
Bp	YTEDGNWDLVGNNTPIFFIRDALIFPSF--IHSQKRNPQTHL-K-----DPDMVWDFWSLRPES					
Om	YTEDGVWDLVGNNTPVFFIRDPIILFPSF--IHTQKRNPQTHL-K-----DPDMFWDFTLRLPES					
Ps	YTREGNYDLVGNFPVFFVHDGMNFPD-M-VHALKPNPQTHI-Q-----ENWRILDFFYNFPES					
Gh	YTREGNFDLVGNFPVFFIRDGMKFPD-M-VHALKPNPKSHI-Q-----ENWRILDFFSHHPES					
At	YTREGNFDLVGNFPVFFIRDGMKFPDI--VHALKPNPKSHI-Q-----ENWRILDFFSHHPES					
Ib	YTRGCNWDLVGNFPVFFIRDGTQFPDV--IHAFKPNPKSHI-Q-----ENWRILDYLSHLPES					
Os	YTREGNWDLVGNFPVFFIRDGIKFPDV--IHAFKPNPRSHV-Q-----EYWRVDFLSSHHPES					
Le	YTREGNFDLVGNFPVFFNRDAKSF PDT--IRALKPNPKSHI-Q-----ENWRILDFFSFLPES					
Zm	YTREGDWDLVGNFPVFFIRDGIKFPD-M-VHALKPNPRTHI-Q-----DNWRILDFFSHHPES					
Zm1	YTREGNFDLVGNMFPVFFIRDGMKFPD-M-VHARKPNPKTNL-Q-----ENWRIVDFFSHHPES					
Zm2	YT-GGNWT-V-ETLPSL-IRDGIK--SGH-VHALKPNPRTHI-Q-----DNWRILDFFSHHPES					
Zm3	YTREGNWDLVGNFPVFFIRDGIKFPDV--IHAFKPNPRSHV-Q-----EYWRVDFLSSHLPES					
Bs	YTEEGNYDLVGNTPVFFIRDAIKFPDF--IHTQKRDPKTHL-K-----NPTAVWDFWLSLSPES					
Lsa	YTEEGNYDLVGNTPVFFVNDPLKFPDF--IHSQKRDPRTHA-R-----SQDMQWDFWLSLSPES					
Lse	YTEEGNYDLVGNLVPVFFIRDAIKFPDV--IHSQKRDPRTNI-Q-----DGNRYWDFWLSLTPES					
Hi	YTEEGNWDLVGNTPVFFLRDPRKFPDL--NKAVKROPRTNM-R--SATNNDWDFWT-L-L-PEA					
Ma	YTEQGNVYDLVGNFPVFFIQDGIKFPVL--VHAVKPEPHNEIPQAQSAHDT--LWDFVSLQPET					
Ec	YTEEGIFDLVGNTPVFFIQDAHKFPDF--VHAVKPEPHWAIPQQAQSAHDT--FDWYVSLQPET					
Ct	YTEEGNLDLVYNNTPVFFIRDPSKFPHF--IHTQKRNPETHL-K-----DANMFWDYLTNEES					
Hp	YTEEGNLDLVYNNTPVFFIRDPIKFPHF--IHTQKRNPATNL-K-----DPNMFWDYLTANDES					
ScA	YTEEGNLDWVYNNTPVFFIRDPSKFPHF--IHTQKRNPQTNL-R-----DADMFWDFLTP-PEN					
ScT	YYEWGNHDWVFNTPVFFLRDAIKFPVF--IHSQKRDPQSHLNQFQ---DTTIYWDYLTLPNDES					
Pv	YVDEGNFDLVGNTPVFFIWDVIEPTLMALHAQKPNPRFHLPRGQQ--DPNRISDNLTARGDS					

	270	280	o 290	300	310	320	330
Mm	SDRGI PDGHRHMNGYGSHTFKLVNADGEAVYCKFHYKTDQGIKNLPVGEAGRLAQEDPDYGLRD						
Rn	SDRGI PDGHRHMNGYGSHTFKLVNANGEAVYCKFHYKTDQGIKNLPVEEAGRLAQEDPDYGLRD						
Hs	SDRGI PDGHRHMNGYGSHTFKLVNANGEAVYCKFHYKTDQGIKNLSVEDAARI SQEDPDYGLRD						
Bp	SDRGI PDGHRHMNGYGSHTFKLVNADGEAVYCKFHYKTDQGIKNLSVEDAARLAHEDPDYGLRD						
Dm	SDRGTDPDGYCHMNGYGSHTFKLINAKGEPYAKHFHKTDQGIKNLDVKTADQLASTDPDYSIRD						
Ps	DDVGVPPQDYRHMDGFGVNTYTLINKAGKSVYVKFHWKPTCGVKCLLEEEAIQVGGSNHSHATKD						
Gh	DDLGYPPQDYRHMEGSGVNTYTLINKAGKAHYVKFHWKPTCGVKCLLEDEAIKVGGANHSHATQD						
At	DDIGIPQDYRHMDGSGVNTYMLINKAGKAHYVKFHWKPTCGVKSLLLEDAIRLGGTNHSHATQD						
Ib	DDVGIPTDYRHMEGFGVHTFTMINKEGKANYVKFHWKPTCGVKCLLEEEAIRIGGENHSHATQD						
Os	DDVGIPTDYRHMDGFGVNTYTFVTRDAKARYVKFHWKPTCGVSLMDDEATLVGGKNHSHATQD						
Le	DDVGLPTDYRHMEGFGVHAYQLINKEGKAHYVKFHWKPTCGVKCMSEEEAIRVGGTNHSHATKD						
Zm	DDVGI PADYRHMDGSGVHTYTLVSRAGTVTYVKFHWKPTCGVRSLLMDEAVR-CGANHSHATKD						
Zm1	DDVGI PLNYRHMEGFGVNTYSLINRDGKPHLVKFWKPTCGVKCLLDNEAVTVGGTCHSHATKD						
Zm2	DDSGI PADVVRPHGGSGVHTYTLVSRAGTVTYVKFHWKPTCGVRSLLMDEAVRVAQPPQ-PRTKD						
Zm3	DDVGVPSDYRHMEGFGVNTYTFVSAAGKAQYVKFHWKPTCGDRCILTDEEARVGGRNHSH-TQD						
Bs	SDRGI PATLRHMHGFGSHTFKWNAEPEGVWIKWHFKTEQGVKNLDVNTAAK IAGENPDYHTED						
Lsa	SDRGI PASYRMMHGFGSHTFKWNAQGEQFVWIFHFKTNGIHNLSNELADELAGKDTDYLDQND						
Lse	SDEGTPASYREIRGSSVHAFKWINEEGKTVYVKLRWVPKAGIVNLSTDAQAAQIQAKEFNHASRD						
Hi	SDRGI PASYRHMHGFGSHTYSFWNEAGERFVVKFHFRTQQGIKNLTDAAEAEI IANDRESHQRD						
Ma	SDRALPRSYRMMQGFVHTFRLVNARGRGTFFVKFHWKPRLGVHSLIWDECCQKIAGKDPDYNRRD						
Ec	SDRGI PRSYRTMEGFGIHTFRLINAEGKATFVRFHWKPLAGKASLVWDEAQKLTGRDPDFHRE						
Ct	SDRGT PASYREMNYSGHTYKWSNNKGEWYVQVHFISDQGIKTLTNEEAGSLAGSNPDYAQED						
Hp	SNRGT PASYRTMNGYSGHTYKWSNKGWYVQVHF IANQGVHNLDEEAGRLAGEDPDHSHTRD						
ScA	SDRGT PANYSRSMHGYSGHTYKWSNKNGDWHYVQVHIRTQGIKNLTIIEATKIAGSNPDYCCQD						
ScT	GDRGT PASWASMNAYS GHSFKMVNKEGKDTYVQFHVLSDTGFETLTGDKAAELSGSHPDYNQAK						
Pv	SERGS PKAYSNTPEPNKHSFRLVTDNGKQFQCSNHWQPLQGFIDLGVEEAWRFPEEGEGYVAEN						

	350	360	370	380	390	400	410
Mm	PSWTFYIQVMTFKEAETFPF	-NPFDLTKVWPHKDYPLI	PVVKLVLNKNPVNYFAEVEQMAFDPSNM	-PPGIEPSP			
Rn	PSWTFYIQVMTFKEAETFPF	-NPFDLTKVWPHKDYPLI	PVVKLVLNRRNPANYFAEVEQMAFDPSNM	-PPGIEPSP			
Hs	PSWTFYIQVMTFNQAEETFPF	-NPFDLTKVWPHKDYPLI	PVVKLVLNRRNPVNYFAEVEQIAFDPSNM	-PPGIEASP			
Bp	PSWTLYIQVMTFSEAEIFPF	-NPFDLTKVWPHGDYPLI	PVVKLVLNRRNPVNYFAEVEQLAFDPSNM	-PPGIEPSP			
Dm	PSWTFYIQVMTYEQAQKFKY	-NPFVTKVWSQKEYPLI	PVVKMLDRNPKNYFAEVEQIAFSPAHL	-VPGVEPSP			
Ps	PEWKLYIQTIIDPAHEDRFDF	-DPLDVTKTWPEDI	ILPLQPVGRMVLNKNIDNFFAENEQLAFCPAIV	-LPGIYYSD			
Gh	PEWKLYIQTIIDPAHEDRFDF	-DPLDVTKTWPEDI	ILPLQPVGRMVLNKNIDNFFAENEQLAFCPAII	-VPGIHYSD			
At	PEWKLYIQVMDPDHEDRFDF	-DPLDVTKTWPEDI	ILPLQPVGRMVLNKNIDNFFAENEMLAMDPANI	-VPGIYFSD			
Ib	PEWKLFVQVIDPEEEERFDF	-DPLDDTKTWPEDEV	PLRPVGRMVLNRRVDNFFNENEQLAFGPGLV	-VPGIYYSD			
Os	PEWKLYIQVMDPEEDVDFDF	-DPLDVTKTWPEDI	ILPLIPVGRMVLNRRIDNFFAENEQLAFNPGHI	-VPGIYYSE			
Le	PEWTLYIQVMDPEMEDPLDLDPLDVTKTWPEDR	FLPLQPVGRMVLNRRIDNFFAENEQLAFCPGLI	-VPGIYYSD				
Zm1	PEWKLYIQVMDPEMEDPLDLDPLDVTKTWPEDI	ILPLQPVGRMVLNKNVDNFFAENEQIAFCAIS	-VPGIHYSD				
Zm2	PEWTLYIQVMDPEMEDRQLDPLDVTKTWPEDV	-PLEPVGRMVLNRRIDNFFAENEQLAFSPGLI	-VPGIYYSD				
Zm3	PEWTLYVQVMDPAQEEQYDF	-DPLDDTKTWPEDE	ILPLRPVGRMVLDRNVDNFFNENEQLAFGPGLV	-VPGIYYSD			
Bs	PAWKLYVQIMPLEDANTYRF	-DPFDVTKVWSQKEYPLI	IEVGRMVLDRNPNENYFAEVEQIAFSPGTL	-VPGIDVSP			
Lsa	PSWTVAVQLVLMKMAELSPK	-DIFDVTKVI	SQKDYPLIEIGQMVLDENPTNFFEDIQELAFSPANL	-VPGTEASP			
Lse	PEWDLVYVQLDPKDLNDF	-NPLDATKDFEDV	FPYEHVGTMTLNRNPDNIFAETESVGFNPGVL	-VPGMLPSE			
Hi	PKWTLFVQIMFEADAQKVPY	-HPFDLTKVWSKDYPLI	IEVGEFELNRNPNENFFADVEQSAFAPSNL	-VPGIGASP			
Ma	PEWELGVQLVAEDDEFDF	-DLLDATKI	IPPEQVPLPVGRMVLNRRPDNFFAETEQAFAHTANV	-VPGIDFTN			
Ec	PEYELGFQLPEEDEFKDF	-DLLDPTKLI	PEELVPVQRVGRMVLNRRPDNFFAENEQAFAHPGHI	-VPGLDFTN			
Ct	PSWTCYIQVMTFAQAKEAEF	-SVFDLTKVWPHGKYPMRRF	GKFTLNENPKNYFAEVEQAFAFSPAHT	-VPHMEPSA			
Hp	PSWECYIQVMTLEQSKKLPF	-SVFDLTKVWPHKDYPLI	RHFGRFTLNENPKNYFAETEQAFAFSPAHT	-VPGMEPSN			
ScA	PSWTVYIQVMTTERDARKL	PF-SVFDLTKVWPHGKDYPLI	RRVGRMVLNRRPDNFFAQAQVEQAFAFSPST	-VPYQEASA			
ScT	PKFNVCYVQVMTPEQATKFRY	-SVNDLTKI	WPHKEFPLRKFGTITLTENVVDNYFQEI	EQVAFSPITNTCIPGIKPSN			
Pv	GDDELEIQSMSPFNDLRERF	-NSSEVTKSSVVR	VPLITQGKLVFNKNIQMIFNEVIGAMFQPGHT	-VRGVDFTE			

	420	430	440	450	460	470	480	490
Mm	DKMLQGRLFAYPDTHRRLG	-PNYLQIPVNC	PYR-----			ARVANYQRDGP	MCM-HDNQGGAPNY	
Rn	DKMLQGRLFAYPDTHRRLG	-PNYLQIPVNC	PYR-----			ARVANYQRDGP	MCM-HDNQGGAPNY	
Hs	DKMLQGRLFAYPDTHRRLG	-PNYLHIVNC	PYR-----			ARVANYQRDGP	MCMQ-DNQGGAPNY	
Bp	DKMLQGRLFAYPDTHRRLG	-PNYLQIPVNC	PYR-----			ARVANYQRDGP	MCM-MDNQGGAPNY	
Dm	DKMLHGRLFSYSDTHRRLG	-PNYLQIPVNC	PK-----			VKIENFQRDGM	NVT-DNQGGAPNY	
Ps	DKMLQTRVFSYADSQRRLG	-PNYLQLPVN	APK-----			WSHHNNHHEG	FMNAIHRDEE--VNY	
Gh	DKLLQTRIFSYSDTQRRLG	-PNYLQLPAN	APK-----			CAHHNNHHEG	FMNFMHRDEE--INNY	
At	DKLLQTRVFSYADTQRRLG	-PNYLQLPVN	APK-----			CAHHNNHHEG	FMNFMHRDEE--VNY	
Ib	DKMLQARVFAYADTHRRLG	-PNYMLLPVN	APK-----			CAHHNNSYDG	MYMNFVHRDEE--VDY	
Os	DKMLQCRVFAYADTQRYRLG	-PNYMLLPVN	APK-----			CAHHNNHYDG	AMNFMHRDEE--VDY	
Le	DKLLQTRIFAYADTQRHRIG	-PNYMQLPVN	APK-----			CGHHNNHHRD	GAMNMTHRDEE--VDY	
Zm	DKLLQTRFFSYSDTQRRLG	-PNYLMLPAN	APK-----			CAHHNNHYDG	SMNFMHRHEE--VDY	
Zm1	DKLLQTRIFSYADTQRRLG	-PNYMLLPVN	APK-----			CAHHNNHHDG	FMNFMHRDEE--VNY	
Zm2	DKLLQTRIFSYSDTQRRLG	-PNYLMLPAN	APK-----			CAHHNNHYDG	SMNFMHRTEE--VDY	
Zm3	DKMLQCRVFAYADTQRYRLG	-PNYMLLPVN	APR-----			CGTHNNHYDG	AMNLMHRDEE--VDY	
Bs	DKMLQGRLFAYHDAHRYRVG	-ANHQALP	INRARN-----			KVNNYQRDGM	RFD-DNGGGSVYY	
Lsa	DKLLQGRLFYKDAERYRLG	-ANYEQLPVN	RPK-----			VPVHNYERD	GAMAQN--QATG-VNY	
Lse	DRLLQGRLFYSYSDTQRHRVG	-PNYLQLP	INSPKT-----			PVDNNQRDGM	PFK--QOTSSINNY	
Hi	DRMLQARLFHYADAQRYRLG	-VCYRQIPVN	RPR-----			CPVHSNQRDGM	GRVD-GTYGSLPHY	
Ma	DPLLQFRNFSYLDTQLIRLGGPNFAQLPVNRP	-----				VAQVRTNQHD	GYAQHAIPQGR-S-SY	
Ec	DPLLQGRLFYSYTDQISRLGGPNFHEIPINRPT	-----				CPYHNFQRDGM	HMRGIDTNP-A-NY	
Ct	DPVLQSRLFYSYADTHRRLG	-TNYTQIPVN	CPVT-----			GAVFNPHMRD	GAMNVN-GNLGNHPNY	
Hp	DPVLQSRLFYSYADTHRRLG	-PNYHQIPVN	CPLK-----			SGSFNP	INRDGPMCVD-GNLGGTPNY	
ScA	DPVLQARLFYSYADHRYRLG	-PNFHQIPVN	CPIA-----			SKFFNPA	INRDGPMNVN-GYGFSEPTY	
ScT	DSVLQARLFYSYADTQRRLG	-ANYQQLPVN	RPRNLGCPYSGDSQYTAEQCPFKAVN	QRDGPMSYVNF			GFPEPNY	
Pv	DPLLQGRLFYSYLDTQLNRHG	-PNIQQLG	FNRP-----			APIHNNNRD	GAGEM-IDLPPFAFV	

	500	510	520	530	540	550
Mm	YPNSFSAPEQQRS	SALEHSV-----	QC--	AVDV-----		
Rn	YPNSFSAPEQQG	SALE-----	HHSQC--	SADV-----		
Hs	YPNSFGAPEQQP	SALEHSI-----	QYSGEV-----			
Bp	YPNSFSAPEHQPS	AL-----	EHR-THFSGDV-----			
Dm	FPNSFNGPEQCP	RRARALSS-----	CCPVTGDV-----			
Ps	FP-SRHDTVRHA	ERVPIPI-----	TTHLSA-----			
Gh	FP-SRYDQVRHA	EMFPIPI-----	PAVCTG-----			
At	FP-SRYDQVRHA	EKYPTPI-----	PAVCSG-----			
Ib	FP-SKFDNTRNA	ERFPTPI-----	LRIVTG-----			
Os	YP-SRHAPLRHA	PPYPTI-----	PRPVVG-----			
Le	LP-SRFDPCRP	AEQYPIPI-----	SCVLNG-----			
Zm	FP-SRYDAVRNA	PRYPIPI-----	TAHIAG-----			
Zm1	FP-SRFDPARHA	EKVPIPI-----	PRVLTG-----			
Zm2	FP-SRYDAVRNA	PRYPIPI--TAH-----	IAG-----			
Zm3	YP-SRHAR-RG	GRAAPTPLP-----	PRPVAG-----			
Bs	EPNSFGG--PK	ESPED-KQAAYPVQGI--	AD-----			
Lsa	EPNSQDG--P	TEVP-----	AAK----I-HGD-----			QI
Lse	EPNSYDTE-P	KENPAYIE----PEQEI-RGD-----				
Hi	EPNSFS-QWQQ	QPDFAEPPLRISGDAAHWDY-----				
Ma	FKNSIGGG---	CPALADEDVFRHYTQRVDGQT-----				
Ec	EPNSINDINWP	RETTPGPKRGGFESYQERVEGNK-----				
Ct	LA-SDKPIEFK	QFS--L----QEDQEVWHGAAT-----				
Hp	ANAYNCP IQY	AV-SPKASGNKP--DEKYTGVEVV-----				
ScA	LAN-DKSYTYI	QQDRPI----QQHQEVWNGPAI-----				
ScT	IS-S-LPMQTL	KFKNEVNDEVS--D-KFKGIVLDEVTEVSVRKQEQDQIRNEHIVDAKI-----				
Pv	ETQEWGAKDI	KQTAVGQNKFDQEHREFSHWKFGV-----				

	570	580	590	600	610	620
Mm	V----RTFYTKVL	NEEERKR-LCENIAGHLKDA-----	QLFIQKKAVKKNFTDVHPDYG-----			
Rn	V----RTFYTKVL	NEEERKR-LCENIANHLKDA-----	QLFIQKKAVKKNFTDVHPDYG-----			
Hs	V----RALFYVNL	NEEQRKR-LCENIAGHLKDA-----	QIFIQKKAVKKNFTDVHPDYG-----			
Bp	V----RTFYTKVL	NEEQRKR-LCENIAGHLKDA-----	QLFIQKKAVKKNFTDVHPDYG-----			
Dm	V----TDI'FWHVL	DKCAKKR-LVQNIAGHLSNA-----	SQ-FIQERAVKKNFTQVHADFG-----			
Ps	AGERYRTWAP---	DRQER--FLRRWVEA-LSD--TDPRI	THE-IRSIWVSYSWSQADRSVG-----			
Gh	PGERYRSWAA---	DRQER--FICRWVDA-LSD--PRV	THE-IRSIWVSYSWSQADKSLG-----			
At	PGERYRTF'FP---	ERQER--FIQRWIDA-LSD--PRI	THE-IRSIWVSYSWSQADKSLG-----			
Ib	PGDRYRSWAP---	DRQDR--FINRWVKA-LSE--PRV	THE-IRSTWISYLTQADRSLG-----			
Os	PGERYRSWAP---	DRQER--FLAVRRR--VAI--PKV	SPE-LRAIWNVYLSQCDESIG-----			
Le	AGERYRSWEP---	DRQDR--YINKWVES-LSD--PRV	THE-IRSIWISYLSQADKSCG-----			
Zm	PGERYRAMDP---	ARQER--FITRWVDA-LSD--PRL	THE-IRTIWLSNWSQADRSVG-----			
Zm1	AGERYRSFDP---	ARQDR--FIQRWVDA-LTD--PRV	THE-HRTIWIYSWSQCDAAALG-----			
Zm2	PGERYRAMDP---	ARQER--FIT-N-GSTLEI--P-G--	Y-RAR-IRNHGHSFC-LC-----			
Zm3	PGERYRSWDA---	DRQDR--FVKAIRR--LA--RTP	KQPE-LRSIWIDLLAKVDASLG-----			
Bs	A----GDL'YRL	MSEDERTR--LVENIVNAMKPV-----	EKEEIKLRQIEHFYKADPEYG-----			
Lsa	A----GKLYRLL	SADQTR--LIENIRMNLGQV-----	TKPEIQIREVKQFYQADPEYG-----			
Lse	A----KEVWKRY	SDAERAA--LVKNI'VDDWEGV-----	REDIKIRNLRNFYQVEPEFA-----			
Hi	P----RALFNLM	-NAEQKQS-LFNNTAAAMGDA-----	PDFIKYRHIRNCHWCDAAYG-----			
Ma	A----RMPFKSM	SFVEAEHI-VAAF'AFELGKVE-----	MPEIRSAVVAQLARVDDQLA-----			
Ec	P----RL'FWSQ	TPFEQRHI-VDGFSFELSKVV-----	RPYIRERVVDDQLAHDITLA-----			
Ct	A----TELWKVL	KKYPNQQEHLAHNVAVHASAA-----	DAPIQDRVIAYFTKVHEDLG-----			
Hp	P----KMPWKVL	GRTPGEQESLVKNVANHVSA-----	DEFIQDRVYEFYFKAEP'IG-----			
ScA	A----RNL'YRVL	GKQPGQKKNLAYNIGIHVEGA-----	CPQIQORVYDMFARVVK--G-----			
ScT	P----RAL'YEK	VYNDEQKK-LFVHNVVCHACKI-----	KDPKVKKRVTYQYFGLLNEDLG-----			
Pv	A----RGP'ETA	PERGQQKKR-VAAFDRMFTVVG-----	LSVDGQQANSDQYADFDAAG-----			

	650	660	670	680	690	700	710	
Mm	KNAIHTYQAGSHMAAKGKANL							
Rn	KNAIHTYVQAGSHIAAKGKANL							
Hs	KNAIHTFVQSGSHLAAREKANL							
Bp	KN							
Dm	SKF							
Zm2	EA							
Lse	IDLK							
Hi	LKARDTDPALGGGLL							
Ma	VDESAPVS-----PALSQVTDG-GDTIASRRIAVLAADG-VDVVGTRQRFTELMEQRC--AV-VEVLAPVAGGT							
Ec	LNETPPPDVNGLKKDPSSLYAI PDGDVKG-FVVAILLNDE-VRSADLLAILKALKAKG---VHAKLLYSRMGE-							
Ct	PRK							
Hp	-RKA---SSPSKI							
ScA	AKHASELSSNSKF							
ScT	LEGYAKTWSIASAN							
Pv	SNVTHPTDVFGEHIAASGTKYGVPEGNTKG-VLLASVKNKPASTAQQAKLQVVASSGDFAEFFISAKQL--NMRE-							
	720	730	740	750	760	770	780	790
Ma	LAGGSGGELRVDRSFTTMAASVLYDAVVVACGPRSVSTLSDDGYAVHFVTEAYKHLKPIGAYGAGVDLLRKAGIDN							
Ec	VTADDGTVLPIAATFAGAPSLTVDAVIVPCG--NIADIADNGDANYYLMEAYKHLKPIALAGDA----RKFKATI							
Pv	VTQGIIPLVVVLKLAKLDLGKTFRFQLMQVG--NIEELERFGFDLPDLTDKQVLD SAMGMFETT----FRPTSRA							
	800	810	820	830	840			
Ma	RLAEDTDVINDQAVVTTKAAADELPERFAEEFAAALAQHRCW---QRRTD--AVPA							
Ec	KIADQGE----EGIVE----ADSADGSFMDLTLMAAHRVWSRIPK-IDK--IPA							
Pv	AQFEQGK----TKLVK----GLQGNFMDRALKQPSNNRE--KIQRFADRFVQD							

Рис. 2. Сравнение первичных структур каталаз. Принятые сокращения: Mm — *Mus musculus*; Rn — *Rattus norvegicus*; Hs — *Homo sapiens*; Bp — *Bos primigenius*; Dm — *Drosophila melanogaster*; Ps — *Pisum sativum*; Gh — *Gossypium hirsutum*; At — *Arabidopsis thaliana*; Ib — *Ipomoea batatas*; Os — *Oriza sativa*; Le — *Lycopersicon esculentum*; Zm — *Zea mays*; Bs — *Bacillus subtilis*; Lsa — *Lactobacillus sake*; Lse — *Listeria seeligeri*; Hi — *Haemophilus influenzae*; Ma — *Micobacterium avium*; Ec — *Escherichia coli*; Ct — *Candida tropicalis*; Hp — *Hansenula polymorpha*; Sc — *Saccharomyces cerevisiae*; Pv — *Penicillium vitale*

соответственно. Аминокислотные последовательности всех каталаз выведены из нуклеотидных последовательностей, кроме каталазы Bp, первичная структура которой выяснена, как и у Pv, химическими методами [12]. Как видно из рис. 2, сравниваемые каталазы можно разделить на две группы: длинные каталазы Pv, Ma и Ec, содержащие 696, 706 и 753 остатка аминокислот соответственно, и короткие, сдержанные около 500 остатков. Общая особенность длинных каталаз состоит в том, что их С-концевая последовательность на 160--200 остатков длиннее, чем коротких. Сравнительные рентгеноструктурные исследования [6] каталаз Pv и Bp показали, что каталаза Pv содержит дополнительный С-концевой домен, по пространственной структуре подобный флаводоксину. Можно полагать, что и каталазы Ec и Ma также содержат

дополнительный по сравнению с короткими каталазами С-концевой «флаводоксиноподобный» домен, поскольку их С-концевая последовательность в той же мере, что и С-концевая последовательность Pv, длиннее, чем у Bp (на 184 и 187 остатков соответственно для Ec и Ma и на 188 остатков для Pv).

Мы подсчитали «степени родства» сравниваемых на рис. 2 каталаз, используя метод [13]. Матрица подсчитанных «степеней родства» приведена на рис. 3. «Степень родства» представляет собой процент «числа соответствий» от общего числа сравниваемых аминокислотных остатков. «Число соответствий», в свою очередь, равно сумме идентичных и гомологичных остатков сравниваемой пары белков. Гомологичные остатки были разбиты на следующие группы: Val, Met, Ile, Leu (гидрофобные); Phe, Tyr, Trp (ароматические);

	Mn	Rn	Hs	Bp	Dm	Ps	Gh	At	Ib	Os	Le	Zn	Zm1	Zm2	Zm3	Bs	Lsa	Lse	Hi	Ma	Ec	Ct	Hp	ScA	ScT
Rn	96
1 Hs	93	92
Kp	93	93	93
2 Dm	73	73	74	74
Ps	52	52	51	52	50
Gh	51	52	51	52	50	92
At	51	52	50	52	50	90	94
Ib	51	51	50	51	49	85	87	88
Os	50	50	49	50	48	80	82	82	82
3 Le	52	52	51	53	50	84	85	85	85	80
Zn	50	50	50	51	48	85	88	86	82	78	81
Zm1	51	51	50	51	49	89	91	90	85	80	84	85
Zm2	43	43	43	43	40	63	66	65	62	61	61	71	65
Zm3	49	49	49	49	47	76	76	77	78	85	75	73	75	58
Bs	64	64	63	63	60	50	50	50	49	50	50	49	49	40	49
Lsa	57	58	58	58	56	46	46	47	45	46	47	46	46	37	45	67
4 Lse	51	51	50	52	49	52	53	54	50	51	53	52	52	44	49	59	53
Hi	54	54	55	56	52	46	46	47	45	46	47	47	45	39	45	59	56	51
Ma	47	47	46	49	46	49	49	50	49	50	48	48	49	39	49	51	47	51	44
Ec	48	48	48	51	47	49	49	50	49	50	50	48	48	40	49	55	50	53	50	59
Ct	55	55	55	56	59	46	45	45	43	44	45	44	45	36	43	52	52	46	49	45	46
5 Hp	55	55	55	55	57	44	45	44	44	44	44	44	45	35	42	53	50	47	48	45	45	72	.	.	.
ScA	53	52	52	54	53	42	42	42	41	42	43	41	42	35	41	52	51	44	48	42	42	70	61	.	.
ScT	45	44	45	45	45	39	39	39	36	39	38	38	39	33	38	45	44	42	40	40	40	51	48	46	.
6 Pv	45	45	45	47	42	39	39	40	38	38	39	38	39	29	40	42	39	39	37	37	43	39	37	35	32

Рис. 3. Матрица «степеней родства» первичных структур каталаз млекопитающих (1), насекомого (2), растений (3), бактерий (4), дрожжей (5), гриба (6)

Lys, Arg (основные); Asp, Glu (кислые); Asn, Gln (амиды); Ser, Thr (оксиаминокислоты).

Как видно из матрицы, «степень родства» каталаз колеблется в большом диапазоне—29—96 %. Очевидно, что каталазы не относятся к числу консервативных белков. Можно отметить некоторые закономерности эволюции первичных структур каталаз. «Степени родства» каталаз внутри каждой таксономической группы выше, чем между группами. Причем наибольшие наблюдаются внутри группы животных (93—96 %), а наименьшие — у бактерий (44—67 %), у растений — 61—92 % и у дрожжей — 46—72 %. Здесь видна некоторая корреляция между эволюцией первичных структур каталаз и организмов. Что касается «степеней родства» между группами, то наибольшей она является у пяти групп (2—6) с группой каталаз животных (1), чем между собой. У каталазы Pv наибольшая «степень родства» наблюдается, как отмечено выше, с каталазами животных (45—47 %), в то время как с каталазами других групп — 29—43 %.

Несмотря на эволюционную нестабильность первичной структуры каталаз, в полипептидной цепи можно выделить относительно консервативный и вариабельный участки. Мы рассчитали «степени родства» С-концевых участков и участка, отмеченного на рис. 2 кружочками, содержащего остатки His, Ser и Asn, входящие в активный центр каталазы Bp [14] (отмечены звездочкой на рис. 2). Величина сравнимых участков равна длине дополнительных С-концевых доменов длинных каталаз (180 остатков аминокислот). «Степень родства» дополнительных С-концевых последовательностей длинных каталаз Pv, Ma и Ec лежит в пределах 21—42 %. «Степени родства» С-концевых последовательностей такой же длины коротких каталаз и соответствующих им участков длинных каталаз лежат в тех же пределах. Однако «степень родства» участков, содержащих активный центр, лежит в пределах 55—80 %. Можно полагать, что С-концевые участки менее важны для основной функции каталаз. Однако, какова же роль дополнительного

С-концевого «флаводоксиподобного» домена длинных каталаз, если степень их родства низка, как и С-концевых участков коротких каталаз? Здесь важно отметить, что «степень родства» флаводоксиподобного родства структур высшего порядка важнее, чем первичных. В этом смысле интересно провести сравнительное исследование каталаз Pv, Es и Ma рентгеноструктурным методом, как это сделано для каталаз Pv и Vp [6].

Е. А. Козлов, Т. Л. Левитина, М. Т. Бобровская, Л. В. Гудкова, М. Ф. Радомский

Вивчення первинної структури каталази *Penicillium vitale*.

5. Реконструкція поліпептидного ланцюга та порівняння його з іншими каталазами

Резюме

На основі пептидів T6, T12, T29, T30, T32, T44, Tm2, Tm11, Tm28, Tm29, Tm35, Tm36, Tm39, Sp1, Sp2, Sp5, Sp11, Sp14, Sp26, Sp27, BrCN3—BrCN10 реконструйовано три фрагменти каталази *P. vitale*. В результаті порівняння амінокислотної послідовності цих фрагментів з первинною структурою інших каталаз побудовано безперервний поліпептидний ланцюг, який складається з 696 залишків амінокислот. Порівнюється первинна структура 26 каталаз з 23 організмів шести таксономічних груп. Побудовано матрицю «ступенів спорідненості» первинних структур досліджуваних каталаз.

Е. А. Козлов, Т. Л. Левитина, М. Т. Бобровская, Л. В. Гудкова, М. Ф. Радомский

Investigation of the *Penicillium vitale* catalase primary structure.

5. Reconstruction of the polypeptide chain and comparison it with other catalases

Summary

Three fragments of *P. vitale* catalase were reconstructed on the basis of T6, T12, T29, T30, T32, T44, Tm2, Tm11, Tm28, Tm29, Tm35, Tm36, Tm39, Sp1, Sp2, Sp5, Sp11, Sp14, Sp26, Sp27, BrCN3—BrCN10 peptides. The peptides structure was published earlier. The uninterrupted polypeptide chain of *P. vitale* catalase was structured from these fragments by means of their comparison with the amino acid sequences of 26 other catalases. This reconstructed polypeptide chain comprises 696 amino acid residues. The primary structure of 26 catalases from 23 origins of 6 taxonomic groups is comparing. The «degree of relatedness» matrix of catalases primary structure was constructed.

СПИСОК ЛІТЕРАТУРИ

1. Гусак Н. М., Левитина Т. Л., Бобровская М. Т. и др. Выяснение первичной структуры каталазы гриба *Penicil-*

lium vitale. 1. Триптические пептиды немодифицированной каталазы // Биополимеры и клетка.—1998.—14. № 1.—С. 62—67.

2. Левитина Т. Л., Гусак Н. М., Мирошниченко О. С. и др. Выяснение первичной структуры каталазы гриба *Penicillium vitale*. 2. Триптические пептиды модифицированной по остаткам лизина каталазы // Там же.—№ 2.—С. 105—110.
3. Левитина Т. Л., Латышко Н. В., Бобровская М. Т. и др. Выяснение первичной структуры каталазы гриба *Penicillium vitale*. 3. Пептиды, полученные путем расщепления каталазы протеиназой из *Staphylococcus aureus* V8 // Там же.—№ 3.—С. 191—195.
4. Бобровская М. Т., Левитина Т. Л., Роднин Н. В. и др. Выяснение первичной структуры каталазы гриба *Penicillium vitale*. 4. Бромциановые фрагменты // Там же.—С. 196—202.
5. Гудкова Л. В., Кириленко М. Т., Левитина Т. Л., Козлов Э. А. Исследование субъединичной структуры каталазы *Penicillium vitale* // Укр. биохим. журн.—1985.—57, № 4.—С. 29—33.
6. Vainshtein B. K., Melik-Adamjan W. R., Barynin V. V. et al. Three-dimensional structure of catalase from *Penicillium vitale* at 2.0 Å resolution // J. Mol. Biol.—1986.—188, N 4.—P. 49—61.
7. VanOssofski J., Hausner G., Loewen P. C. Molecular evolution analysis based on the amino acid sequence of catalase // J. Mol. Evol.—1993.—37.—P. 71—76.
8. Bethards L. A., Skadsen R. W., Scandalios J. G. Isolation and characterization of a cDNA clone for the Cat2 gene in maize and its homology with other catalases // Proc. Nat. Acad. Sci. USA.—1987.—84.—P. 6830—6834.
9. Redinbaugh M. G., Wadsworth G. J., Scandalios J. G. Characterization of catalase transcripts and their differential expression in maize // Biochim. et biophys. acta.—1988.—951.—P. 104—116.
10. Bishai W. R., Smith H. O., Barcak G. J. A peroxide ascorbate-inducible catalase from *Haemophilus influenzae* is homologous to the *Escherichia coli* katE gene product // J. Bacteriol.—1994.—176, N 10.—P. 2914—2921.
11. Milano A., De Rossi E., Gusberti L. et al. The katE gene, which encodes the catalase HP2 of *Mycobacterium avium* // Mol. Microbiol.—1996.—19, N 1.—P. 113—129.
12. Schroeder W. A., Shelton J. R., Shelton J. B. et al. The complete amino acid sequence of bovine liver catalase and the partial sequence of bovine erythrocyte catalase // Arch. Biochem. and Biophys.—1982.—214, N 1.—P. 397—421.
13. Haber J. E., Koshland D. E. An evaluation of the relatedness of proteins based on comparison of amino acid sequences // J. Mol. Biol.—1970.—50.—P. 617—639.
14. Fita J., Rossmann M. G. The active centre of catalase // Ibid.—1985.—185, N 1.—P. 21—37.

Поступила в редакцию 28.08.97