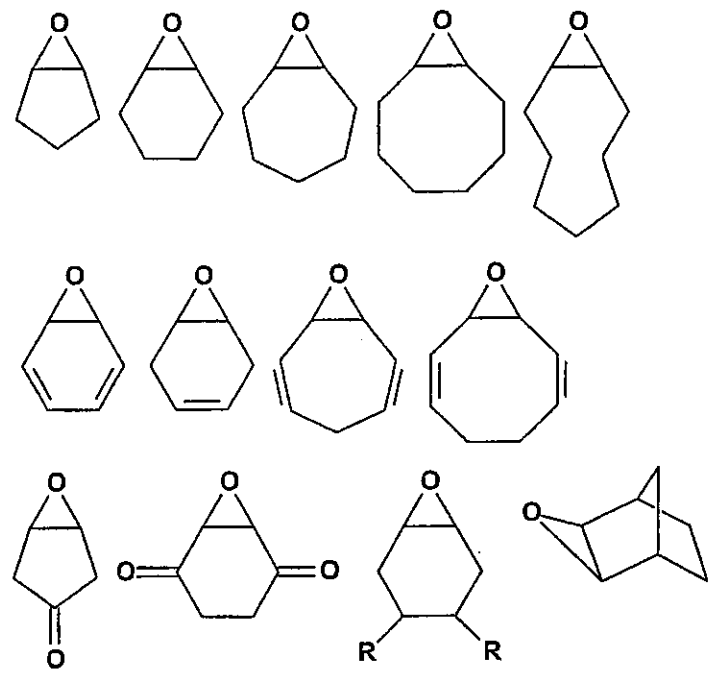



FIG 1

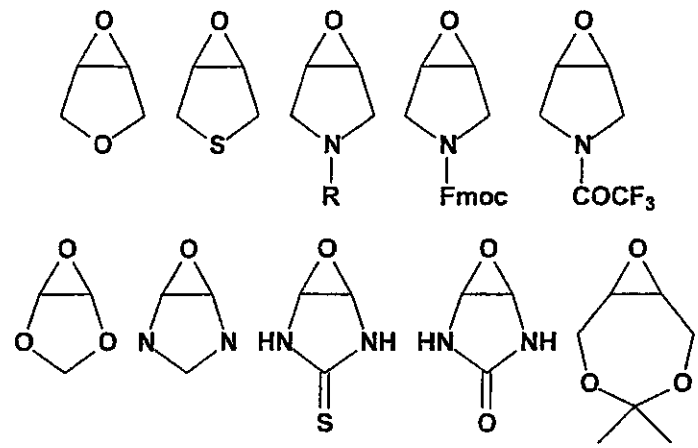



FIG 2

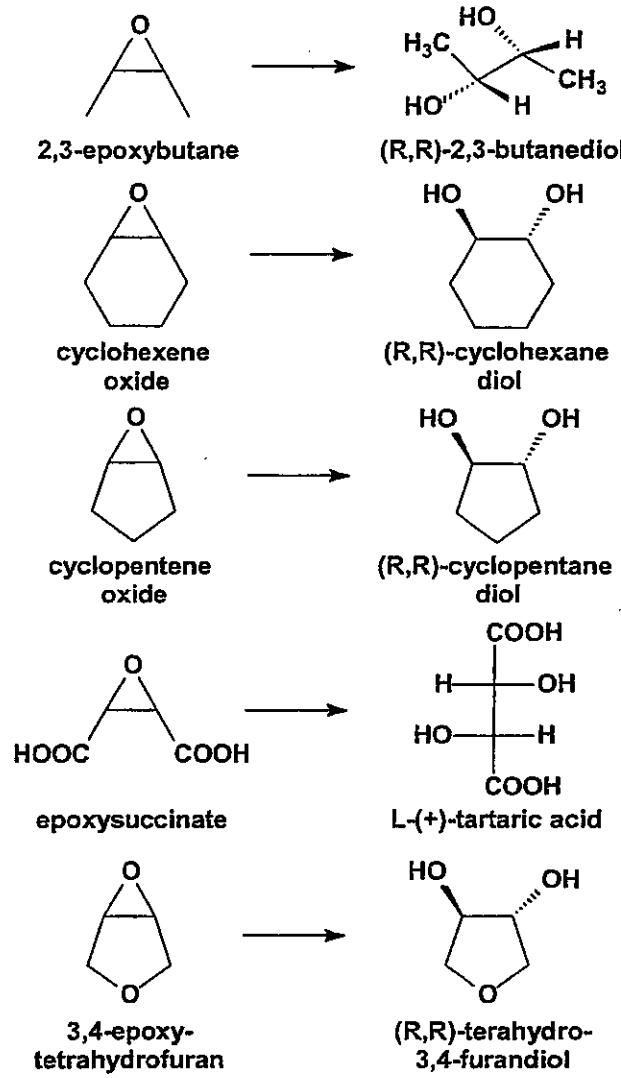

2,3-epoxybutane
(R,R)-2,3-butanediol
cyclohexene oxide
(R,R)-cyclohexane diol
cyclopentene oxide
(R,R)-cyclopentane diol
epoxysuccinate
L-(+)-tartaric acid
3,4-epoxy-tetrahydrofuran
(R,R)-terahydro-3,4-furandiol

FIG 3

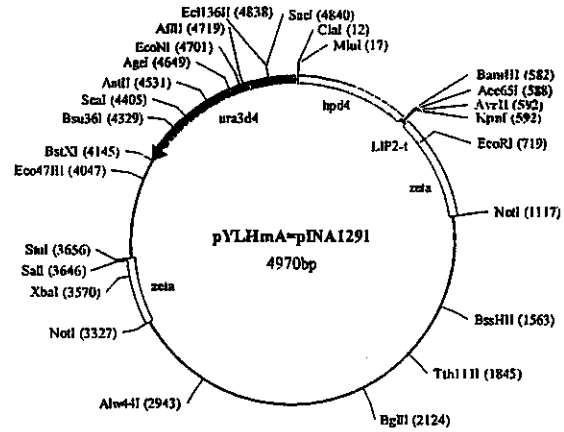

Ecl136II (4838)
SacI (4840)
AflII (4719)
ClaI (12)
EcoNI (4701)
MluI (17)
AgeI (4649)
AatII (4531)
ScaI (4405)
Bsu36I (4329)
ura3d4
hpd4
BamHI (582)
Acc65I (588)
AvrII (592)
KpnI (592)
LIP2-t
EcoRI (719)
BstXI (4145)
Eco47III (4047)
zeta
NotI (1117)
pYLHmA=pINA1291
4970bp
SalI (3656)
SalI (3646)
XbaI (3570)
zeta
NotI (3327)
BssHII (1563)
TthI11I (1845)
Alw44I (2943)
BglII (2124)

FIG 4

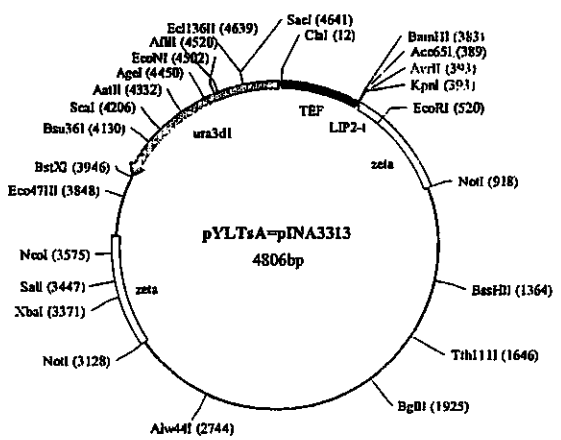

Ecl136II (4639)
SacI (4641)
AflII (4520)
ClaI (12)
EcoNI (4502)
BamHI (383)
Acc65I (389)
AgeI (4450)
AvrII (393)
AatII (4332)
KpnI (393)
ScaI (4206)
TEF
EcoRI (520)
Bsu36I (4130)
ura3d1
LIP2-t
BstXI (3946)
zeta
Eco47III (3848)
NotI (918)
pYLTsA=pINA3313
4806bp
NcoI (3575)
SalI (3447)
zeta
BssHII (1364)
XbaI (3371)
TthI11I (1646)
NotI (3128)
BglII (1925)
Alw44I (2744)

FIG 5

Sample 146

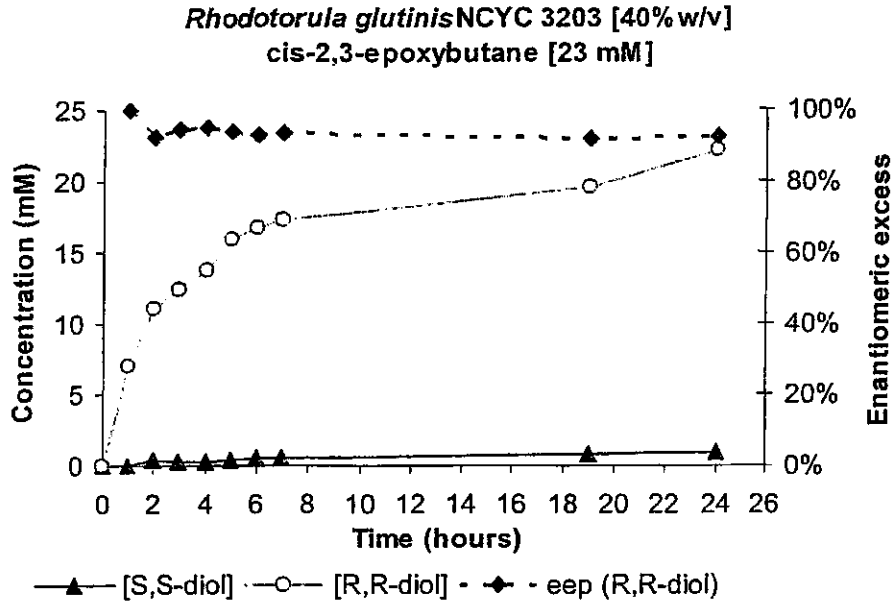



FIG 6

Samples 147

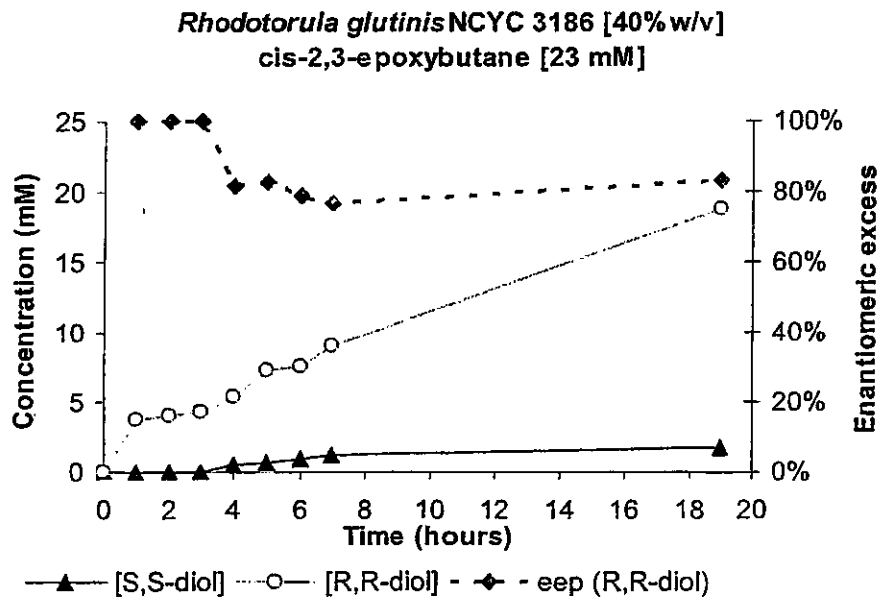



FIG 7

Samples 148

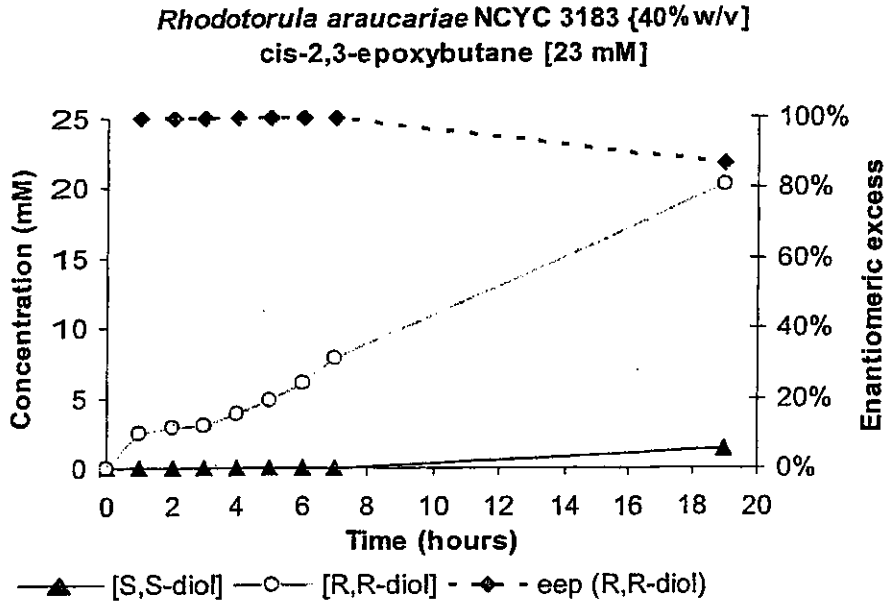



FIG 8

Sample 149

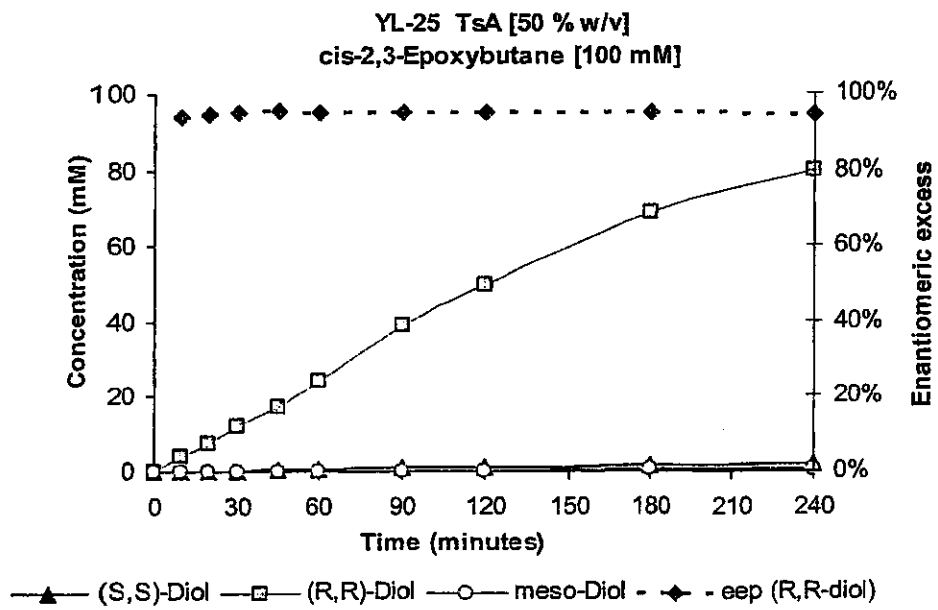



FIG 9

Sample 150

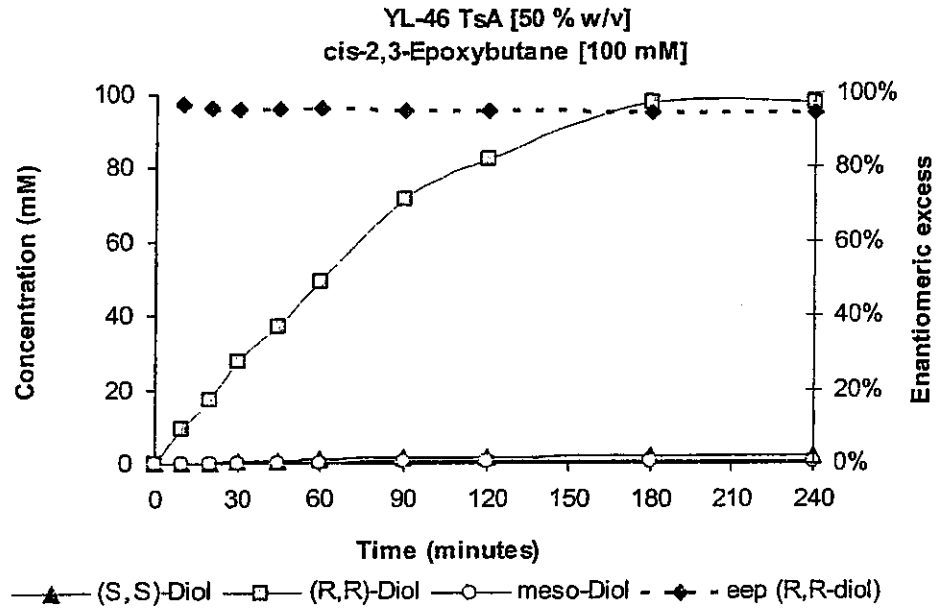



FIG 10

Sample 151

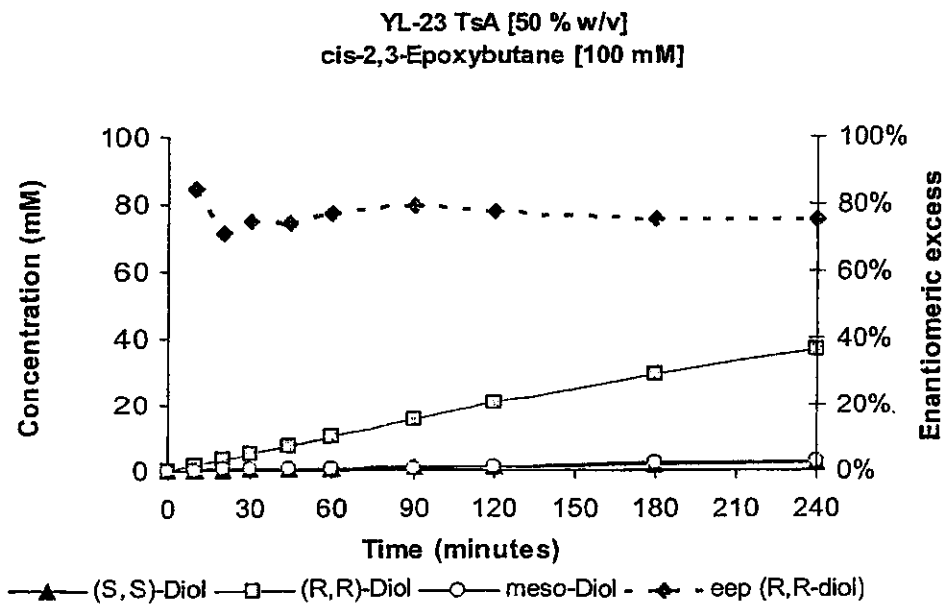



FIG 11

Sample 152

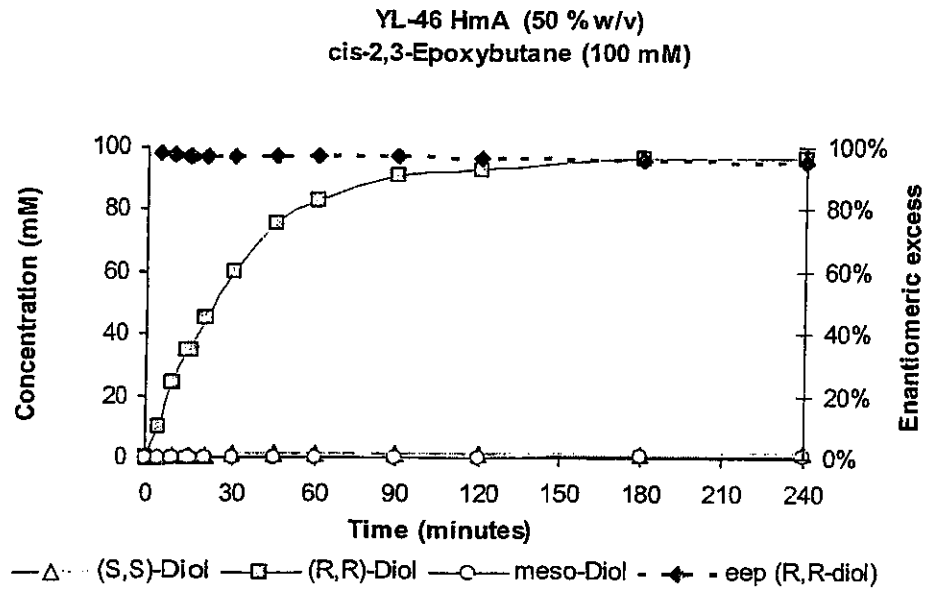



FIG 12

Sample 153

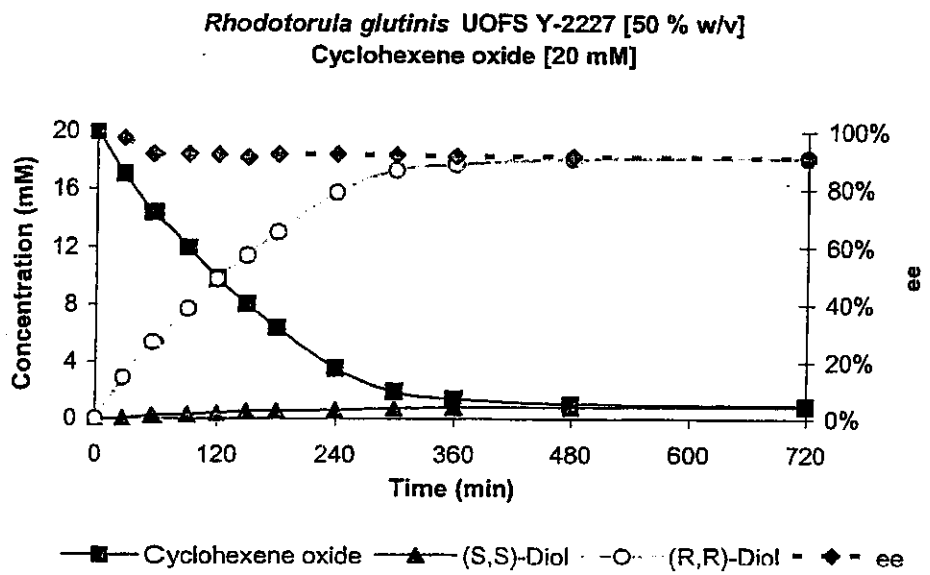



FIG 13

Sample 154

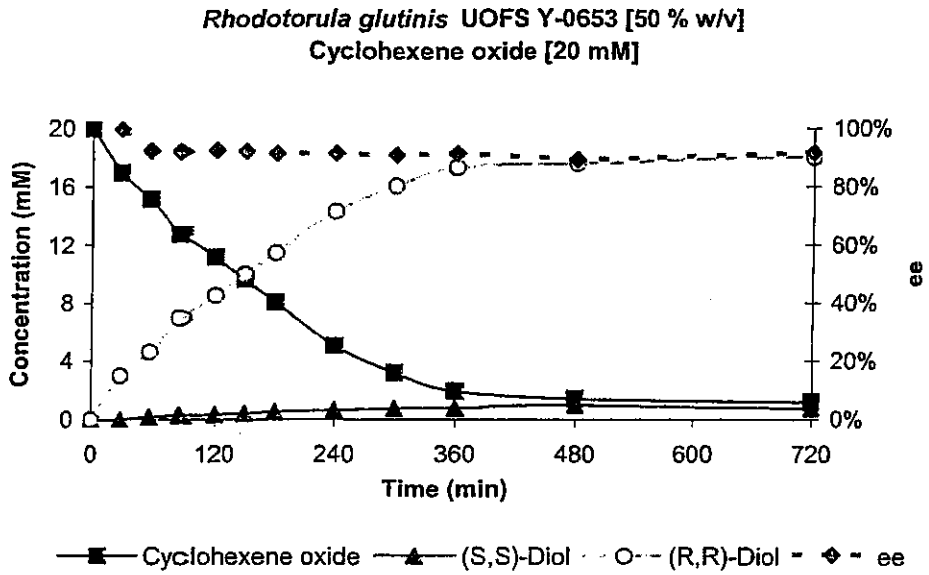



FIG 14

Sample 155

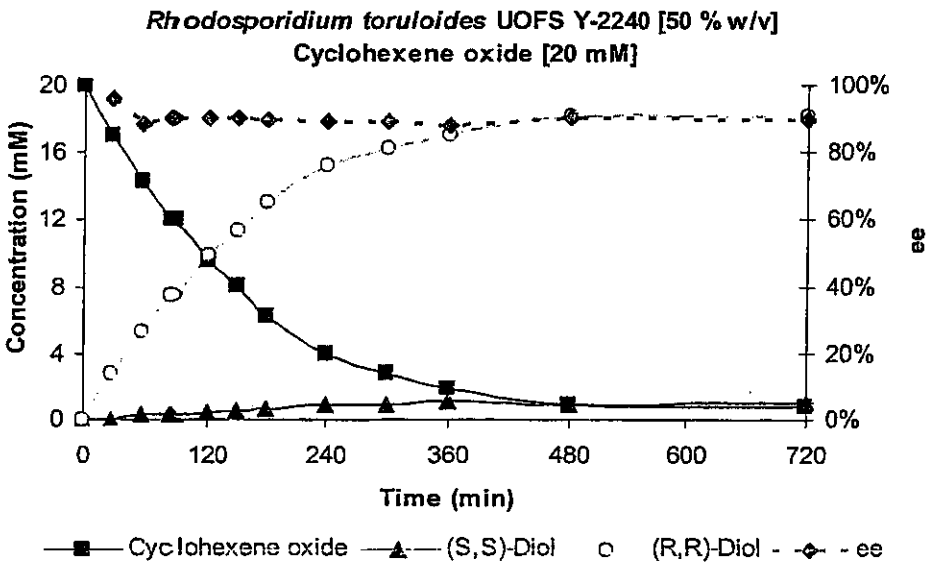



FIG 15

Sample 156

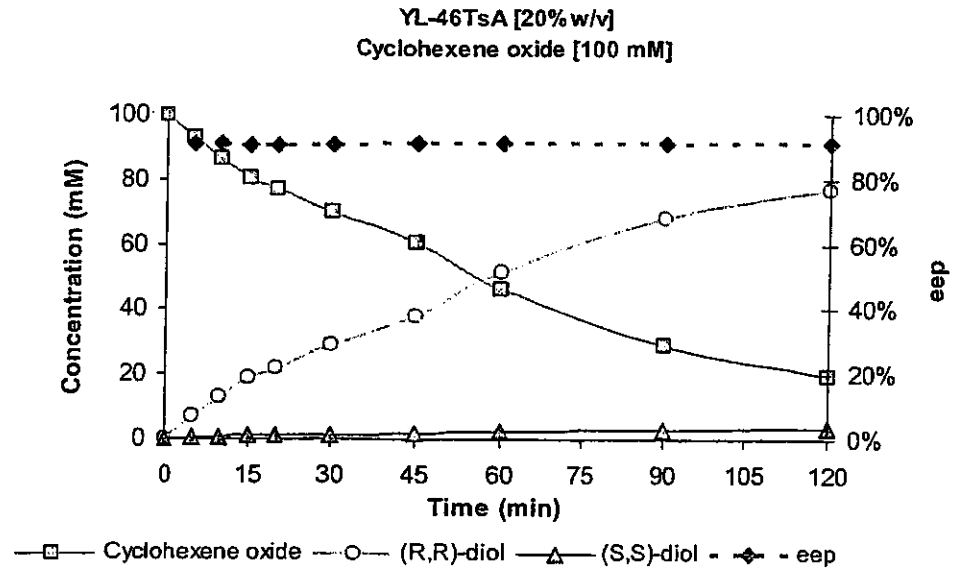



FIG 16

Sample 157

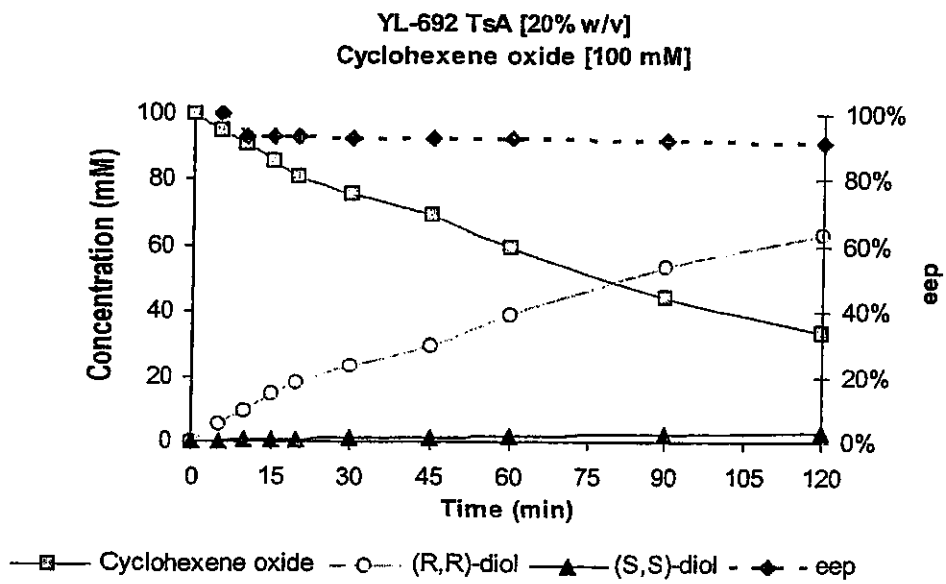



FIG 17

Sample 158

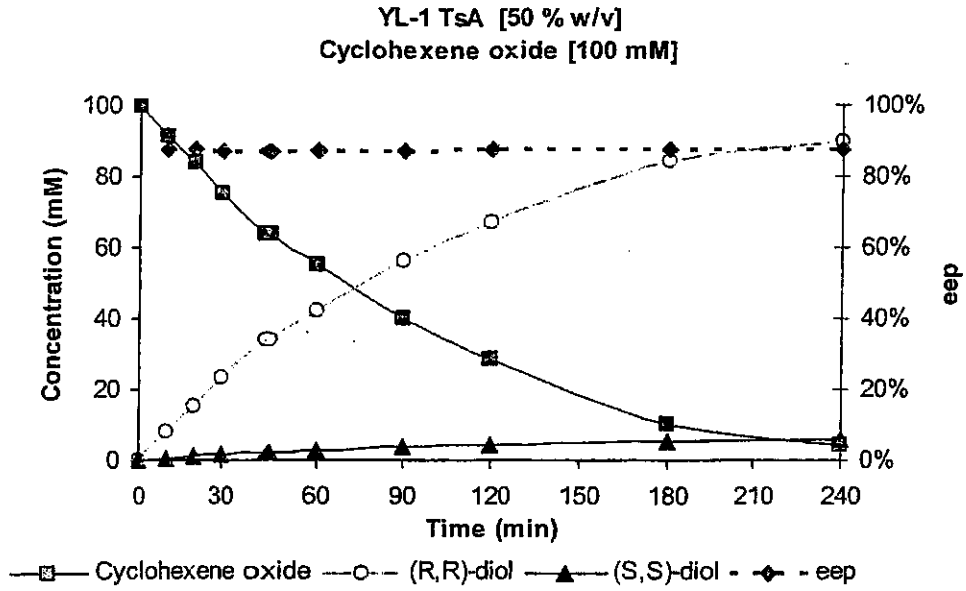



FIG 18

Sample 159

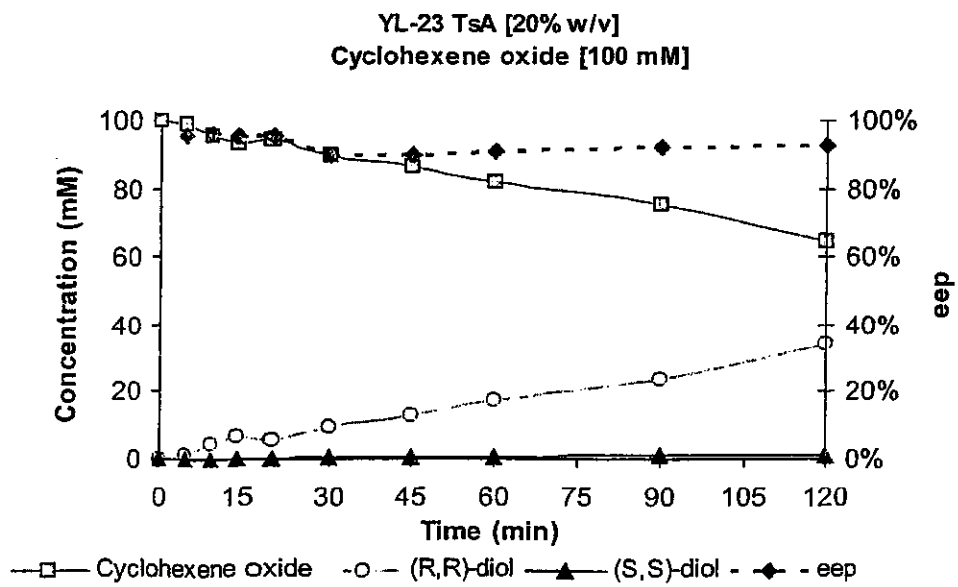



FIG 19

Sample 160

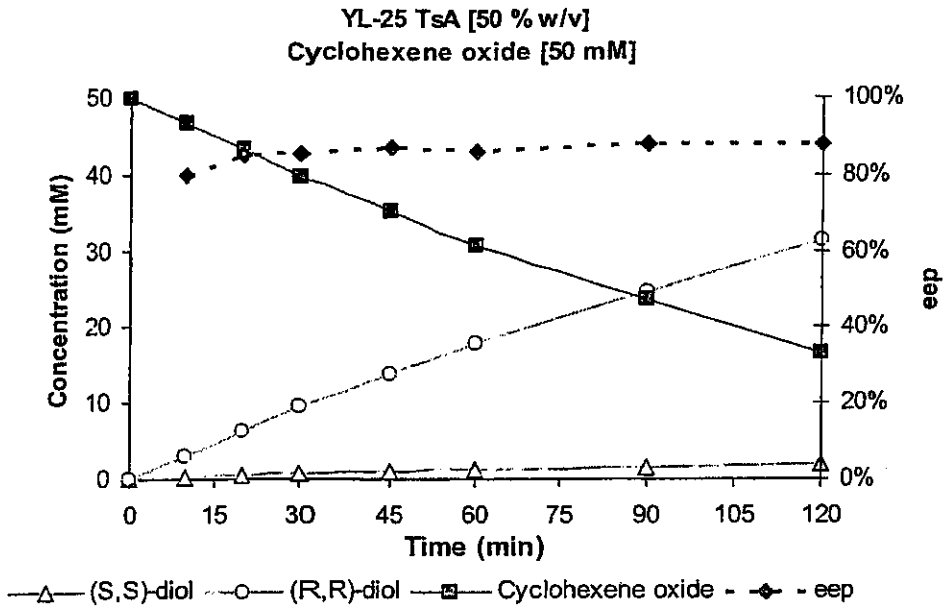



FIG 20

Sample 161

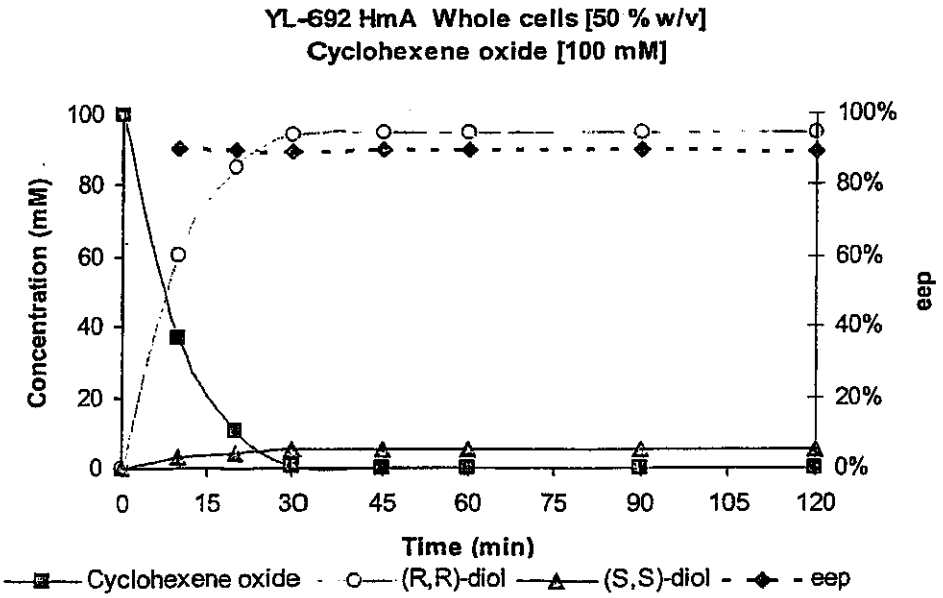



FIG 21

Sample 162

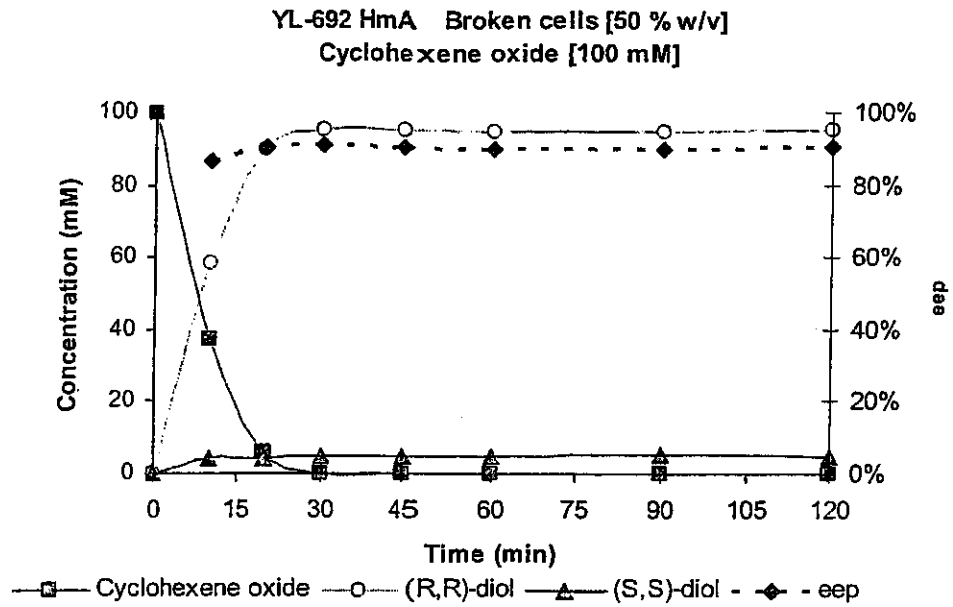



FIG 22

Sample 163

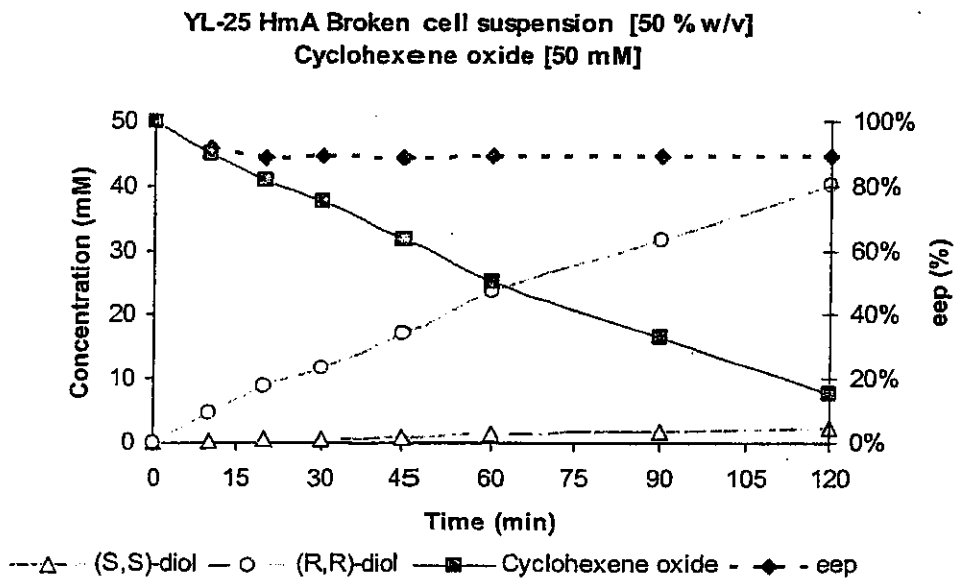



FIG 23

Sample 164

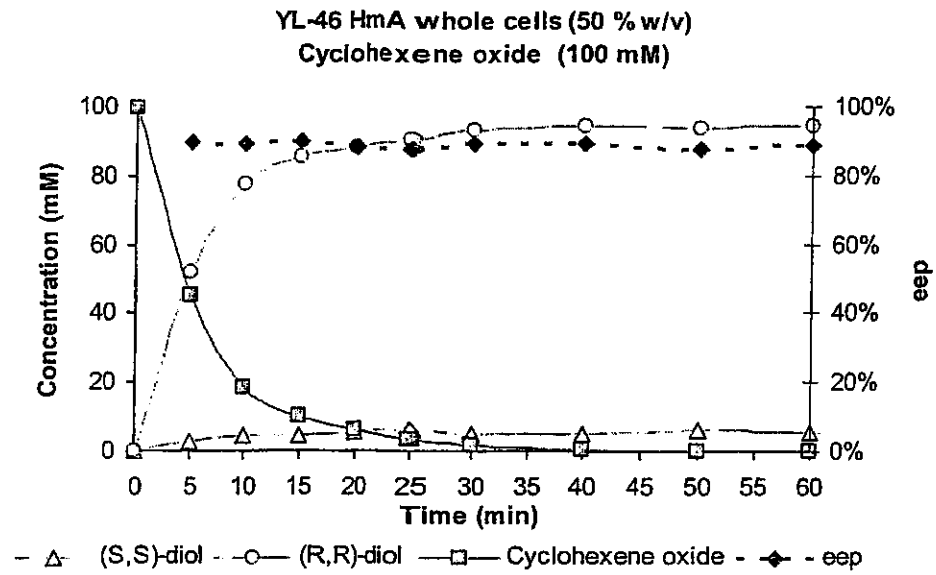



FIG 24

Sample 165

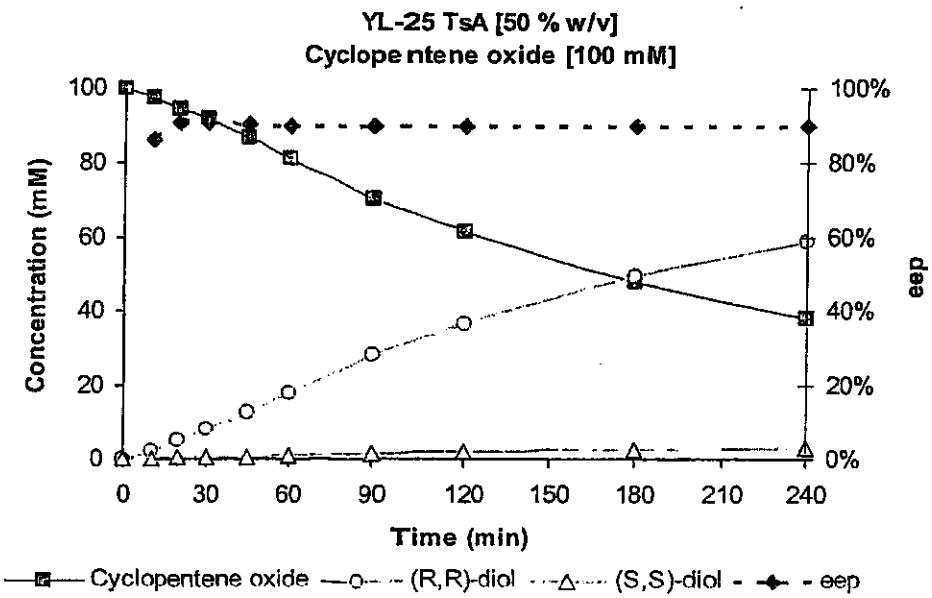



FIG 25

Sample 166

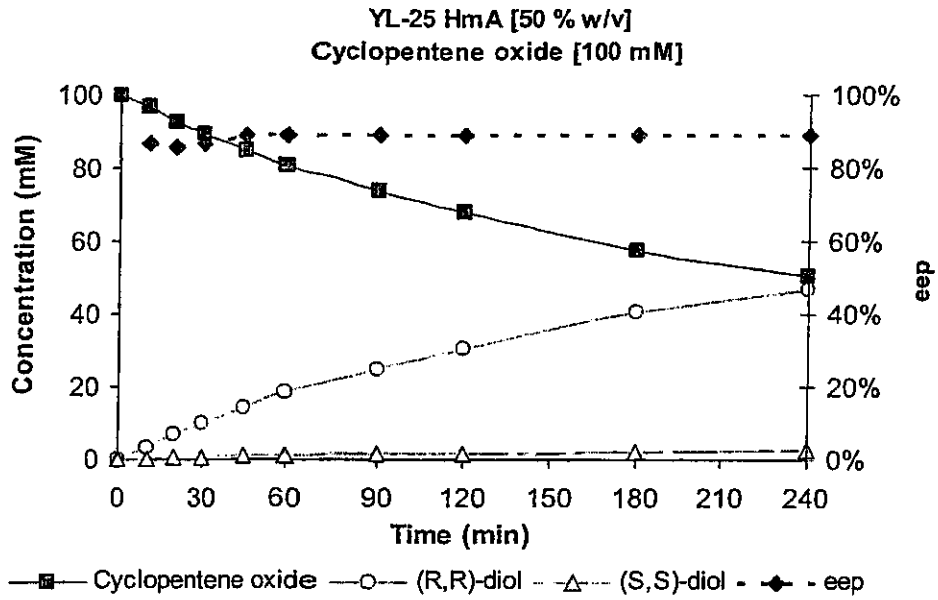



FIG 26

Sample 167

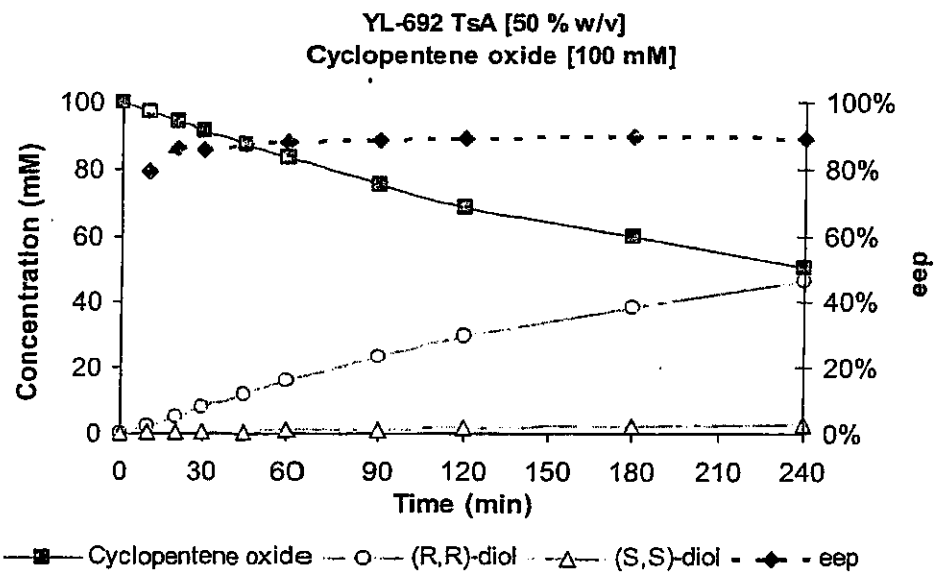



FIG 27

Sample 168

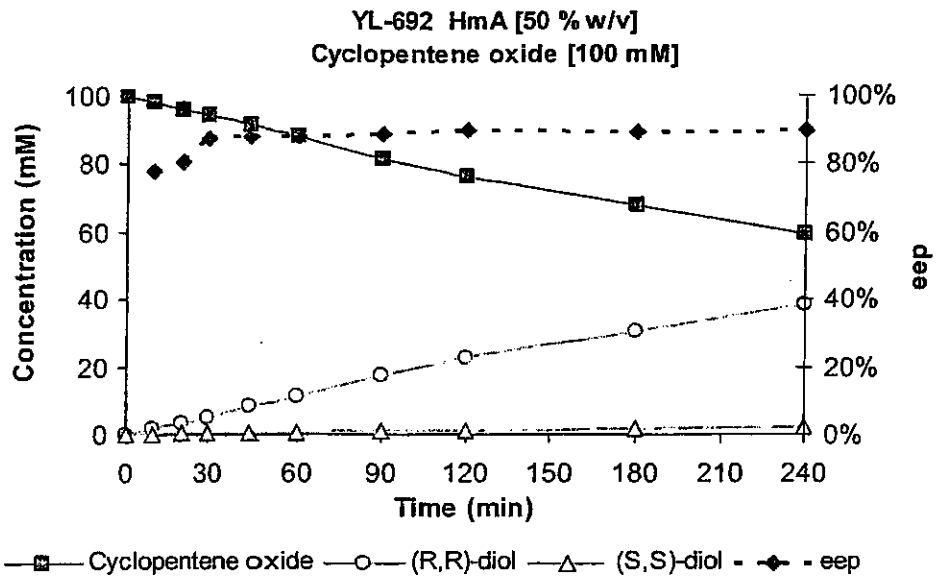



FIG 28

Sample 169

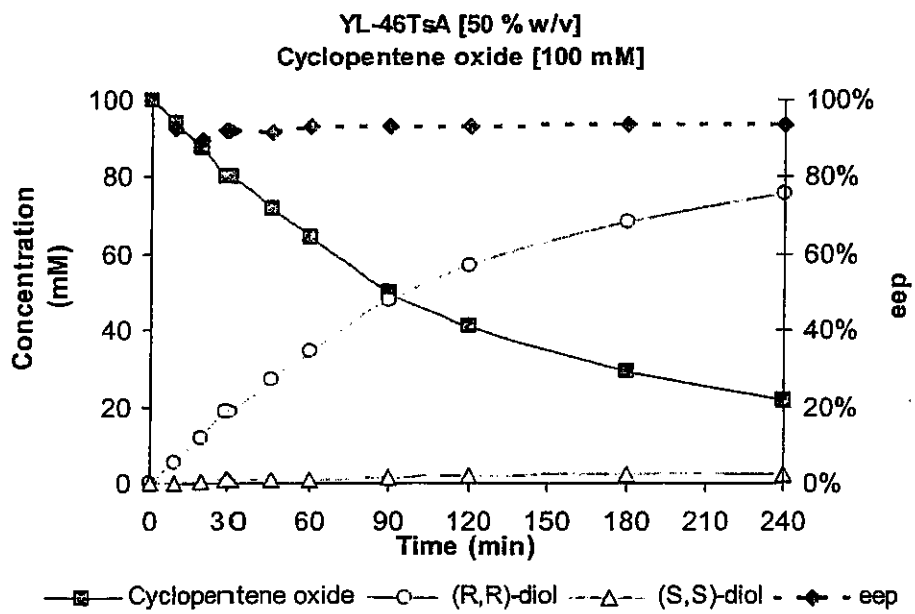



FIG 29

Sample 170

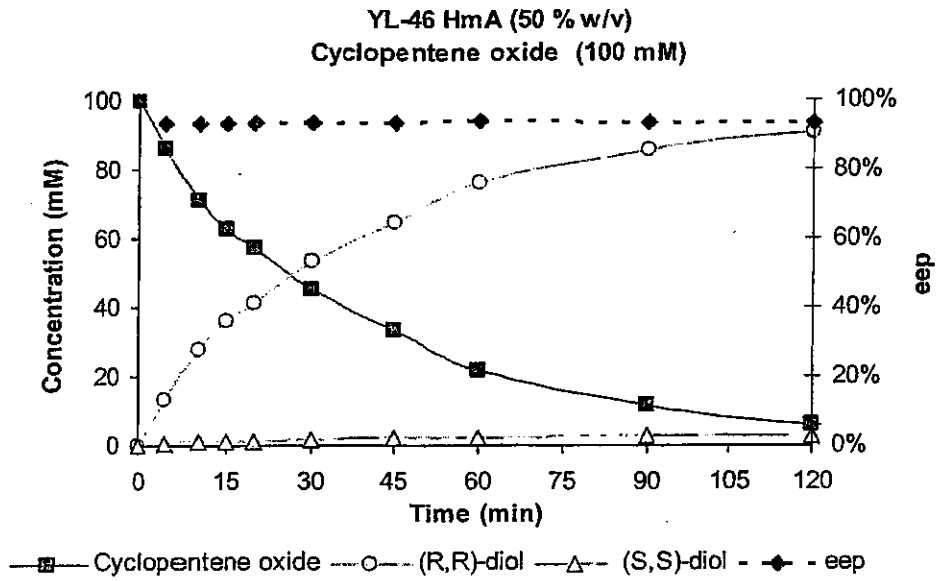



FIG 30

Sample 171

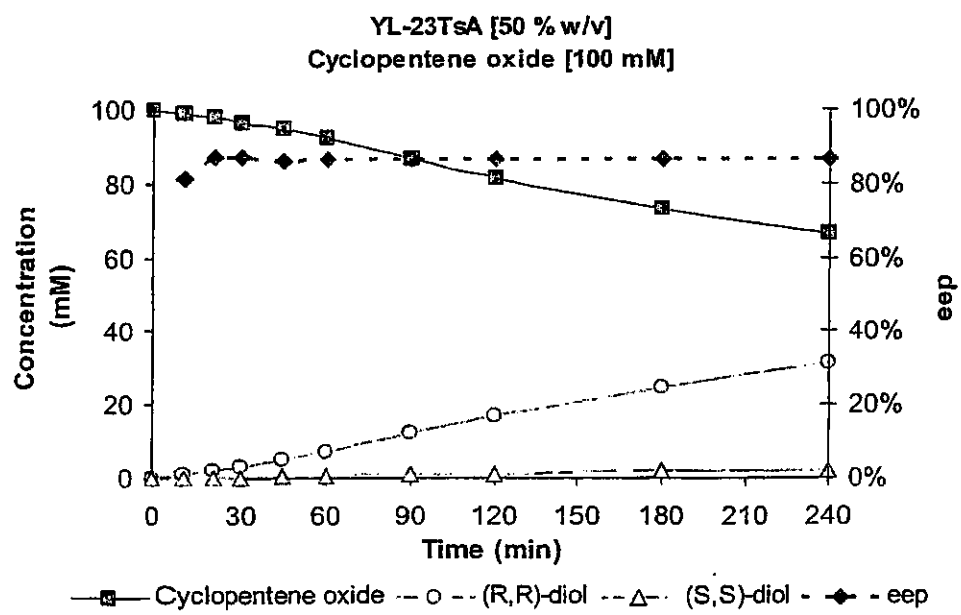



FIG 31

Sample 183

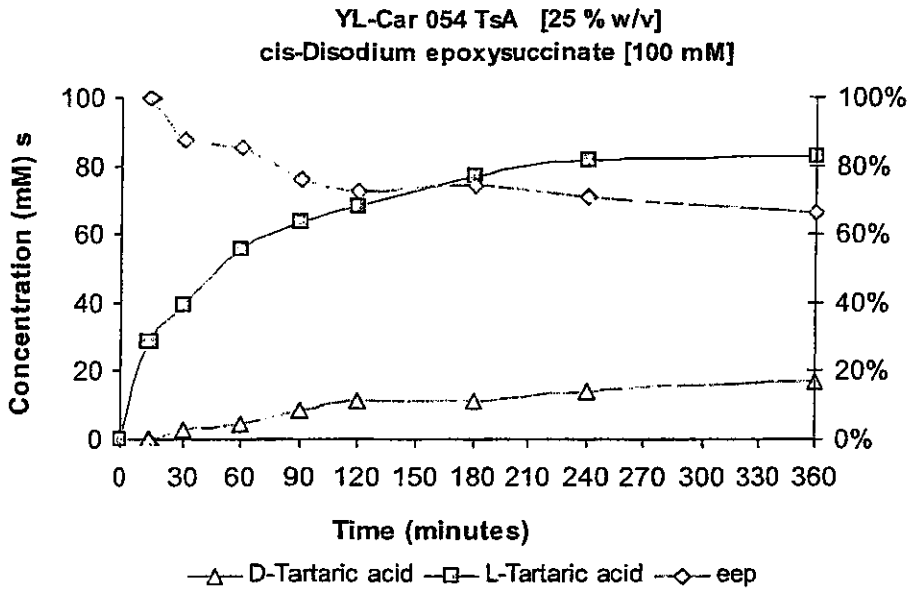



FIG 32

Sample 184

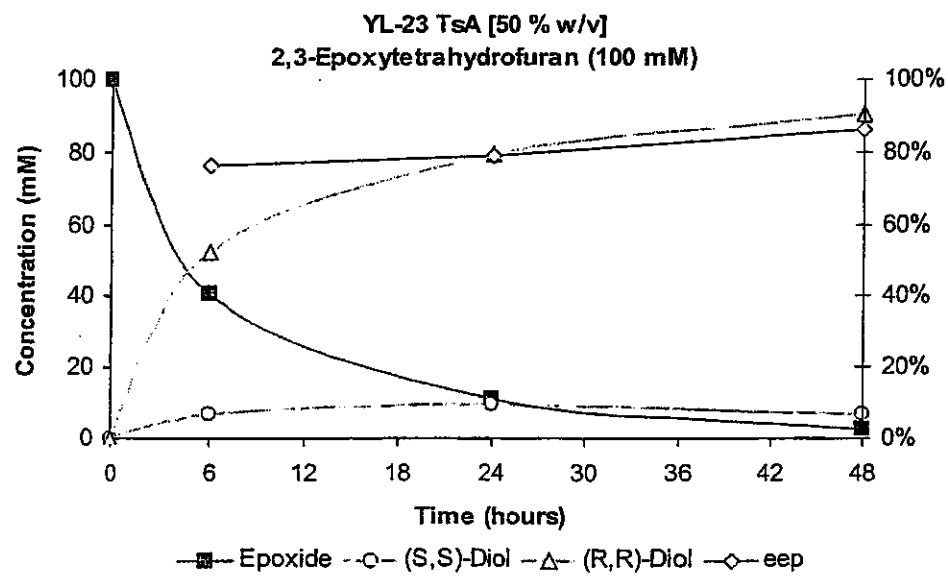



FIG 33

SEQ. ID. NO. 1

***Rhodosporidium toruloides NCYC 3181 (#1)***

Amino acid sequence

ORIGIN

```
1       MATHTFASPP TRFTVDIPQS ELDELDFRLD KTRWPATEIV PEDGADDPTA FGLGAGPTLP
61      LMKELAKGWR EFDWKKAQDH LNTFEHYTVE IEDLSIHFLH HRSTRPNAVP LILCHGWPGH
121     FGEFLNVIPL LTEPSDPSAQ AFHVVAPSMP GYAWSSPPPS SKWSMPDTAR VFDKLMTGLG
181     YEKYMAQGGD WGSIAARCLG SLHKDHCKAV HLNFLPVFPP VPMWLINPHT LLAWAPRFLV
241     PEKQAARMKR GLAYLEKGSA YYVMQQLTPR TPAYGLTDSP VGLLAWIGEK FEPTIQEASK
301     QAQPTLTRDE LYFTCSLYWF TRSIGTSFLP YSLNPHFTTF LTDSSTTCPT LPCPSTRARS
361     TARQNGTLSV PATSSDQGRA GGRTLCCARE ARRVCRPSQG GVRRHVGEV
```

**FIG 34**

SEQ. ID. NO. 2

***Rhodosporidium toruloides UOFS Y-0471 (#46)***

Amino acid sequence

ORIGIN

```
1       MKMATHTFAS PPTRFTVDIP QSELDELHSR LDKTRWPATE IVPEDGTDDP TAFGLGAGPT
61      LPLMKELAKG WREFDWKKAQ DHLNTFEHYM VEIEDLSIHF LHHRSTRPNA VPLILCHGWP
121     GHFGEFLNVI PLLTEPSDPS AQAFHVVAPS MPGYAWSLPP PSSKWNMPDT ARVFDKLMTG
181     LGYEKYMAQG GDWGSIAARC LGSLHKDHCK AVHLNFLPVF PPVPMWLINP HTLLAWAPRF
241     LVPEKQAARM KRGLAYLEKG SAYYVMQQLT PRTPAYGLTD SPVGLLAWIG EKFEPTIQEA
301     SKQAQPTLTR DELYFTCSLY WFTRSIGTSF LPYSLNPHFT TFLTDSKYHL PNFALSLYPG
361     EIYCPAERDA KRTGNLKWIK DAPEGGHFAA LEKPDVFVEH LREAFGVMWE K
```

**FIG 35**

SEQ. ID. NO. 3

***Rhodotorula araucariae NCYC 3183 (#25)***

Amino acid sequence

ORIGIN

```
1       MSEHSFEAPP QPFTVDFAPH IEDLHRRLDN ARWPTQEIVP VDVSETEHHN AFGLGMGPQL
61      NLMKELANGW RAFDQSALQD HLNSFNNWKV EIEGLSIHFL HHRSTRAGAL PLILCHGWPG
121     GYHEFLHVVQ LLTEPEGADA QAFHLVVPSM PGYAFSSPPP TAKWGMEDTA RVFDKLMTGL
181     GYNKYVAQGG DWGSITARCL GALHKDHCIA VHLNFCPVPP PFPFDQFNPR TLLNWMPRFV
241     ISDQQRAKLE RGLAYLEKGS AYYVMQQLTP RTPAYALNDS PIGLLAWIGE KMIPGINEAS
301     AQPNPTLNRD ALYTTLSLYW FTNSIGTSFL PYSLNPHFST FLTSPRYRLP NFALSSFPGE
361     LFSPTPRDAA RTGVMRWYKE ADDGGHFAAL EKPDVFSQHV REAVKAMLSS
```

**FIG 36**

SEQ. ID. NO. 4

***Cryptococcus curvatus NCYC 3158* (Car 054)**

Amino acid sequence

ORIGIN

```
1       MATHTFASPP TRFTVDIPQS EVDQLHSRLD KTRWPGTEIV PEDGADDPTA FGLGAGPTLP
61      LMKELAKGWR DFDWKKAQDH LNTFEHYTVE IEDLSIHFLH HRSTRPNAVP LILCHGWPGH
121     FGEFLHVIPL LTEPSDPSAQ AFHVVAPSMP GYAWSSPPPT SKWNMPDTAR VSDKLMNGLG
181     YEKYMAQGGD WGSIAARCLG ALHKDHCKAV HLNFLPVFPP VPMWLINPHT LLAWAPRFLV
241     PEKQAARMKR GLAYLEKGSA YYVMQQLTPR TSAYGLTDSP VGLLAWIGEK FEPTIQEASK
301     QAQPTLTRDE LYFTCSLYWF TRSIGTSFLP YSLNPHFTTF LTDSKYYLPN FALSLYPGEI
361     YCPAERDAKR TGNLKWIKDA PEGGHFAALE KPDVFVDHLR EAFGVMWEK
```

**FIG 37**

SEQ. ID. NO.5

***Rhodosporidium paludigenum* NCYC 3179 ( #692)**

Amino acid sequence

ORIGIN

```
1       MAAHSFTAPP APYNIDFAPQ VNDLHRRLDA ARWPEHDVVP DDVDHGEHGA FGLGAGPSLA
61      LMKELAQEWR GQDQKQLQDH LNSYKNYRVE IEGLNIHFLH YPSSRADAFP LILCHGWPGG
121     YHEFLHVLER LTEPEDQGSR AFHVVVPSMP GYAFSSPPKT AKWGMEDTAR VFDKLMTGLG
181     YAKYAAQGGD WGSITARCLG SLHKENCVAV HLNFCPVPPP FPLNMFNPRT LLDWMPRFVL
241     PDQRRAKIER GVAYIERGSA YYAMQNLTPR TPAYGLNDSP IGLLAWIGEK MIPGIDKAVK
301     HPNATLNREA LFTTLSIYWF TGSIGSSFLP YALNPHFSTF LVSPRHQLPN FALSNFPDEL
361     FTPEERDARR TGNLRWYKDA EDGGHFAALE KPEVFAEHVR EAMGVLLSNQ A
```

**FIG 38**

SEQ. ID. NO. 6

***Rhodotorula mucilaginosa* NCYC 3190 (#23)**

Amino acid sequence

ORIGIN

```
1       MPARSLTLRP FSPSFTAPEL DGLARSLESS RLPAETYASR QAKYGIKHAW MKNALQRWKD
61      GFDWKKHEQD INEVDHYMVQ VQSDGIQHDL HVIYHESKDP NAIPLLLLHG WPGSAFEFIE
121     AIKILRKSTS PAFHLIAPME PGYGWSTPPP LDRGFNMNDC TALMNDLMVG LGYGDGYAAQ
181     GGDIGSGLAR LLAVNYDACK CININYMPAV APPEDAPERH QIKPHEEDAL RRADEFQKTG
241     RGYANMHATR PGTVGIVVGS SPVALLAWIA EKYLAWTDED PPLDTILAIC TIWWIRDSYP
301     SSIWAYADFL ETGISALHND PKYKLDKKPF GFSSFKEEIS ATPEAWAGRN GNLQFYRYHD
361     KGGHFAALEQ PEAFAQDMQD CFGKIWPLSQ EQKS
```

**FIG 39**

SEQ. ID. NO. 7

***Yarrowia lipolytica* NCYC 3229 -2**

Amino acid sequence

ORIGIN

```
1       MSSLDLSYYE PFHRNAVLNG KRWHYVDIPG DSSGSLGRGK TLLLVHGFPD TWYGWRHQVP
61      VFRKHGFRLL IPSLLGFPRS EAPLTHPGVA TGEKFDGHNV HKELGLEDEN IQELECYTAD
121     FFAKSMVQLL DQLGIDKVCS FGHDWGAVFA PRLWLNYPER VECVSSACWY YQMPMEGFVD
181     LKDFAEVAPS LRYQLYWGGD APQEVIQKPL KEFMDRMYHH PTENHLLTSE AEYNNILKEF
241     QYGGKTIAPM FPLYKSRKVA YDIDERDFLT KGRTEEDLAV DVPYLFIGAE FDIALQPGME
301     SVLEGYVKEG LLEKQWVPCG HWVLFEEPEK MNKIYIDFLK KVFGSSKL
```

**FIG 40**

SEQ. ID. NO. 8

***Debaromyces hansenii NCYC 3167* (#113)**

Amino acid sequence

ORIGIN

```
1       MMQGQYKIQL RNGVRIFSTF SNYDNHDVFC SGGERKWTRV IFLLHGFPDN NSSYNSVWTV
61      ILDSVPEDEK VLLLAPSMRG YEPSSQGEQC DYKMSDLAGD VSSWIKNISP DGIPVHLVGH
121     DWGAIVAFKT ASMYPEMITS MACLAIPYIT NIRLWEFAWY FPDQIYHSSY MFTMQFGWLY
181     KKRLSDTSSN SYLDGLWRCW SPTWNYTVDE IDSVKRTLQE PGVIDASTAY YRCIANPLNI
241     SNRRWNVDFK KVPTLLLGGE RDGCMNEKLY RLEASKLANN PYVKVKTISN VGHFLHREDP
301     TKVGELITDW FSEYP*
```

**FIG 41**

SEQ. ID. NO. 9

***Filobasidium neoformans var neoformans* (#777) (GenBank XM 568708)**

Amino acid sequence

MSYSDLPHKP TIPVEPFKLS VPHEDLNGLL TLLKSTRIAK ESYENVSAHE NKFGITRKWL VNMKDEWIKQ
DWRKQEERIN SLPAFKAKVK NSDGSVFSIH FTALFSKKKV AIPIILSHGW PGSFYEFVPM MEMVKKKYSP
EDLPFHLIVP SLPGWLFSTP PPNDREFNVT DVGYLFNGLM EGLGFGDGYV AQGGDIGSYV TNELGAKYPA
CKIIHVNYSN PPPRPLPSPG SPGQEASPPS AEDLLELLQK FGYALEHSTR PATVGLVVGS NPLSLLAWVG
EKFLEWTDES PSEETILTMT SLYWFTDCFT TSIYTYRYGL GAKRHESAKQ ASYQKCPLGY SQFPKEIVEI
PAEWVKAQSN MVWSKKHESG GHFAALEKPE LLWADIEEFV HSQWENYKGN Y

**FIG 42**

SEQUENCE ID NO 10

***Rhodosporidium toruloides NCYC 3181 (#1)***

Nucleotide sequence

ORIGIN

```
1       ATGGCGACAC ACACATTCGC TTCGCCTCCC ACGCGCTTCA CCGTCGACAT CCCACAGTCA
61      GAACTCGACG AACTCGACTT CCGACTCGAC AAGACCCGCT GGCCGGCGAC AGAGATCGTT
121     CCAGAGGATG GGGCGGACGA CCCGACGGCG TTTGGGCTCG GAGCAGGGCC GACGCTGCCG
181     CTCATGAAGG AACTGGCGAA GGGTTGGCGC GAGTTCGACT GGAAGAAGGC GCAGGACCAC
241     CTCAACACCT TTGAGCACTA CACGGTCGAA ATCGAGGACC TCTCCATCCA CTTCCTCCAC
301     CACCGCTCGA CTCGCCCGAA TGCTGTTCCG CTCATCCTCT GCCACGGCTG GCCAGGCCAC
361     TTCGGCGAGT TCCTGAACGT CATACCGCTC TTGACGGAGC CGTCGGACCC GTCCGCCCAG
421     GCGTTCCACG TCGTCGCGCC TTCGATGCCC GGTTATGCTT GGTCTTCGCC TCCTCCGTCC
481     TCCAAGTGGA GCATGCCTGA TACCGCGAGG GTCTTCGACA AGCTCATGAC CGGGCTTGGC
541     TACGAGAAGT ACATGGCGCA GGGCGGAGAC TGGGGCAGCA TCGCTGCTCG CTGCCTTGGA
601     TCGCTTCACA AAGACCACTG CAAAGCCGTC CACCTCAACT TCCTCCCCGT CTTCCCACCC
661     GTCCCGATGT GGCTTATCAA CCCGCACACG CTCCTTGCCT GGGCACCGCG CTTCCTCGTG
721     CCGGAGAAGC AGGCTGCGCG TATGAAGCGC GGGTTGGCGT ACCTTGAGAA GGGCTCCGCC
781     TACTACGTCA TGCAGCAGTT GACGCCTCGC ACGCCTGCGT ACGGCCTGAC CGACAGTCCC
841     GTCGGCTTGC TGGCCTGGAT CGGCGAGAAG TTCGAGCCGA CCATTCAGGA GGCGAGCAAG
901     CAAGCCCAGC CGACCCTGAC TCGCGACGAG CTCTACTTCA CCTGCTCGCT CTACTGGTTC
961     ACCCGCTCAA TCGGCACCTC CTTCCTTCCC TACTCGCTCA ACCCGCACTT CACCACCTTC
1021    CTGACCGACA GCAGGTACCA CCTGCCCAAC TTTGCCCTGT CCCTCTACCC GGGCGAGATC
1081    TACTGCCCGG CAGAACGGGA CGCCAAGCGT ACCGGCAACC TCAAGTGGAT CAAGGACGCG
1141    CCCGATGGAG GACACTTTGC TGCGCTCGAG AAGCCCGATG TGTTTGTCGA GCATCTCAGG
1201    GAGGCGTTTG GCGTCATGTG GGAGAAGTAG
```

**FIG 43**

SEQUENCE ID 11

***Rhodosporidium toruloides UOFS Y-0471 (#46)***

Nucleotide sequence

ORIGIN

```
1       ATGGCGACAC ACACATTCGC TTCGCCTCCC ACCCGCTTCA CCGTCGACAT CCCACAGTCG
61      GAACTCGACG AACTTCACTC GCGACTCGAC AAGACCCGCT GGCCGGCGAC AGAGATCGTT
121     CCAGAGGATG GGACGGACGA TCCGACGGCG TTTGGGCTCG GAGCAGGGCC GACGCTGCCG
181     CTCATGAAGG AATTGGCGAA GGGTTGGCGC GAGTTCGACT GGAAAAAGGC GCAGGACCAC
241     CTCAACACCT TCGAGCACTA CATGGTCGAA ATTGAGGACC TCTCGATCCA CTTCCTCCAC
301     CATCGCTCGA CTCGCCCGAA CGCTGTTCCC CTCATCCTCT GCCACGGCTG GCCAGGCCAC
361     TTTGGCGAGT TCCTGAACGT TATCCCGCTC TTGACGGAGC CGTCGGACCC CTCCGCTCAG
421     GCGTTCCACG TCGTCGCCCC TTCGATGCCT GGCTATGCTT GGTCTTTGCC TCCTCCGTCC
481     TCCAAGTGGA ACATGCCTGA CACCGCGAGG GTCTTCGACA AGCTCATGAC CGGGCTTGGC
541     TACGAGAAGT ACATGGCGCA GGGCGGAGAC TGGGGAAGCA TCGCCGCTCG CTGCCTTGGA
601     TCGCTGCACA AGGACCATTG CAAAGCCGTC CACCTCAACT TCCTCCCCGT CTTCCCACCC
661     GTCCCGATGT GGCTTATCAA CCCGCACACG CTCCTTGCCT GGGCACCGCG CTTCCTCGTG
721     CCGGAGAAGC AGGCTGCGCG TATGAAGCGC GGGTTGGCGT ACCTTGAGAA GGGCTCCGCC
781     TACTACGTCA TGCAGCAGTT GACGCCTCGC ACGCCTGCGT ACGGCCTGAC CGACAGTCCC
841     GTCGGCTTGC TGGCCTGGAT CGGCGAGAAG TTCGAGCCGA CCATTCAGGA GGCGAGCAAG
901     CAAGCCCAGC CGACCCTGAC TCGCGACGAG CTCTACTTCA CCTGCTCGCT CTACTGGTTC
961     ACCCGCTCAA TCGGCACCTC CTTCCTTCCC TACTCGCTCA ACCCGCACTT CACCACCTTC
1021    CTGACCGACA GCAAGTACCA CCTGCCCAAC TTTGCCCTCT CGCTTTACCC AGGCGAGATC
1081    TACTGCCCCG CCGAGCGGGA CGCCAAGCGC ACCGGCAACC TCAAGTGGAT CAAGGACGCG
1141    CCTGAGGGAG GACACTTTGC TGCGCTCGAA AAGCCGGATG TGTTTGTCGA GCACCTCAGG
1201    GAGGCGTTTG GCGTCATGTG GGAGAAGTAG
```

**FIG 44**

SEQUENCE ID NO 12

***Rhodotorula araucariae NCYC 3183 (#25)***

Nucleotide sequence

ORIGIN

```
1       ATGAGCGAGC ACAGCTTCGA GGCCCCGCCA CAGCCGTTTA CGGTCGACTT TGCTCCCCAC
61      ATCGAGGATC TCCACCGCCG TCTCGACAAT GCGCGCTGGC CGACGCAAGA GATTGTCCCC
121     GTCGACGTGT CCGAGACGGA GCATCACAAC GCGTTCGGAC TCGGGATGGG CCCGCAGCTC
181     AACCTTATGA AGGAGCTCGC CAACGGCTGG CGCGCGTTCG ACCAGTCGGC GCTCCAGGAC
241     CACCTCAACA GCTTCAACAA CTGGAAGGTC GAGATCGAGG GATTGTCGAT CCACTTCCTC
301     CACCATCGCT CGACGCGCGC CGGCGCTCTC CCGCTCATCC TGTGCCATGG CTGGCCCGGC
361     GGGTACCACG AGTTCCTCCA CGTCGTCCAG CTCCTCACCG AACCAGAGGG GGCGGATGCG
421     CAGGCGTTTC ACCTCGTCGT CCCCTCGATG CCCGGGTACG CCTTCTCGTC TCCGCCGCCG
481     ACGGCCAAGT GGGGCATGGA AGACACTGCA AGGGTTTTTG ACAAGCTCAT GACCGGTTTG
541     GGGTACAACA AGTATGTCGC GCAGGGCGGT GACTGGGGGT CCATCACGGC GCGATGCCTC
601     GGCGCGCTGC ACAAGGACCA CTGCATTGCT GTCCACCTCA ACTTCTGCCC CGTCCCGCCG
661     CCGTTCCCAT TCGACCAGTT CAACCCGCGC ACGCTGCTCA ACTGGATGCC GCGCTTCGTG
721     ATCTCGGACC AGCAGCGTGC GAAGCTCGAG CGTGGGCTGG CGTACCTCGA GAAGGGGTCT
781     GCTTACTATG TCATGCAGCA GCTCACACCG CGTACCCCGG CCTACGCTCT CAATGACAGC
841     CCGATTGGCC TGCTCGCCTG GATTGGCGAA AAGATGATCC CAGGCATCAA CGAGGCGAGC
901     GCGCAGCCGA ACCCGACGCT CAATCGCGAT GCGTTGTACA CCACGCTCTC GCTGTACTGG
961     TTCACCAACT CCATCGGCAC CTCTTTCCTC CCCTACTCGC TTAACCCGCA CTTCAGCACG
1021    TTCCTCACCT CGCCCCGCTA TCGCCTGCCG AACTTTGCGC TGTCTTCCTT CCCGGGCGAG
1081    CTGTTCTCGC CGACGCCGCG CGATGCTGCG AGGACGGGCG TGATGCGCTG GTACAAGGAG
1141    GCGGACGATG GCGGGCACTT TGCGGCGCTC GAGAAGCCCG ATGTGTTCAG CCAGCATGTC
1201    AGGGAGGCAG TCAAGGCCAT GCTGTCGTCG TGA
```

**FIG 45**

SEQUENCE ID NO 13

***Cryptococcus curvatus NCYC 3158* (Car 054)**

Nucleotide sequence

ORIGIN

```
1       ATGGCGACAC ACACATTCGC TTCGCCTCCC ACCCGCTTCA CCGTCGACAT TCCGCAGTCG
61      GAAGTTGACC AACTTCACTC GCGACTCGAT AAGACTCGCT GGCCAGGGAC AGAGATCGTT
121     CCAGAGGATG GGGCGGACGA CCCGACGGCG TTTGGACTCG GAGCAGGACC AACGCTGCCG
181     CTCATGAAGG AACTGGCGAA GGGTTGGCGC GACTTCGACT GGAAGAAGGC GCAGGACCAC
241     CTCAACACCT TCGAGCACTA CACGGTGGAG ATCGAGGACC TCTCGATCCA CTTCCTCCAC
301     CATCGCTCGA CTCGTCCGAA TGCTGTTCCC CTCATCCTCT GCCACGGCTG GCCAGGACAC
361     TTCGGCGAGT TTCTACACGT TATACCACTC TTGACGGAGC CGTCGGACCC GTCCGCTCAG
421     GCGTTTCACG TCGTCGCGCC TTCGATGCCT GGTTATGCTT GGTCTTCGCC TCCTCCTACC
481     TCCAAGTGGA ATATGCCTGA CACCGCAAGG GTGTCCGACA AGCTTATGAA CGGACTCGGC
541     TACGAGAAGT ACATGGCGCA GGGCGGAGAC TGGGGCAGCA TCGCCGCTCG CTGCCTCGGA
601     GCGCTGCACA AAGATCACTG CAAAGCCGTC CACCTCAACT TCCTCCCCGT CTTCCCGCCC
661     GTCCCTATGT GGCTCATCAA CCCACACACA CTCCTCGCAT GGGCTCCGCG CTTCCTTGTG
721     CCGGAGAAGC AGGCTGCGCG CATGAAGCGC GGGTTGGCGT ACCTCGAAAA GGGCTCCGCC
781     TACTACGTCA TGCAGCAATT GACGCCTCGC ACGTCTGCGT ACGGCCTTAC TGACAGTCCC
841     GTCGGCTTGC TGGCCTGGAT CGGCGAGAAG TTTGAGCCGA CCATTCAGGA AGCGAGCAAG
901     CAAGCCCAGC CGACCCTCAC TCGCGACGAG CTCTACTTCA CCTGCTCCTT GTACTGGTTC
961     ACCCGCTCAA TCGGCACCTC CTTCCTGCCC TACTCGCTCA ACCCGCACTT CACGACCTTC
1021    CTGACCGACA GCAAGTATTA CCTGCCCAAT TTTGCCCTCT CCCTGTACCC GGGCGAGATC
1081    TACTGCCCTG CCGAGCGGGA CGCCAAGCGC ACTGGCAACC TCAAGTGGAT CAAGGACGCG
1141    CCTGAGGGAG GACACTTTGC GGCGCTCGAG AAGCCCGACG TGTTTGTCGA CCATCTCAGG
1201    GAGGCATTTG GCGTTATGTG GGAGAAGTAG
```

**FIG 46**

SEQUENCE ID NO 14

***Rhodosporidium paludigenum*** NCYC 3179 ( #692)

Nucleotide sequence

ORIGIN

```
1       ATGGCTGCCC ATTCCTTTAC TGCACCTCCT GCACCCTACA ACATCGACTT TGCGCCCCAG
61      GTAAATGACC TGCACCGCCG TCTCGATGCT GCCCGCTGGC CGGAACACGA CGTGGTGCCC
121     GACGATGTGG ATCACGGAGA GCACGGCGCA TTCGGACTCG GCGCTGGTCC CAGCCTCGCC
181     CTCATGAAGG AGCTCGCGCA GGAATGGAGG GGCCAGGACC AGAAGCAGCT GCAGGACCAC
241     CTCAACTCCT ACAAGAACTA TCGCGTCGAG ATCGAGGGTC TCAACATCCA CTTCCTGCAC
301     TACCCGTCGT CTCGCGCCGA TGCGTTCCCG CTCATCCTGT GCCACGGCTG GCCTGGCGGC
361     TACCACGAGT TCCTGCACGT CCTAGAGCGC CTCACGGAGC CCGAGGATCA GGGGTCGCGG
421     GCCTTCCATG TCGTCGTGCC TTCCATGCCG GGTTACGCCT TCTCCTCGCC GCCCAAGACG
481     GCAAAATGGG GCATGGAGGA CACGGCTCGC GTGTTCGACA AGCTCATGAC GGGGCTAGGT
541     TACGCCAAGT ATGCGGCCCA AGGCGGTGAC TGGGGGTCTA TCACGGCGCG CTGCCTAGGT
601     TCGCTGCACA AGGAGAACTG CGTCGCTGTC CACCTCAACT TCTGCCCGGT TCCCCCGCCG
661     TTCCCGCTCA ACATGTTCAA CCCGCGCACA CTTCTGGACT GGATGCCTCG CTTTGTCCTG
721     CCTGATCAAC GGCGGGCCAA GATTGAGCGC GGCGTGGCCT ATATCGAGCG CGGCTCTGCC
781     TACTACGCCA TGCAAAACTT GACGCCGCGC ACGCCTGCGT ACGGCTTGAA CGATAGTCCG
841     ATTGGTTTGC TCGCGTGGAT TGGCGAGAAG ATGATTCCGG GCATTGACAA GGCTGTCAAG
901     CATCCGAACG CAACCCTCAA TCGCGAAGCT CTTTTCACGA CACTCTCGAT CTACTGGTTC
961     ACGGGCTCGA TTGGCTCCTC CTTCCTGCCA TACGCTCTCA ACCCGCACTT CTCTACCTTC
1021    CTCGTCTCGC CGCGGCACCA ACTGCCGAAC TTTGCTCTGT CCAACTTTCC CGACGAGCTG
1081    TTCACGCCCG AAGAACGCGA TGCTCGCCGA ACCGGAAACT TGCGGTGGTA CAAGGATGCA
1141    GAGGATGGAG GGCACTTCGC GGCGCTGGAG AAGCCCGAGG TCTTCGCCGA GCACGTAAGG
1201    GAGGCGATGG GGGTCTTGCT GTCGAACCAG GCCTGA
```

**FIG 47**

SEQUENCE ID NO 15

***Rhodotorula mucilaginosa* NCYC 3190 (#23)**

Nucleotide sequence

ORIGIN

```
1       ATGCCCGCCC GCTCGCTCAC GCTGCGCCCG TTCTCGCCGT CGTTCACGGC TCCGGAACTG
61      GACGGTCTCG CTCGCTCGCT CGAGTCGTCG CGCTTGCCCG CCGAGACGTA CGCTTCCCGC
121     CAGGCCAAAT ACGGCATCAA GCATGCTTGG ATGAAGAATG CCCTCCAACG GTGGAAGGAC
181     GGGTTCGATT GGAAGAAGCA CGAGCAGGAC ATCAACGAGG TCGACCACTA TATGGTGCAG
241     GTCCAGTCCG ATGGCATTCA ACACGACCTC CATGTGATCT ATCACGAATC GAAAGACCCG
301     AATGCGATCC CGCTCTTGCT GCTGCATGGG TGGCCCGGTT CCGCGTTCGA GTTTATCGAG
361     GCGATCAAGA TCCTTCGCAA GAGTACCTCG CCCGCGTTCC ACCTGATCGC GCCCATGGAG
421     CCCGGCTACG GGTGGAGTAC TCCGCCGCCA CTCGACCGCG GTTTCAACAT GAACGATTGC
481     ACAGCGCTCA TGAACGACTT GATGGTTGGA CTCGGGTACG GAGACGGTTA CGCTGCTCAG
541     GGTGGCGACA TCGGTTCGGG ACTCGCAAGA CTCCTCGCCG TCAACTATGA CGCATGCAAA
601     TGCATCAACA TCAACTACAT GCCTGCCGTT GCACCACCAG AGGACGCTCC GGAGCGGCAC
661     CAGATCAAAC CGCACGAGGA GGATGCGCTC CGACGTGCGG ACGAGTTTCA GAAGACGGGC
721     AGGGGGTATG CCAACATGCA TGCAACGAGA CCCGGTACGG TCGGCATCGT CGTCGGTAGT
781     TCGCCGGTCG CACTGCTCGC TTGGATCGCT GAAAAGTACC TCGCTTGGAC CGATGAGGAT
841     CCGCCCCTCG ACACGATCCT CGCAATCTGC ACCATCTGGT GGATCCGCGA CTCGTACCCT
901     TCTTCAATCT GGGCCTACGC CGACTTTCTC GAGACGGGCA TCTCGGCCCT GCACAACGAC
961     CCGAAGTACA AACTTGACAA GAAACCGTTC GGGTTCTCGA GCTTCAAGGA GGAGATCAGC
1021    GCGACTCCCG AGGCGTGGGC GGGCAGGAAC GGCAACTTGC AGTTCTATCG GTACCACGAC
1081    AAGGGAGGTC ACTTTGCGGC GCTCGAGCAG CCGGAAGCGT TCGCGCAAGA CATGCAGGAT
1141    TGCTTCGGCA AAATCTGGCC TCTCTCTCAG GAGCAAAAAT CGTAG
```

**FIG 48**

SEQUENCE ID NO 16

***Yarrowia lipolytica* NCYC 3229-2**

Nucleotide sequence

ORIGIN

```
1       ATGTCATCAC TCGACCTATC TTATTACGAA CCGTTCCACA GAAATGCCGT TCTTAATGGC
61      AAAAGATGGC ACTATGTAGA CATTCCTGGA GACTCCAGTG GCTCTCTTGG TAGAGGAAAG
121     ACACTTTTGC TGGTCCATGG GTTCCCAGAC ACTTGGTACG GGTGGCGGCA TCAAGTGCCT
181     GTGTTCCGTA AGCACGGATT TCGGCTCTTG ATTCCCTCCT TGCTTGGTTT CCCAAGAAGT
241     GAAGCTCCTC TTACTCACCC TGGAGTTGCC ACAGGAGAGA AGTTCGATGG TCACAATGTT
301     CACAAGGAGC TGGGTCTGGA AGATGAAAAC ATTCAGGAAC TTGAATGTTA CACGGCTGAC
361     TTCTTTGCCA AGAGCATGGT CCAACTGCTG GACCAATTGG GGATCGACAA GGTCTGTTCC
421     TTTGGACATG ACTGGGGAGC TGTCTTTGCT CCCCGACTCT GGCTCAATTA TCCTGAGAGA
481     GTCGAATGTG TGTCATCAGC ATGCTGGTAC TACCAAATGC CAATGGAGGG CTTCGTCGAC
541     TTGAAAGATT TCGCAGAGGT GGCTCCTTCA TTGAGATACC AATTGTACTG GGGAGGAAGAC
601     GCTCCTCAAG AAGTCATCCA GAAGCCGCTC AAGGAATTTA TGGACAGAAT GTACCACCAT
661     CCTACAGAGA ATCATCTTCT GACATCGGAA GCTGAATACA ACAACATTCT CAAGGACGTTT
721     CAATATGGAG GAAAGACCAT TGCTCCCATG TTTCCTCTCT ACAAGTCACG AAAGGTAGCC
781     TACGATATCG ATGAAAGGGA CTTTCTGACC AAGGGGAGAA CAGAGGAAGA CTTGGCGGTG
841     GATGTGCCGT ACTTGTTTAT CGGGGCCGAG TTTGACATTG CTCTTCAGCC CGGAATGGAG
901     TCTGTTTTGG AGGGTTATGT GAAGGAGGGA CTACTAGAGA AGCAATGGGT CCCCTGTTGGA
961     CATTGGGTCT TGTTTGAGGA GCCTGAGAAG ATGAACAAGA TTTATATTGA CTTTCTCGAAG
1021    AAGGTGTTTG GATCGTCGAA GTTGTAG
```

**FIG 49**

SEQ. ID. NO. 17

***Debaromyces hansenii NCYC 3167* (#113)**

Nucleotide sequence

ORIGIN

```
1     ATGATGCAAG GTCAATACAA GATTCAACTT CGCAACGGAG TAAGAATATT TTCTACCTTC
61    AGCAATTATG ATAACCACGA TGTATTTTGT AGTGGTGGTG AGCGTAAATG GACTCGAGTT
121   ATATTCCTCC TCCATGGGTT TCCCGATAAT AACAGTTCAT ATAATAGCGT TTGGACAGTG
181   ATATTAGATT CTGTCCCAGA AGATGAAAAG GTATTGTTGC TAGCGCCATC AATGAGAGGT
241   TACGAACCGT CTAGTCAAGG TGAACAATGT GATTATAAGA TGTCAGATCT TGCGGGCGAT
301   GTAAGCTCAT GGATTAAAAA TATTTCACCA GATGGTATTC CAGTACACCT TGTTGGTCAT
361   GATTGGGGAG CGATAGTGGC TTTTAAGACC GCTAGTATGT ACCCTGAGAT GATAACATCA
421   ATGGCTTGTT TGGCCATTCC ATATATCACC AATATTCGAT TATGGGAGTT TGCATGGTAT
481   TTTCCAGACC AGATTTATCA TTCCAGCTAT ATGTTCACCA TGCAGTTTGG ATGGCTCTAT
541   AAAAAGAGGC TTTCTGATAC GAGTAGCAAC AGTTATTTGG ACGGATTATG GCGCTGCTGG
601   TCTCCGACAT GGAATTATAC CGTAGATGAA ATCGATAGTG TTAAGAGAAC ACTTCAAGAA
661   CCTGGAGTTA TCGATGCATC AACTGCATAC TACAGATGTA TTGCTAACCC CCTAAACATT
721   TCTAATAGGC GCTGGAACGT AGACTTTAAA AAAGTTCCAA CGTTGTTGCT CGGCGGAGAG
781   CGGGATGGGT GCATGAATGA AAAGTTGTAT CGTTTAGAAG CCAGTAAGCT TGCAAATAAT
841   CCATATGTTA AGGTCAAGAC AATTTCAAAT GTCGGGCATT TCTTGCATCG TGAAGACCCT
901   ACAAAAGTTG GTGAATTGAT AACTGACTGG TTCAGTGAAT ATCCTTAG
```

**FIG 50**

SEQ. ID. NO. 18

***Filobasidium neoformans var neoformans (#777)***

Nucleotide sequence

ORIGIN

```
1       ATGTCGTATT CAGACCTTCC CCACAAGCCG ACCATTCCTG TCGAGCCCT T CAAACTATCC
61      GTACCCCATG AGGACCTCAA TGGGCTCCTC ACTCTACTCA AATCTACCA G AATTGCCAAA
121     GAAAGCTATG AGAATGTCTC TGCGCACGAG AACAAATTTG GAATAACGA G GAAATGGCTG
181     GTGAATATGA AAGATGAATG GATCAAACAG GACTGGCGGA AACAAGAAG A GCGAATCAAT
241     TCTCTTCCTG CGTTCAAGGC AAAGGTAAAA AACTCAGATG GTTCGGTAT T TTCAATCCAT
301     TTTACCGCCC TTTTCTCGAA AAAGAAGGTC GCAATCCCTA TCATCCTCA G TCACGGCTGG
361     CCAGGTAGCT TTTACGAATT CGTTCCAATG ATGGAGATGG TGAAGAAGA A ATATAGTCCT
421     GAGGATCTTC CGTTCCATTT AATCGTACCA TCCCTTCCCG GATGGCTCT T CTCTACACCT
481     CCTCCTAATG ATCGAGAATT CAATGTGACC GATGTTGGGT ATCTCTTCA A TGGTCTAATG
541     GAGGGCCTAG GTTTCGGAGA CGGCTATGTT GCCCAAGGCG GTGATATTG G GAGCTACGTG
601     ACGAACGAAC TTGGTGCCAA GTACCCAGCT TGCAAAATCA TTCACGTCA A TTACTCTAAC
661     CCTCCTCCTC GCCCACTTCC CTCTCCAGGC TCTCCTGGAC AAGAGGCTT C ACCGCCCTCT
721     GCTGAAGATC TTCTTGAGCT ACTGCAGAAG TTTGGATATG CTCTCGAGC A TTCTACTCGA
781     CCAGCCACTG TAGGCTTGGT AGTAGGCTCC AATCCACTCA GCTTACTAG C ATGGGTCGGC
841     GAGAAATTCC TCGAGTGGAC TGATGAGTCC CCAAGTGAGG AAACAATTT T GACAATGACA
901     AGCCTTTATT GGTTCACTGA TTGTTTCACG ACGTCGATTT ACACATACA G ATATGGCCTT
961     GGTGCTAAAC GCCATGAAAG TGCCAAACAA GCTTCTTACC AGAAATGTC C TCTGGGCTAC
1021    AGCCAGTTCC CCAAAGAGAT CGTCGAGATA CCTGCCGAGT GGGTCAAGG C GCAATCCAAT
1081    ATGGTTTGGT CTAAGAAACA TGAATCCGGG GGTCATTTTG CTGCTCTCG A AAAGCCAGAA
1141    TTGTTGTGGG CGGATATAGA GGAGTTTGTA CACTCGCAGT GGGAGAATT A CAAAGGTAAT
1201    TACTGA
```

**FIG 51**

Alignments (homology at amino acid level)

| Strain | NCYC 3181 (#1) | NCYC 3183 (#25) | UOFS Y-0471 (#46) | NCYC 3179 (#692) | NCYC 3158 (Car54) | NCYC 3190 (#23) | NCYC 3229-2 (Jen46-2) | NCYC 3167 (#113) | #777 |
|---|---|---|---|---|---|---|---|---|---|
| NCYC 3181 (#1) | | 68.37 | 98.04 | 63.11 | 95.84 | 29.95 | 10.51 | 12.17 | 24.23 |
| NCYC 3183 (#25) | 68.37 | | 67.64 | 74.51 | 67.64 | 26.42 | 13.27 | 12.38 | 26.95 |
| UOFS Y-0471 (#46) | 98.04 | 67.64 | | 62.62 | 96.33 | 30.66 | 9.54 | 12.65 | 27.63 |
| NCYC 3179 (#692) | 63.11 | 74.51 | 62.62 | | 62.14 | 27.53 | 11.58 | 13.43 | 25.77 |
| NCYC 3158 (Car 054) | 95.84 | 67.64 | 96.33 | 62.14 | | 29.72 | 10.51 | 12.17 | 26.64 |
| NCYC 3190 (#23) | 29.95 | 26.42 | 30.66 | 27.53 | 29.72 | | 13.30 | 12.91 | 36.67 |
| NCYC 3229 (Jen46-2) | 10.51 | 13.27 | 9.54 | 11.58 | 10.51 | 13.30 | | 16.85 | 12.84 |
| NCYC 3167 (#113) | 12.17 | 12.38 | 12.65 | 13.43 | 12.17 | 12.91 | 16.85 | | 10.47 |
| Filobasidium #777 | 24.23 | 26.95 | 27.63 | 25.77 | 26.64 | 36.67 | 12.84 | 10.47 | |

*Alignments done in DNAMAN Version 5.2.9 (Lynnon Biosoft) – multiple sequence alignment mode (Fast alignment/Quick alignment settings)*

**FIG 52**

## Alignments (homology at nucleic acid level)

| Strain | NCYC 3181 (#1) | NCYC 3183 (#25) | UOFS Y-0471 (#46) | NCYC 3179 (#692) | NCYC 3158 (Car54) | NCYC 3190 (#23) | NCYC 3229-2 (Jen462) | NCYC 3167 (#113) | #777 |
|---|---|---|---|---|---|---|---|---|---|
| NCYC 3181 (#1) | | 69.58 | 96.18 | 67.88 | 92.36 | 39.98 | 26.25 | 21.59 | 34.28 |
| NCYC 3183 (#25) | 69.58 | | 68.69 | 72.32 | 67.72 | 41.47 | 23.36 | 23.37 | 34.65 |
| UOFS Y-0471 (#46) | 96.18 | 68.69 | | 66.34 | 92.60 | 40.34 | 19.81 | 23.11 | 38.63 |
| NCYC 3179 (#692) | 67.88 | 72.32 | 66.34 | | 66.10 | 39.18 | 25.87 | 22.34 | 35.02 |
| NCYC 3158 (Car 054) | 92.36 | 67.72 | 92.60 | 66.10 | | 41.41 | 24.85 | 21.14 | 36.35 |
| NCYC 3190 (#23) | 21.59 | 23.37 | 23.11 | 22.34 | 21.14 | | 18.73 | 22.11 | 44.28 |
| NCYC 3229 (Jen46-2) | 26.25 | 23.36 | 19.81 | 25.87 | 24.85 | 18.73 | | 18.66 | 28.88 |
| NCYC 3167 (#113) | 21.59 | 23.37 | 23.11 | 22.34 | 21.14 | 22.11 | 18.66 | | 25.25 |
| Filobasidium #777 | 34.28 | 34.65 | 38.63 | 35.02 | 36.35 | 25.25 | 28.88 | 25.25 | |

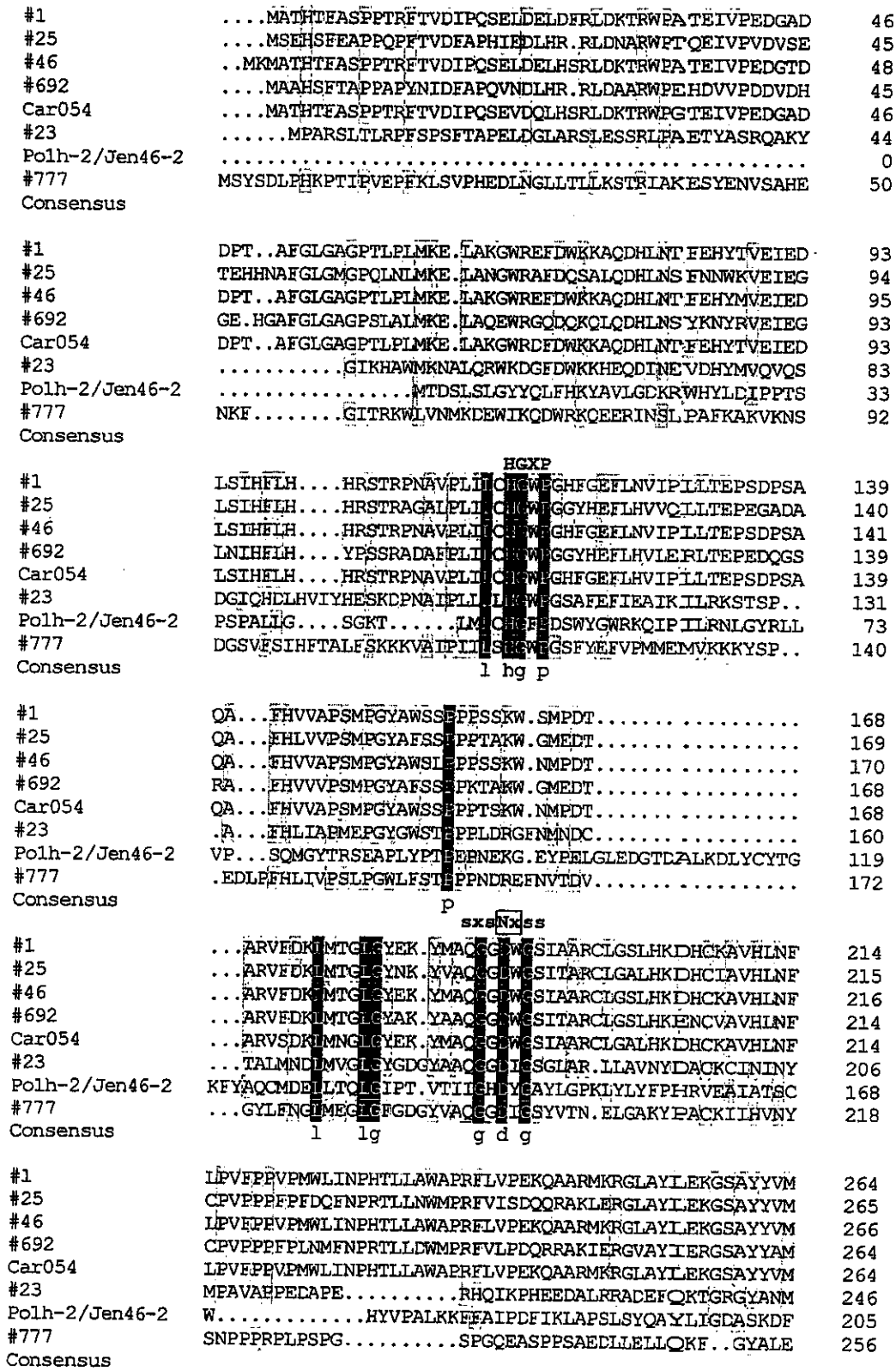
*Alignments done in DNAMAN Version 5.2.9 (Lynnon Biosoft) – multiple sequence alignment mode (Fast alignment/Quick alignment settings)*

FIG 53



Amino acid alignments of EH proteins, indicating conserved sequence motifs and regions surrounding the catalytic triad. The nucleophile, acid and base of the catalytic triad are indicated by N, A and B, respectively. HGXP represents the region of the oxy-anion hole of the enzyme. sxNxss represents the genetic motif found in α/β-hydrolase fold enzymes. (P = 100% homology; * = identity of 75-100%; # = identity of 50-75%; - = gap).

**FIG 54**


#1 ....MATHTFASPPTRFTVDIPQSELDELDFRLDKTRWPATEIVPEDGAD 46
#25 ....MSEHSFEAPPQPFTVDFAPHIEDLHR.RLDNARWPTQEIVPVDVSE 45
#46 ..MKMATHTFASPPTRFTVDIPQSELDELHSRLDKTRWPATEIVPEDGTD 48
#692 ....MAAHSFTAPPAPYNIDFAPQVNDLHR.RLDAARWPEHDVVPDDVDH 45
Car054 ....MATHTFASPPTRFTVDIPQSEVDQLHSRLDKTRWPGTEIVPEDGAD 46
#23 ......MPARSLTLRPFSPSFTAPELDGLARSLESSRLPAETYASRQAKY 44
Polh-2/Jen46-2 ..... 0
#777 MSYSDLPHKPTIPVEPFKLSVPHEDLNGLLTLLKSTRIAKESYENVSAHE 50
Consensus
#1 DPT..AFGLGAGPTLPLMKE.LAKGWREFDWKKAQDHLNTFEHYTVEIED 93
#25 TEHHNAFGLGMGPQLNLMKE.LANGWRAFDQSALQDHLNSFNNWKVEIEG 94
#46 DPT..AFGLGAGPTLPLMKE.LAKGWREFDWKKAQDHLNTFEHYMVEIED 95
#692 GE.HGAFGLGAGPSLALMKE.LAQEWRGQDQKQLQDHLNSYKNYRVEIEG 93
Car054 DPT..AFGLGAGPTLPLMKE.LAKGWRDFDWKKAQDHLNTFEHYTVEIED 93
#23 .....GIKHAWMKNALQRWKDGFDWKKHEQDINEVDHYMVQVQS 83
Polh-2/Jen46-2 .....MTDSLSLGYYQLFHKYAVLGDKRWHYLDIPPTS 33
#777 NKF.....GITRKWLVNMKDEWIKQDWRKQEERINSLPAFKAKVKNS 92
Consensus
HGXP
#1 LSIHFLH....HRSTRPNAVPLLICHGWPGHFGEFLNVIPLLTEPSDPSA 139
#25 LSIHFLH....HRSTRAGALPLLCHGWPGGYHEFLHVVQLLTEPEGADA 140
#46 LSIHFLH....HRSTRPNAVPLLCHGWPGHFGEFLNVIPLLTEPSDPSA 141
#692 LNIHFLH....YPSSRADAFPLLCHGWPGGYHEFLHVLERLTEPEDQGS 139
Car054 LSIHFLH....HRSTRPNAVPLLCHGWPGHFGEFLHVIPLLTEPSDPSA 139
#23 DGIQHLLHVIYHESKDPNAIPLLLHGWPGSAFEFIEAIKILRKSTSP.. 131
Polh-2/Jen46-2 PSPALLG....SGKT.....LMLCHGFDSWYGWRKQIPILRNLGYRLL 73
#777 DGSVFSIHFTALFSKKKVAIPLLLSHGWPGSFYEFVPMMEMVKKKYSP.. 140
Consensus
l hg p
#1 QA...FHVVAPSMPGYAWSSPPPSSKW.SMPDT..... 168
#25 QA...FHLVVPSMPGYAFSSPPPTAKW.GMEDT..... 169
#46 QA...FHVVAPSMPGYAWSLPPPSSKW.NMPDT..... 170
#692 RA...FHVVVPSMPGYAFSSPPKTAKW.GMEDT..... 168
Car054 QA...FHVVAPSMPGYAWSSPPPTSKW.NMPDT..... 168
#23 .A...FHLIAPMEPGYGWSTPPPLDRGFNMNDC..... 160
Polh-2/Jen46-2 VP...SQMGYTRSEAPLYPTPPPNEKG.EYPELGLEDGTDALKDLYCYTG 119
#777 .EDLPFHLIVPSLPGWLFSTPPPNDREFNVTDV..... 172
Consensus
p
sxsNxss
#1 ...ARVFDKLMTGLGYEK.YMAQGGDWGSIAARCLGSLHKDHCKAVHLNF 214
#25 ...ARVFDKLMTGLGYNK.YVAQGGDWGSITARCLGALHKDHCIAVHLNF 215
#46 ...ARVFDKLMTGLGYEK.YMAQGGDWGSIAARCLGSLHKDHCKAVHLNF 216
#692 ...ARVFDKLMTGLGYAK.YAAQGGDWGSITARCLGSLHKENCVAVHLNF 214
Car054 ...ARVSDKLMNGLGYEK.YMAQGGDWGSIAARCLGALHKDHCKAVHLNF 214
#23 ...TALMNDLMVGLGYGDGYAAQGGDIGSGLAR.LLAVNYDACKCININY 206
Polh-2/Jen46-2 KFYAQCMDELLTQLGIPT.VTIIGHDYGAYLGPKLYLYFPHRVEAIATSC 168
#777 ...GYLFNGLMEGLGFGDGYVAQGGDIGSYVTN.ELGAKYPACKIIHVNY 218
Consensus
l lg g d g
#1 LPVFPPVPMWLINPHTLLAWAPRFLVPEKQAARMKRGLAYLEKGSAYYVM 264
#25 CPVPPPFPFDQFNPRTLLNWMPRFVISDQQRAKLERGLAYLEKGSAYYVM 265
#46 LPVFPPVPMWLINPHTLLAWAPRFLVPEKQAARMKRGLAYLEKGSAYYVM 266
#692 CPVPPPFPLNMFNPRTLLDWMPRFVLPDQRRAKIERGVAYIERGSAYYAM 264
Car054 LPVFPPVPMWLINPHTLLAWAPRFLVPEKQAARMKRGLAYLEKGSAYYVM 264
#23 MPAVAPPEDAPE.....RHQIKPHEEDALRRADEFQKTGRGYANM 246
Polh-2/Jen46-2 W.....HYVPALKKFFAIPDFIKLAPSLSYQAYLIGDASKDF 205
#777 SNPPPRPLPSPG.....SPGQEASPPSAEDLLELLQKF..GYALE 256
Consensus