



EtAS W257,

EtAS Y259

S2

EtAS : CMRLLGEGPNG-GQDNACSRARKWIIDHGGATYIPSWGKTWLSILGVYEWISGNSNMPPEFWILP TFLPMHPAKM WYCRMVYMPMSYLYGKRFVGPITPL : 282  
 PNY1 : CMRILGEGPDG-GVNNACARGRKWILDHGSVTAIPSWGKTWLSILGVYEWIGSNMPPEFWILP SFLPMHPAKM WYCRMVYMPMSYLYGKRFVGPITPL : 284  
 BPY : CMRILGEGPDG-GQDNACARARKWILDHGGVTHMPSWGKTWLSILGIFEWIGSNMPPEFWILP SFLPMHPAKM WYCRMVYMPMSYLYGKRFVGPITPL : 282  
 PSY : CMRILGEGPDG-GEDNACVRARNWIRQHGGSVTHIPSWGKTWLSILGVEDWLGSNMPPEFWILP SFLPMHPAKM WYCRMVYMPMSYLYGKRFVGPITPL : 282  
 AtLUP1 : CLRMLGENP----EQDACKRARQWILDRGGVIFIPSWGKFWLSILGVYDWSGTNPTPELLMLP SFLPIHHPGKILYCRMVYMPMSYLYGKRFVGPITPL : 280  
 TRW : TLRLLGEEADSVAEADMALVGRKWILDHGGAVGIPSWGKFWLTIILGVYEWGGCNMPPEFWLMBKFFPIHHPGKMLYCRMVYMPMSYLYGKRFVGPITEL : 283  
 OEW : TLRLLGEGQED-GEDKAVARGRKWILDHGGAVGIPSWGKFWLTVLGVYEWGDCNMPPEFWLLENFSPHHPGKMLYCRMVYMPMSYLYGKRFVGPITGL : 281  
 BPW : ALRILGEGLED-GEDGAMAKSRKWILDHGGGLVAIPSWGKFWVTVLGVYEWSGCNPLPPEFWFLDIFPIHHPGKMLYCRMVYMPMSYLYGKRFVGPITGL : 280  
 AtCAS1 : TLRLLGEGPNDG--DGDMEKGRDWILNHGCATNITSWGKMWLSVLCAYEWSGNNPLPPEIWLLEPYLPIHHPGRMWHCRMVYLPMSYLYGKRFVGPITST : 280  
 PsCAS : TLRLLGEGPNDG--EGDMEKGRDWILEHGGATYITSWGKMWLSVLCAYEWSGNNPLPPEIWLLEPYLPIHHPGRMWHCRMVYLPMSYLYGKRFVGPITPT : 280  
 PgCAS : TLRLLGEGANDG--QGAMKGRQWILDHGSATAITSWGKMWLSVLCAYEWSGNNPLPPEIWLLEPYLPIHHPGRMWHCRMVYLPMSYLYGKRFVGPITPT : 280  
 BpCAS1 : TLRLLGE-PEDG--MGAVEKARKWILDHGGATAITSWGKMWLSVLCAYEWSGNNPLPPEIWLLEPYLPIHHPGRMWHCRMVYLPMSYLYGKRFVGPITST : 290  
 AtLAS1 : ALRLMGEELDGG--DGAMESARSWIHGGATFIIPSWGKFWLSVLCAYEWSGNNPLPPELWLLPYSLPFHPGRMWHCRMVYLPMSYLYGKRFVGPITNGT : 280  
 HsLAS : SLRILGVGPD----DPDLVRARNILHKKGGAVAIIPSWGKFWLAVLNIVYSWEGNLTLFPEMWFPEWVAHPSTIWHCRMVYLPMSYLYGKRFVGPITSAEDPL : 255  
 MmLAS : ALRILGIGPD----DPDLVRARNVILHKKGGAVAIIPSWGKFWLAVLNIVYSWEGNLTLFPEMWFPEWVAHPSTIWHCRMVYLPMSYLYGKRFVGPITSAEDPL : 256  
 ScLAS : ILRLGLPKD----HPVCAKARSTLLRLGCAIGSPHWKILWLSALNIYKWEVNEAPPETWLLPYSLPMHPGRMWHCRMVYLPMSYLYGKRFVGPITSAEDPL : 257  
 SHC : ALKYIGMSRD----EEPQKALRFIQSQGGIESSRVETRMWLAIVGEYPMWEKVPMPPEIMFLGKRMPLNIYEBGWARATVVALSIIVMSRQ---PVFPL : 189  
 64 6G r 6 Gg s5g4 W6 6g 5 W g np pPE w l f

Erg 7 W232

PNY1 Y261, Erg. H234

EtAS : ILQLR-QELHTQPYHHINWTKTRHLCAHEDVYYPHPLIQDLWDSLYIFTEPLLTRWPFNKIIRKKALEVTMKHIHYEDENSRYITIGCVER--VLCMLA : 379  
 PNY1 : ILQLR-EELYGQPYNEINWTKTRRVCAGEDIYYPHPLIQDLLWDSLYVLTPELLTRWPFNKI-REKALQTTMKHIHYEDENSRYITIGCVER--VLCMLV : 380  
 BPY : ILQLR-EELYTQPYHQVNWKKVRHLCAKEDIYYPHPLIQDLLWDSLYIFTEPLLTRWPFNKIIVREKALQVTMKHIHYEDENSRYITIGCVER--VLCMLA : 379  
 PSY : ILQLR-EELHTEPYEKNWTKTRHLCAKEDIYYPHPLIQDLIWDLSLYIFTEPLLTRWPFNKIIVREKALQVTMKHIHYEDENSRYITIGCVER--VLCMLA : 379  
 AtLUP1 : ILLLR-EELYLEPYEEINWKKSRRLYAKEDMYYAHPVQDLLSDTLQNFVEPLLTRWPLNKLIVREKALQVTMKHIHYEDENSHYITIGCVER--VLCMLA : 377  
 TRW : VRDLR-QELYTDPYDEINWTKARNTCAKEDLYYPHPLVQDMVWGVLDHNVVEPVLTSRPISTL-REKALKVAMDVHYEDKSSRYLCIGCVER--VLCMLA : 379  
 OEW : VLSLR-QEITYTEPYHGINWTKARNTCAKEDLYYPHPLAQMMLWGLDHHFAEPVLTTRWPFSKL-REKALKVAMDVHYEDMNSRYLCIGCVER--VLCMLA : 377  
 BPW : IQSLR-QELYNEPYHQINWTKARNTVAKEDLYYPHPLIQDLLWGLDHHVAEPVLTTRWPF SML-REKALKAAIGVHYEDENSKYLCIGSVER--VLCMLA : 376  
 AtCAS1 : VLSLR-KELETVPYHEVNWNEARNLCAKEDLYYPHPLVQDILWASLHKIVEPVLMRWPGANL-REKAIKTAIEHIHYEDENTRYICIGPVNK--VLNMLC : 376  
 PsCAS : VLSLR-KELETVPYHDIWQARNLCAKEDLYYPHPLVQDILWATLHKFVEPVFMNWPGKKL-REKAIKTAIEHIHYEDENTRYICIGPVNK--VLNMLC : 376  
 PgCAS : VLSLR-KEVESVPYHEIDWQARNLCAKEDLYYPHPLIQDILWASLKVWVEPIFMHWPAKKL-REKSLRTVMEHIHYEDENTRYICIGPVNK--VLNMLC : 376  
 BpCAS1 : IQSLR-KELYTVPYHEIDWTKARNDCAKEDLYYPHPLVQDILWASLYYAYEPIFMYWPAKRL-REKALDTVMQHIHYEDENTRYICIGPVNK--VLNMLC : 386  
 AtLAS1 : ILSLR-RELYTIPYHHIDWDTARNQCAKEDLYYPHPKTIQDVLWSDLNKFGEPLLERWPLNLL-RNHALQTVMQHIHYEDENTRYICIGPVNK--VLNMLC : 376  
 HsLAS : VQSLR-QELYVEDEASIDWLAQRNNVAPDELYYTPHPSWLLRVVYALLN----LYEHHHSAHL-RQAVQKLYBHVADDRFTKISIGPISK--TINMLV : 346  
 MmLAS : VQSLR-QELYVQDYASIDWPAQRNNVSPDEMYTTPHPSWLLRVVYALLN----LYERFHTSL-RKAVQMLYBHVAAADDCTKISIGPISK--TINMLV : 347  
 ScLAS : LEELR-NEIYTKPFDKINWTKARNTVCGVDLYYPHSTTLNIA-NSLVVIFYEKYLRNRFIYSLSKKRVYDL----IKTELQNTDSLCLABVWQ--AFCALV : 349  
 SHC : PERARVPELYETDV-----PPRRRGAKGG----GGWIFDALDRALHGYQK--LSVHPFRRAAEIRALDWLLERQAGD-----SWGGIQPPWFYALIA : 271  
 lR E6 p R a d yph L e p a h e d i g 6 k 6

## EtAS F413

## EtAS F474

EtAS : CW-AEDPNGVPEKKHLARI PDYMWVAEDGMKMQSF-GSQQWDTGFAIQALLASNLTE--EIGQVLKKGHDFFIKKSQVKENP-SGDFKSMHRHISKGSWTF : 474  
 PNY1 : CW-VEDPNGDYERKHLARI PDYIWWVAEDGMKMQSF-GSQEWDGTGFSIQALLSDSLTH--EIGPTLMKKGHDFFIKKSQVKDNP-SGDFKSMYRHHISKGSWTF : 475  
 BPY : CW-VEDPNGDYEEKHIARI PDYIWWVAEDGIKMQSF-GSQEWDGTGFAIQALLASNLTD--EIGPTLARGHDFFIKKSQVKDNP-SGDFESMHRHISKGSWTF : 474  
 PSY : CW-VEDPNGDAEKKHIARV PDYLWVISEDGMTMQSF-GSQEWDAGFAVQALLATNLIE--EIKPALARGHDFFIKKSQVTENP-SGDFKSMHRHISKGSWTF : 474  
 AtLUP1 : CW-VBNPNGDYEEKHLARI PDYMWVAEDGMKMQSF-GCQLWDTGFAIQALLASNLDP--ETDDALKRGHNYIKASQVRENK-SGDFRSMYRHHISKGSWTF : 472  
 TRW : TW-VEDPNGDAYKRHLARI PDYFVVAEDGMKMQSF-GCQMWDAFAIQATFSSNLTE--EYGPTLKKAHAEFVKASQVRDNP-PGDFSKMYRHTSKGSWTF : 474  
 OEW : CW-VEDPNSEAYKRHIARI PDYFVVAEDGLKMQSF-GCQMWDAFAIQAILSSNLAE--EYGPTLMKAHAEFVKASQVQENP-SGDFNEMYRHTSKGSWTF : 472  
 BPW : CW-AEDPNGEAYKHLGRIPDNYWVAEDGLKIQSF-GCQMWDAFAIQAILSCNLNE--EYWPTLRKAHAEFVKASQVPENP-SGDFKAMYRHINKGSWTF : 471  
 AtCAS1 : CW-VEDPNSEAEKHLHPRIHDFLWVAEDGMKMQCYNGSQLWDTGFAIQAILATNLVE--EYGPVLEKAHAEFVKNSQVLEDC-PGDLNYWYRHHISKGSWTF : 472  
 PsCAS : CW-VEDPNSEAEKHLHPRIYDYLWVAEDGMKMQCYNGSQLWDTAFAAQAIISTNLID--EFGPTLKRKAHAFIKNSQVSEDC-PGDLNYSKWRHHISKGSWTF : 472  
 PgCAS : CW-VEDPNSEAEKHLHPRLHDFLWVAEDGMKMQCYNGSQLWDTAFAVQAIISTNLAE--EYGPTLRKAHAFMKNNSQVLDCC-PGDLDAWYRHHISKGSWTF : 472  
 BpCAS1 : CW-AEDPNSEAEKHLHPRIILDYLVVAEDGMKMQCYNGSQLWDTTFAVQAIISTNIAE--EYQTLRKAHEYIKDSQVLEDC-PGDLNFWYRHHISKGSWTF : 482  
 AtLAS1 : CW-VESNSEAEKSHLSRIKDYLVVAEDGMKMQCYNGSQLWDTLAVQAILATNLVD--DYGLMLKKAHNYIKNTQIRKDT-SGDPGLWYRHHISKGSWTF : 472  
 HsLAS : RMYVDGPASTAEQEHVSRI PDYLWVGLDGMKMQCTNGSQIWDTAFAIQALLEAGGHRPEFSSCLQKAHEFLRLSQVDPNP-P-DYQKYRQMRKGGGSE : 444  
 MmLAS : RWSVDGPPSSPAEQEHVSRIKDYLVWGLDGMKMQCTNGSQIWDTSFAIQALLEAGAHHRPEFLPCLQKAHEFLRLSQVPENC-P-DYQKYRHRMRKGGGSE : 445  
 ScLAS : TLIEGVDSEAEQRLQYRFKDALFHGPOGMITMGTNGVQTDWCAFAIQYEFVAGLAERPEFYNTIIVSAYKELCHAQEDTECVPGS----YRDKRKGGGSE : 445  
 SHC : LKILDMTQHPAEIKGWEGLELVGVELDYGGWVFCASISPVNDTGLAVLALRAAGLPADHD---RLVRAGEMILLDROIT---VPGDWAVKRPNLKGGEFAE : 365

w e b 5 h r d g q WD fa qa e 6 h 56 Q -- -h kG 5 F

## EtAS V483

D<sup>485</sup>CTAE

## Erg 7 T384

## EtAS W534

## EtAS E563

## EtAS C564

## Erg 7 F445

EtAS : SDQDHGWCVSDCTAEGKCCLLFSMPP-EIVGKMDAQHLYNAVNIILISLQS----KNGGLAAMEPAGAQQWLEMLNPTFEFFADIVIEHEYVECTASAI : 569  
 PNY1 : SDQDHGWCVSDCTAEGKCCLLFSMPPE-EIVGKKIKPERLYDSVNVLLSLQS----KNGGLSAMEPAGAQEWLEMLNPTFEFFADIVIEHEYVECTSSAI : 570  
 BPY : SDQDHGWCVSDCTAEGKCCLLFSIMPP-EIVGKMEPEQLYDSVNVLLSLQS----KNGGLAAMEPAGAQEWLEMLNSTEFAADIVIEHEYVECTASAM : 569  
 PSY : SDQDHGWCVSDCTAEGKCCLLLSLPP-EIVGKMEPERLFDVNVLLSLQS----KNGGLAAMEPAGAQEWLEMLNPTFEFFADIVVEHEYVECTGSAI : 569  
 AtLUP1 : SDRDHGWCVSDCTAEGKCCLLLSMMSA-DIVGQKIDDEQLYDSVNVLLSLQS----GNGGVNAMEPSPRAYKWLEMLNPTFEFFADIVVEHEYVECTSSVI : 567  
 TRW : SIQDHGWCVSDCTAEGKVSLLYSQMPN-KLVGKQVETEHLYDAVNVILSLQS----ENGGFPAMEPQRAYAWLEMLNPTFEFFEDVLIERYVECTSSAI : 569  
 OEW : SMQDHGWCVSDCTAEGKVAALLFSQMPI-ELVGAETETGHLYDAVNVILSLQS----ASGGFPAMEPQKAYRWLEMLNPTFEFFEDVLIERYVECTSSAV : 567  
 BPW : SMQDHGWCVSDCTAEGKVAALLFSQMPN-DLVGKIEKERLYDAVNVILSLQS----SNGGFPAMEPQRAYGWLEMLNPTFEFFEDVLIERYVECTSPAV : 566  
 AtCAS1 : STADHGWEI SDCTAEGKVAALLSKVPK-AIVGEPIDAKRLYEAVNVIIISLQN----ADGGLATYELTRSYPWLEMLNPAETEGDIVIDYPPVECTSAAI : 567  
 PsCAS : STADHGWEI SDCTAEGKVAALLSKVAP-EIVGEPIDSKRLYEAVNVIIISLQN----ENGGLATYELTRSYTWLEMLNPAETEGDIVIDYPPVECTSAAI : 567  
 PgCAS : STADHGWEI SDCTAEGKVAALLSKVPS-ELVGEPLDAKRLYEAVNVIIISLQN----SDGGYATYELTRSYSWLEMLNPAETEGDIVIDYPPVECTSAAI : 567  
 BpCAS1 : STADHGWEI SDCTAEGKVAALLSQFPS-ETVGSVDVVRLYDAVNVILSLQN----TDGGFATYELTRSYHWLEMLNPAETEGDIVIDYPPVECTSAAI : 577  
 AtLAS1 : STGDNFWEI SDCTAEGKVAALLSQMPV-NLVGEPMPPEHLVDAVNEIISLQN----KNGGFASYELTRSYPELEVINPSETEGDIIIDYQYVECTSAAI : 567  
 HsLAS : STLDCGWIIVSDCTAEGKVAALLLQEKCP-HVT-EHPRERLCDAVAVLLNMRN----PDGGFATYETKRGHLEMLNPNSEVEGDIMIDYTYVECTSAVI : 538  
 MmLAS : STLDCGWIIVSDCTAEGKVAALLLQNCQP-SIT-EHPRERLCDAVDVLLSLRN----ADGGFATYETKRGHLEMLNPNSEVEGDIMIDYTYVECTSAVI : 539  
 ScLAS : STKTQYTVADCTAEAIKAIIMVKNSPVFSEVHHMISSEERLEFEGIDVLLNIGNISGFYEGSFATYEKIKAPLAMETLNPAEVEGDIMVEYPPVECTDSSV : 545  
 SHC : QFDNVVYIPDVEDTA-----VWVWALNLTLPDERRRRDAMTKGFRWLVGMQS----SNGGWGAYDNDNTSDLPNHI-PFCDEGEVT-DPPSEIVTAHVIL : 453

sDcTAE k 6 v 6 66 6q Ga 5e e np e f

## Erg 7. V454

## DXDD for SHC

## Erg 7 Y510

## Erg 7 E539

## Erg 7 C540

```

*           420           *           440           *           460           *           480           *           500
EtAS   : CM-AEDPNGVPEKKHLARI PDYMWVAEDGMKMQSF-GSQQWDTGFAIQALLASNLTE--EIGQVLKKGHDFIKKSQVKENP-SGDFKSMHRHISKGSWTF : 474
PNY1   : CM-VEDPNGDYERKHLARI PDYIWWVAEDGMKMQSF-GSQEWDGTGFSIQALLSDSLTH--EIGPTLMKGHDFIKKSQVKDNP-SGDFKSMYRHHISKGSWTF : 475
BPY    : CM-VEDPNGDYEEKHIARI PDYIWWVAEDGIKMQSF-GSQEWDGTGFAIQALLASNLTD--EIGPTLARGHDFIKKSQVKDNP-SGDFESMHRHISKGSWTF : 474
PSY    : CM-VEDPNGDAEKKHIARV PDYLWISEDGMTMQSF-GSQEWDAGFAVQALLATNLTIE--EIKPALARGHDFIKKSQVTENP-SGDFKSMHRHISKGSWTF : 474
AtLUP1 : CM-VBNPNGDYEEKHLARI PDYMWVAEDGMKMQSF-GCQLWDTGFAIQALLASNLPD--ETDDALKRGHNYIKASQVRENP-SGDFRSMYRHHISKGAWTF : 472
TRW    : TW-VEDPNGDAYKRHLARI PDYFWVAEDGMKMQSF-GCOMWDAFAIQATFSSNLTE--EYGPTLKKAHAEFVKASQVRDNP-PGDFSKMYRHTSKGAWTF : 474
OEW    : CM-VEDPNSEAYKRHIARI PDYFWVAEDGLKMQSF-GCOMWDAFAIQAILSSNLAE--EYGPTLMKAHAEFVKASQVQENP-SGDFNEMYRHTSKGAWTF : 472
BPW    : CM-AEDPNGEAYKHLGRIPDNYWVAEDGLKIQSF-GCOMWDAFAIQAILSCNINE--EYWPTLRKAHAEFVKASQVPENP-SGDFKAMYRHHINKGAWTF : 471
AtCAS1 : CM-VEDPNSEAEKHLHPRIHDFLWLAEDGMKMQGYNGSQLWDTGFAIQAILATNLVE--EYGPVLEKAHAEFVKNSQVLEDC-PGDLNYWYRHHISKGAWPF : 472
PsCAS  : CM-VEDPNSEAEKHLHPRIYDYLWVAEDGMKMQGYNGSQLWDTAFAAQAIISTNLID--EFGPTLKRKAHAFIKNSQVSEDC-PGDLKQYRHHISKGAWPF : 472
PgCAS  : CM-VEDPNSEAEKHLHPRLHDFLWLAEDGMKMQGYNGSQLWDTAFAVQAIISTNLAE--EYGPTLRKAHAFMKNSQVLDLDC-PGDLDAWYRHHISKGAWPF : 472
BpCAS1 : CM-AEDPNSEAEKHLHPRIILDYLVIAEDGMKMQGYNGSQLWDTTFAVQAIISTNIAE--EYQTLRKAHEYIKDSQVLEDC-PGDLNFYRHHISKGAWPF : 482
AtLAS1 : CM-VESNSEAEKSHLSRIKDYLVVAEDGMKMQGYNGSQLWDTLAVQAILATNLVD--DYGLMLKKAHNYIKNTQIRKDT-SGDPGLWYRHHISKGAWPF : 472
HsLAS  : RMYVDGPASTAEQEHVSRI PDYLWMLGLGMKMQGTNGSQIWDTAFAIQALLEAGGHRPEFSSCLQKAHEFLRLSQVDPNP-P-DYQKYRQMRKGGFSEF : 444
MmLAS  : RWSVDGPPSSPAEQEHVSRIKDYLVWGLGLGMKMQGTNGSQIWDTSFAIQALLEAGAHHRPEFLPCLQKAHEFLRLSQVPENC-P-DYQKYRHRMRKGGFSEF : 445
ScLAS  : TLIEGVDSEAEQRLQYRFKDALFHGPOGMITMGTVGVTWDCAFAIQYFFVAGLAERPEFYNTIIVSAYKELCHAQEDTECVPGS----YRDKRKGAWGF : 445
SHC    : LKILDMTQHPAEIKGWEGLELVGVELDYGGWVFQASISPVNDTGLAVLALRAAGLPADHD---RLVRAGEMILLDROIT---VPGDWAVKRPNLKEGGFAF : 365
w e p 5 h r d dG k6q g q WD fa qa e 6 h 56 Q gd rh kG 5 F

```

```

*           520           *           540           *           560           *           580           *           600
EtAS   : SDQDHGWQVSDCTAEGLKCLLFSMMPPEIVGKEMDAQHLYNAVNIILISLQS----KNGGLAAMEPAGAQQWLEMLNPTFEFFADIVIEHEHYVECTASAI : 569
PNY1   : SDQDHGWQVSDCTAEGLKCLLFSMMPPEIVGKKIKPERLYDSVNVLLSLQS----KNGGLSAMEPAGAQEWLEMLNPTFEFFADIVIEHEHYVECTSSAI : 570
BPY    : SDQDHGWQVSDCTAEGLKCLLFSIMPP-EIVGKEMPEQLYDSVNVLLSLQS----KNGGLAAMEPAGAQEWLEMLNPTFEFFADIVIEHEHYVECTASAM : 569
PSY    : SDQDHGWQVSDCTAEGLKCLLFSLLPP-EIVGKEMPERLFDVNVLLSLQS----KNGGLAAMEPAGAQEWLEMLNPTFEFFADIVIEHEHYVECTGSAM : 569
AtLUP1 : SDRDHGWQVSDCTAEGALKCLLFSMMSA-DIVGQKIDDEQLYDSVNVLLSLQS----GNGGVNAMEPSRAYKWLEMLNPTFEFFADIVIEHEHYVECTSSVI : 567
TRW    : SIQDHGWQVSDCTAEGLKVSLLYSQMNP-KLVGKIVETEHLVDVNVVILSLQS----ENGGFPAMEPQRAYAWLEMLNPTFEFFADIVIEREYVECTSSAI : 569
OEW    : SMQDHGWQVSDCTAEGLKAAALLFSQMPI-ELVGAETETGHLYDAVNVVILSLQS----ASGGFPAMEPQKAYRWLEMLNPTFEFFADIVIERDYVECTSSAV : 567
BPW    : SMQDHGWQVSDCTAEGLKVAALLFSQMPPEIVGKIEKERLYDAVNVVILSLQS----SNGGFPAMEPQRAYAWLEMLNPTFEFFADIVIEREYVECTSPAV : 566
AtCAS1 : STADHGWPI SDCTAEGLKAAALLSKVPK-AIVGEPIDAKRLYEAVNVIIISLQN----ADGGLATYELTRSYPWLEMLNPTFEFFADIVIEHEHYVECTSAAI : 567
PsCAS  : STADHGWPI SDCTAEGLKAVLLLSKIAP-EIVGEPIDSKRLYDAVNVVILSLQN----ENGGLATYELTRSYTWLEMLNPTFEFFADIVIEHEHYVECTSAAI : 567
PgCAS  : STADHGWPI SDCTAEGFKAVLQLSKLPS-ELVGEPLDAKRLYDAVNVVILSLQN----SDGGYATYELTRSYSWLEMLNPTFEFFADIVIEHEHYVECTSAAI : 567
BpCAS1 : STADHGWPI SDCTAEGLKAVLLLSQFPS-ETVGKSVVVKRLYDAVNVVILSLQN----TDGGFATYELTRSYHWLEMLNPTFEFFADIVIEHEHYVECTSAAI : 577
AtLAS1 : STGDNFMPVSDCTAEGALKAAALLSQMPV-NLVGEPMPPEHLVDVNVVILSLQN----KNGGFASYELTRSYPELEVINPSETEGDIIIDYQYVECTSAAI : 567
HsLAS  : STLDCEMIVSDCTAEGALKAVLLLQEKCP-HVT-EHIPRERLCDAVAVLLNMRN----PDGGFATYETKRGHLEMLNPTFEFFADIVIEHEHYVECTSAAM : 538
MmLAS  : STLDCEMIVADCTAEGALKAVLLLQNCQP-SIT-EHIPRERLCDAVDVLLSLRN----ADGGFATYETKRGHLEMLNPTFEFFADIVIEHEHYVECTSAAM : 539
ScLAS  : STKTQYTVADCTAEAIKAIIMVKNSPVFSEVHHMISSERLFEIGIDVLLNLRNIGSFYEGSFATYETKIKAPLAMETLNPAEVEFGDIMVEYVECTDSSV : 545
SHC    : QFDNVIYVDPVDDTA-----VWVWALNLTLPDERRRRRDAMTKGFRWLVGMQS----SNGGFGAYDVDNTSDLPNHIFPFCDEGEVTDPPSEDVTAHV : 453
s g5 sDcTAE k 6 v 6 66 6q Gg 5e e n p e f e c t 6

```



EtAS W612

		*	620	*	640	660	*	680	*	700				
EtAS	:	HALIMFKKLY	PGHRKKEIENFI	TNAVKYLEDVQ	TADGGMYGNW	GVCFITYGTWFAVGG	LAAAGKNYNN	---	CAAMRRAVDFL	IRTKQKQDGGW	GESYLSCP			
PNY1	:	QALVLEKFKLY	PGHRKKEIDNEFI	TNAVRYLEDT	QMPDGSWYGNW	GVCFITYGSWFALGG	LAAAGKTYYN	---	CAAVRRAVEFLL	KSQMDDGGW	GESYLSCP			
BPY	:	OTLVLEKFKLY	PGHRKKEIENFI	KNAAQFLQVI	QMPDGSWYGNW	GVCFITYGTWFAVGG	LAAVGGKTYNN	---	CLAVRRAVDFL	LRAQRDNGG	WGESYLSCP			
PSY	:	QALVLEKFKLY	PGHRKKEIENFI	ENAVRELEDT	QTEDGSWYGNW	GVCFITYGSWFALGG	LAAAGKTYTN	---	CAAIRRGVKFLL	TTQOREDDGG	WGESYLSS			
AtLUP1	:	QALDLEKFKLY	PDHRKKEINRSI	EKAVQFIQDN	QTPDGSWYGNW	GVCFITYATWFAVGG	LAAAGETYN	---	CLAMRNGVHFLL	TTQRDDGGW	GESYLSCSE			
TRW	:	OGLTLEKFKLY	PGHRTKEIEHCI	SRAVKYVEDT	QESDGSWYGNW	GVCFITYGTWFAVDAL	VACGKNYHN	---	CPALQKACKFLL	SKQLPDGGW	GESYLSSSN			
OEW	:	QALKLEKQLHP	PGHRRKEIASCI	SKAIQYIEAT	QNPDGSWYGNW	GVCFITYGTWFAVGG	LAAAGKTYNN	---	SPTLRRACEFLL	SKQLPDGGW	GESYLSSSN			
BPW	:	HGLALEKFKLY	PGHRTKEIEHCI	SRAVKYVEDT	QESDGSWYGNW	GVCFITYGTWFAVGG	LAAAGKTYNN	---	SPTLRRACEFLL	SKQLPDGGW	GESYLSSSN			
AtCAS1	:	QALISERKLY	PGHRKKEVDECI	EKAVKFIESI	QAADGSWYGNW	GVCFITYGTWFGV	KGLVAVGKTL	---	SPHVARACEFLL	SKQPPSGGW	GESYLSCQD			
PcCAS	:	QALATEFKLY	PGHREIEIQCCI	EKAVAFIEKI	QASDGSWYGNW	GVCFITYGTWFGI	KGLIAAGKNE	---	CLSIKFACEFLL	SKQLPSGGW	AESEYLSCQ			
PgCAS	:	QALTAEFKLY	PGHREIEIQHSI	EKAALFIEKI	QSSDGSWYGNW	GVCFITYGTWFGI	KGLVAVGKTL	---	CASIKFACEFLL	SKQVAVGGW	GESYLSCQ			
BpCAS1	:	QALTFEKKLY	PGHREIEIENCI	AKAAEFIENI	QASDGSWYGNW	GVCFITYGTWFGI	KGLVAVGKTL	---	CSSIHFACEFLL	SKELASGGW	GESYLSCQ			
AtLAS1	:	OGLVLEF	TTLNNSYKRKEI	VGSIKAVEFIE	KTQLPDGSWYGNW	GVCFITYATWFGI	KGLVAVGKTL	---	SLCIRKACG	FLLSKQLCCGG	WGESYLSCQ			
HsLAS	:	QALKYFHKRF	FEHRAAEIRET	LQGLEFCRR	QQRADGSWEGS	WGVCFITYGTWFG	LEAFACMGQ	TYRDTG	TACA	EVSRACDFLL	SRQADGGW	GEDFESCEE		
MmLAS	:	QALKHEHEHF	FDYRAAEVRET	LNQGLDFCRR	KQRADGSWEGS	WGVCFITYGTWFG	LEAFACMGHT	YQDGA	ACAEVAQ	ACNFLLSQ	QADGGW	GEDFESCEE		
ScLAS	:	LGLTYFHKYF	-DYRKEEIRTR	IRIAIEFIKKS	QLPDGSWYGNW	GVCFITYGTWFG	LEAFACMGHT	YQDGA	ACAEVAQ	ACNFLLSQ	QADGGW	GEDFESCEE		
SHC	:	ECFGSE	-----GY--	DDAWKVI	RRAVEYLLKRE	QKPDGSWEGS	WGVCFITYGTWFG	LEAFACMGHT	YQDGA	ACAEVAQ	ACNFLLSQ	QADGGW	GEDFESCEE	
		1	F	p	e	6	a	5	Q	DGSW	6	G	GGWgEs	S

Erg 7 W587

EtAS F728, M729

EtAS L734

EtAS Y736

		*	720	*	740	760	800												
EtAS	:	KKYVPLE	EDNRSNLVHT	SWALMGLI	SAGQMDR	PTPLHRAAKLL	LINSQLEDGDF	PQCEITGV	EMRNCMLHY	AAYRN	IYPLWAL	ABYRNRV	PLPSTTL	---					
PNY1	:	KVYVPLE	EGNRSNLVHT	GWALMGLI	HSEQAER	PTPLHRAAKLL	LINSQMEDGDF	PQCEITGV	EMRNCMLHY	AAYRN	IYPLWAL	ABYRRRV	PLPSLGT	---					
BPY	:	KEYVPLE	EGNKSNLVHT	AWAMMGLI	HAGQAER	PTPLHRAAKLL	LINSQLEDGDF	PQCEITGV	EMRNCMLHY	AAYRN	IYPLWAL	ABYRKHV	PLP-LG	KNLN					
PSY	:	KIYVPLE	EGNRSNLVHT	AWAMMGLI	HAGQAER	PTPLHRAAKLL	LINSQLEQGDW	PQCEITGV	EMRNCMLHY	PMYRDI	YPLWAL	ABYRRRV	PLP-	-----					
AtLUP1	:	QRYIP	SEGERSNLV	QTSWAMMALI	HGTQAER	DLIPLHRAAKLL	LINSQLENGDF	PQCEITGA	EMRNCMLHY	ATYRN	TFPLWAL	ABYRKVV	FEI	-----					
TRW	:	KVYTN	LEGNRSNLV	HTSWALL	SLIKAGQAE	IDPTPI	SNGVRL	LINSQMBEGDF	PQCEITGV	EMRNCMLHY	SSERN	IFPIWAL	GEYRRIV	-----					
OEW	:	KVYTN	LEGNRSNLV	QTSWALL	SLIKAGQVE	IDPGPI	IHRGIR	LLVNSQMEDGDF	PQCEITGA	EMRNCMLHY	SSYRN	IFPIWAL	GEYRRRI	-----					
BPW	:	KVYTN	LEGNRANLV	QTSWALL	SLIDARQAE	IDPTPI	IHRGVR	VLLNSQMEDGDF	PQCEITGV	EMRNCMLHY	SSYRN	IFPIWAL	GEYRRRV	-----					
AtCAS1	:	KVYSN	LDGNRSHV	NTAWAMLALI	GAGQAEV	DRKPLHRAA	RYLINAQ	MENGDF	PQCEITGV	EMRNCMLHY	AAYRN	IFPIWAL	GEYRCQV	-----					
PcCAS	:	KVYSN	LEGNRSHV	NTAWAMLALI	EAEQAKR	PTPLHRAA	VCLINSQ	LENGDF	PQCEITGV	EMRNCMLHY	AAYRC	IFPIWAL	GEYR-RV	-----					
PgCAS	:	KVYTN	LEGNRSHV	NTAWAMLALI	DAGQAER	DATPLHRAA	KLINSQ	MENGDF	PQCEITGV	EMRNCMLHY	AAYRN	IFPIWAL	GEYRCRV	-----					
BpCAS1	:	KVYTN	LKDNRP	PHIVNTAWAMLALI	DAGQAER	PTPLHRAA	RIILINSQ	MENGDF	PQCEITGV	EMRNCMLHY	AAYRN	IFPIWAL	GEYRCRV	-----					
AtLAS1	:	KVYTN	LPGNKSHI	VNTSWALLALI	EAGQASR	DPMLHRAA	KSLLINSQ	MEDGDY	PQCEITGV	EMRNCMLHY	AAYRN	IFPIWAL	GEYRKLM	-----					
HsLAS	:	RRY--	LQSAQS	QHNTCWAMMGL	MAVRHP--	DI	EAQER	GVRCLLEK	QLENGDW	PQENIAGV	ENKSCA	ISYTSYRN	IFPIWAL	GRFSQLYPERALAGHP-					
MmLAS	:	RRY--	VQSARS	QVHSTCWALMGL	MAVRHP--	DI	TAQER	GIRCLLGK	QLENGDW	PQENISGV	ENKSCA	ISYTSYRN	IFPIWAL	GRFSNLYPDNTLAGHI-					
ScLAS	:	HSY--	VDSEKSL	VVQTAWALLALI	FAEYPNKEV--	IDR	GIDLLK	NRQEE	SGEWKFE	SVVEGV	ENHSCA	ISYPSYR	ELFKAL	GMYSRAYETHTL					
SHC	:	PAYAGK	GASTPS--	QAWAMMALI	AGGRAE	SEAA--	RRG	VQYLVET	QRPDGGW	DEPYTY	GTGPE	DFYLSY	MYRHV	FTLALGRYKQAI					
		5	s	v	T	WA66	L6	d	r	6	n	Q	e	Gd5p	Gv	54	5P	wAL	5

Erg 7 L705

Erg 7 F699, N700

Erg 7 Y707

```

          *
EtAS      : ----- : -
PNY1     : ----- : -
BPY      : QVVNCIGQSPLYKKYK : 779
PSY      : ----- : -
AtLUP1   : -----VN : 757
TRW      : -----QNI-- : 758
OEW      : -----LHAQT : 758
BPW      : -----LFA-- : 755
AtCAS1   : -----LLQQGE : 759
PsCAS    : -----L--QAC : 756
PgCAS    : -----L-QGPS : 758
BpCAS1   : -----L--KAL : 767
AtLAS1   : -----L---SL : 756
HsLAS    : ----- : -
MmLAS    : ----- : -
ScLAS    : ----- : -
SHC      : -----ERR : 631

```

The triterpene cyclases amino acids sequences were aligned using Clustal W, as implemented in the (CLC Sequence Viewer (CLC bio), and the figure was made by GenDoc (<http://www.nrbsc.org/gfx/genedoc/>). EtAS: *Euphorbia tirucalli*  $\beta$ -amyirin synthase (AB206469), PNY1: *Panax ginseng*  $\beta$ -amyirin synthase (AB009030), BPY: *Betula platyphylla*  $\beta$ -amyirin synthase (AB055512), PSY: *Pisum sativum*  $\beta$ -amyirin synthase (AB034802), AtLUP1: *Arabidopsis thaliana* multifunctional triterpene cyclase (At1g78970), TRW: *Taraxacum officinale* lupeol synthase (AB025345), OEW: *Olea europaea* lupeol synthase (AB025343), BPW: *Betula platyphylla* lupeol synthase (AB055511), AtCAS1: *Arabidopsis thaliana* cycloartenol synthase (At2g07050), PsCAS: *Pisum sativum* cycloartenol synthase (D89619), PgCAS: *Panax ginseng* cycloartenol synthase (AB009029), BpCAS1: *Betula platyphylla* cycloartenol synthase (AB055509), AtLAS1: *Arabidopsis thaliana* lanosterol synthase (At3g45130), HsLAS: *Homo sapiens* lanosterol synthase (P48449), MmLAS: *Mus musculus* lanosterol synthase (AK044016), ScLAS: *Saccharomyces cerevisiae* lanosterol synthase (P38604), SHC: *Alicyclobacillus acidocaldarius* squalene-hopene cyclase

**Table S1.** Functions of the active site residues. The locations of each of the residues in the reaction cavity are also shown in Scheme 2 and Fig. 1.

EtAS				<i>S. cerevisiae</i> LaS (Erg7)			
Catalytic Sites	Constructed Mutants	Product Profile (see Fig. 2) and Site Locations	Functions	Sites Equivalent to EtAS (see Fig. S1 in the ESI)	Constructed Mutants	Product Profile	Functions
W257 <sup>23</sup>	A, V, L, F, Y	Pentacyclic lupeol <b>26</b> as major product. Located at E-ring site.	Cation- $\pi$ complex with oleanyl cation <b>9</b> and CH- $\pi$ complex with Me-30 of substrate, which is quite different from W232 in Erg7.	W232 <sup>29</sup>	W232X	Protosta-12,24-dien-3 $\beta$ -ol ( <b>I</b> ), lanosterol ( <b>II</b> ), parkeol ( <b>III</b> ).	Discussed based on the $\pi$ -electron and the steric effect on the H234:Y510-hydrogen bond network. Located in proximate to H234, thus mutations at this position impair the H234:Y510 catalytic dyad structure.
Y259 <sup>23</sup>	A, V, I, L, H, F, W	Tetracycle <b>18</b> as major products. Located at D-ring site.	Cation- $\pi$ complex with baccharenyl cation <b>7</b> .	H234 <sup>30-32, 34</sup>	H234X	Achilleol A ( <b>11</b> ), tricycle (13 $\alpha$ H)-isomalabarica-14(26), 17 <i>E</i> , 21-trien-3 $\beta$ -ol ( <b>IV</b> ), <b>I</b> and <b>II</b> .	Located at C/D-ring site. Hydrogen bonding with Y510. Responsible for the formation of H234:Y510 catalytic dyad structure.
F413 <sup>23</sup>	A, V, M, S, T, Y, W	Tetracycle <b>18</b> as major products. Located at D-ring site.	Appropriate steric bulk, not $\pi$ -electron.	T384 <sup>36, 38</sup>	T384Y	Tetracycles with double bonds different from lanosterol (e.g., <b>III</b> ,	Discussed by combination with V454 and Q450, <sup>38</sup> but the function is not clearly assigned. Located at D-ring site.

						9 $\beta$ -lanosta-7,24-dien-3 $\beta$ -ol <b>V</b> , and <b>II</b> .	
F474 <sup>21</sup>	G, A, V, L, M, T, H, Y, W	Bicyclic products <b>12–16</b> were produced from the Gly and Ala mutants by chair-boat and chair-chair folding conformations. Located at B-ring site.	Steric bulk, not $\pi$ -electron, significantly differing from Erg7 F445 residue.	F445 <sup>39</sup>	F445X	<b>II, III, IV</b> and <b>V</b>	Importance of $\pi$ -electron density was proposed (perhaps cation- $\pi$ interaction), differing from the functional analysis of the EtAS F474, but the function is not clearly specified. Interacts with Tyr99, Trp390, Tyr510, Tyr707 and Tyr710 residues to form a $\pi$ -electron-rich pocket. The function of the F445Y mutant has not been reported. Located at A/B-ring site.
V483 <sup>26</sup>	G, A, S, I, F	Monocycles <b>10</b> and <b>11</b> from Gly and Ala mutants, but $\beta$ -amyrin <b>2</b> from the Ile variant. Located at A-ring site.	Steric bulk.	V454 <sup>35-37</sup>	G, A	Camelliol C ( <b>10</b> ) and <b>11</b> from Gly and Ala mutants, but lanosterol <b>II</b> from Leu and Ile mutants. Camelliol C synthase: <sup>37</sup> which is replaced by Ala at this position.	Steric bulk, the same function as in EtAS. Initiation (A-ring) site.



D485 <sup>12</sup>	N, E	No enzymatic products generated. Initiation site for the cyclization reaction. D485C486TA motif.	Proton attack at epoxide ring.	D456 <sup>18a, b</sup>  (human lanosterol D455 <sup>9</sup> )	N	—  (not reported)	Presumed based on human lanosterol X-ray analysis but no mutation analysis. Proton attack on the epoxide ring, initiation site for the cyclization reaction, D455C456TA motif.
C486 <sup>12</sup>	A	Decreased activity by ~50%.  Initiation site for the cyclization reaction.	Increases the acidity of D485.	C457 <sup>34</sup>  (human lanosterol C456 <sup>9</sup> )	C457G <sup>34</sup>	<b>11</b>	Presumed based on human lanosterol X-ray analysis. Increases the acidity of D455. <sup>9</sup>
W534 <sup>26</sup>	A, V, I, M, H, F, Y	The activities of all the mutants decreased significantly.	Steric bulk keeps the substrate at the correct position in the reaction cavity; hydrogen bonding between Y259 and W534 is unlikely, unlike Erg7 (H232:Y510 dyad).	Y510 <sup>29-34</sup>	Y510X	Y510F: <sup>33</sup> <b>II</b> and tricyclic iso-malabarica-triere skeleton <b>IV</b> . Y510K and Y510W: <sup>31</sup> <b>10</b> and <b>11</b> . Y510A: <sup>31</sup> <b>11</b> , <b>II</b> and <b>III</b> .	Hydrogen bonding with H234. The Y510:H-234 complex plays a key role in constructing the active-site structure for lanosterol scaffold biosynthesis. Located at C8-C9 (sterol numbering). This residue possibly abstracts H-9 proton.
E563 <sup>26</sup>	D, Q, N, I	Gln, Asn and Ile mutants gave negligible activity, while Asp mutant retained activity.  Initiation site.	Increases the acidity of D485.	E539 (not reported)	—	—	—
C564 <sup>12</sup>	A	Negligible activity. Initiation site.	Increases the acidity of D485.	C540 (not reported)	—	—	X-ray analysis: . <sup>9</sup> Increases the acidity of D455

				C533 <sup>9</sup> (human lanosterol)			
W612 <sup>26</sup>	A, V, I, M, H, F, Y	Aliphatic variants showed negligible activity. Aromatic mutants had high activity. A small amount of <b>28</b> was confirmed by W612Y and W612F variants. Located at A- and/or B-ring site.	Cation- $\pi$ complex with monocyclic cation.	W587 <sup>34</sup>	W587Y, W587F	<b>I, III, IV</b> , protostra-13(17),24-dien-3 $\beta$ -ol <b>VI</b> .	Stabilization of monocyclic and/or bicyclic cations via cation- $\pi$ interaction (CH- $\pi$ and/or $\pi$ - $\pi$ stacking interactions among Y707, Y99 and F699).
F728 <sup>22</sup>	A, I, M, H, Y, W	As major products, aliphatic mutants afforded tetracycle <b>21</b> , whereas the aromatic Tyr variant gave $\beta$ -amyryn <b>2</b> . Located at D/E-ring site.	Cation- $\pi$ interaction as the main function; the effect of steric size on cyclization is minimal.	F699 <sup>42-44</sup>	F699X (X=T, S, Q, C, Y, L, I, N, H, M)	Tricycles <b>IV</b> and <b>VIII</b> produced from a chair-boat-chair conformation. Tricycle <b>VII</b> were generated from chair-chair-chair structure, <sup>43</sup> and tetracycles <b>VI</b> , <b>IX</b> and <b>X</b> from intermediates bearing 17 $\alpha$ H and 17 $\beta$ H. <sup>43</sup>	Phe699 is located near C-17 of D-ring, but the function of Phe699 was discussed based on the combination of the Tyr99, Tyr707 and His234 residues. <sup>43</sup>
M729 <sup>26</sup>	G, A, V, L, F, W, H, N, Q	G, A mutants: <sup>43</sup> <b>28</b> (major); V, L, F variants: $\beta$ -amyryn <b>2</b> (major), E-ring site.	Steric bulk keeps the substrate at the correct position in the reaction cavity.	N700 (not reported)	—	—	—

L734 <sup>26</sup>	A, G, V, I, M, F, V, W	Taraxerol <b>30</b> from Gly, Ala, Val mutants, in addition to $\beta$ -amyrin <b>2</b> . Located at D-ring site.	Steric bulk keeps the substrate at the correct position in the reaction cavity. Possibly CH- $\pi$ interaction with Y736	I705 <sup>44</sup>	I705X	The following products were generated: <b>II, IV, VII, VI, VIII, IX, X, XI, XII</b> and <b>XIII</b> ,	The function of Ile705 was discussed based on the combination of Tyr99 and Tyr707 residues. Stabilization of the ring structures consisting of tricycle, and tetracycle. Located at C/D-ring site.
Y736 <sup>26</sup>	A, S, L, M, F, W	Phe variant produced only $\beta$ -amyrin <b>2</b> in high yield, but other mutants did not give aberrant cyclization products. Located at C-ring site.	Possibly CH- $\pi$ interaction with Y736. Phe mutant has high activity, possibly indicating that the effect of the OH group on polycyclization is not significant.	Y707 <sup>45</sup>	Y707X	<b>13, II, III</b> and <b>V</b> .	OH group of the Tyr is proximal to the C-8/C-9 cation, stabilizing the C-8/C-9 cation of bicyclic and protosteryl C-8/C-9 cation. Located B-ring site.

—: not reported.

Structure of the enzymatic products (**I–XI** and **10, 11** and **13**) were produced by Erg 7 mutants are shown in the Figure S2.

**Figure S2.** Structures of enzymatic products **I–XI** and **10**, **11** and **13**, which were generated by Erg 7 mutants. These products were reported by the Wu and the Matsuda groups. Structures of compounds **10**, **11** and **13** are identical to those of Figure 2 in the Text, including the compound numbering.

