SEQUENCE LISTING

<110> THE REGENTS OF THE UNIVERSITY OF MICHIGAN

<120> BIOCATALYST AND METHODS FOR SYNTHESIZING MIXED DISULFIDE CONJUGATES OF THIENOPYRIDINE COMPOUNDS

<130> UM-35567/WO-1/ORD

<150> US 62/624,494
<151> 2018-01-31

<160> 3

<170> PatentIn version 3.5

<210> 1
<211> 3150
<212> DNA
<213> artificial sequence

<220>
<223> synthetic

<400> 1
atgacaatta aagaaatgcc tcagccaaaa acgtttggag agcttaaaaa tttaccgtta 60

ttaaacacag ataaaccggt tcaagctttg atgaaaattg cggatgaatt aggagaaatc 120

tttaaattcg aggcgcctgg tcgtgtaacg cgctacttat caagtcagcg tctaattaaa 180

gaagcatgcg atgaatcacg ctttgataaa aacttaagtc aagcgcttaa atttgtacgt 240

gattttgcag gagacgggtt atttacaagc tggacgcatg aaaaaaattg gaaaaaagcg 300

cataatatct tacttccaag cttcagtcag caggcaatga aaggctatca tgcgatgatg 360

gtcgatatcg ccgtgcagct tgttcaaaag tgggagcgtc taaatgcaga tgagcatatt 420

gaagtaccgg aagacatgac acgtttaacg cttgatacaa ttggtctttg cggctttaac 480

tatcgcttta acagctttta ccgagatcag cctcatccat ttattacaag tatggtccgt 540

gcactggatg aagcaatgaa caagctgcag cgagcaaatc cagacgaccc agcttatgat 600

gaaaacaagc gccagtttca agaagatatc aaggtgatga acgacctagt agataaaatt 660

attgcagatc gcaaagcaag cggtgaacaa agcgatgatt tattaacgca tatgctaaac 720

ggaaaagatc cagaaacggg tgagccgctt gatgacgaga acattcgcta tcaaattatt 780

acattcttaa ttgcgggaca cgaaacaaca agtggtcttt tatcatttgc gctgtatttc 840

ttagtgaaaa atccacatgt attacaaaaa gcagcagaag aagcagcacg agttctagta 900

gatcctgttc caagctacaa acaagtcaaa cagcttaaat atgtcggcat ggtcttaaac 960

gaagcgctgc gcttatggcc aactgctcct gcgttttccc tatatgcaaa agaagatacg 1020

gtgcttggag gagaatatcc tttagaaaaa ggcgacgaac taatggttct gattcctcag 1080

cttcaccgtg ataaaacaat ttggggagac gatgtggaag agttccgtcc agagcgtttt 1140

gaaaatccaa gtgcgattcc gcagcatgcg tttaaaccgt ttggaaacgg tcagcgtgcg 1200

tgtatcggtc agcagttcgc tcttcatgaa gcaacgctgg tacttggtat gatgctaaaa 1260

cactttgact ttgaagatca tacaaactac gagctggata ttaaagaaac tttaacgtta 1320

aaacctgaag gctttgtggt aaaagcaaaa tcgaaaaaaa ttccgcttgg cggtattcct 1380

tcacctagca ctgaacagtc tgctaaaaaa gtacgcaaaa aggcagaaaa cgctcataat 1440

acgccgctgc ttgtgctata cggttcaaat atgggaacag ctgaaggaac ggcgcgtgat 1500

ttagcagata ttgcaatgag caaaggattt gcaccgcagg tcgcaacgct tgattcacac 1560

gccggaaatc ttccgcgcga aggagctgta ttaattgtaa cggcgtctta taacggtcat 1620

ccgcctgata acgcaaagca atttgtcgac tggttagacc aagcgtctgc tgatgaagta 1680

aaaggcgttc gctactccgt atttggatgc ggcgataaaa actgggctac tacgtatcaa 1740

aaagtgcctg cttttatcga tgaaacgctt gccgctaaag gggcagaaaa catcgctgac 1800

cgcggtgaag cagatgcaag cgacgacttt gaaggcacat atgaagaatg gcgtgaacat 1860

atgtggagtg acgtagcagc ctactttaac ctcgacattg aaaacagtga agataataaa 1920

tctactcttt cacttcaatt tgtcgacagc gccgcggata tgccgcttgc gaaaatgcac 1980

ggtgcgtttt caacgaacgt cgtagcaagc aaagaacttc aacagccagg cagtgcacga 2040

agcacgcgac atcttgaaat tgaacttcca aaagaagctt cttatcaaga aggagatcat 2100

ttaggtgtta ttcctcgcaa ctatgaagga atagtaaacc gtgtaacagc aaggttcggc 2160

ctagatgcat cacagcaaat ccgtctggaa gcagaagaag aaaaattagc tcatttgcca 2220

ctcgctaaaa cagtatccgt agaagagctt ctgcaatacg tggagcttca agatcctgtt 2280

acgcgcacgc agcttcgcgc aatggctgct aaaacggtct gcccgccgca taaagtagag 2340

cttgaagcct tgcttgaaaa gcaagcctac aaagaacaag tgctggcaaa acgtttaaca 2400

atgcttgaac tgcttgaaaa atacccggcg tgtgaaatga aattcagcga atttatcgcc 2460

cttctgccaa gcatacgccc gcgctattac tcgatttctt catcacctcg tgtcgatgaa 2520

aaacaagcaa gcatcacggt cagcgttgtc tcaggagaag cgtggagcgg atatggagaa 2580

tataaaggaa ttgcgtcgaa ctatcttgcc gagctgcaag aaggagatac gattacgtgc 2640

tttatttcca caccgcagtc agaatttacg ctgccaaaag accctgaaac gccgcttatc 2700

atggtcggac cgggaacagg cgtcgcgccg tttagaggct ttgtgcaggc gcgcaaacag 2760

ctaaaagaac aaggacagtc acttggagaa gcacatttat acttcggctg ccgttcacct 2820

catgaagact atctgtatca agaagagctt gaaaacgccc aaagcgaagg catcattacg 2880

cttcataccg ctttttctcg catgccaaat cagccgaaaa catacgttca gcacgtaatg 2940

gaacaagacg gcaagaaatt gattgaactt cttgatcaag gagcgcactt ctatatttgc 3000

ggagacggaa gccaaatggc acctgccgtt gaagcaacgc ttatgaaaag ctatgctgac 3060

gttcaccaag tgagtgaagc agacgctcgc ttatggctgc agcagctaga agaaaaaggc 3120

cgatacgcaa aagacgtgtg ggctgggtaa 3150

<210> 2
<211> 1048
<212> PRT
<213> artificial sequence

<220>
<223> synthetic

<400> 2

Thr Ile Lys Glu Met Pro Gln Pro Lys Thr Phe Gly Glu Leu Lys Asn
1 5 10 15

Leu Pro Leu Leu Asn Thr Asp Lys Pro Val Gln Ala Leu Met Lys Ile
20 25 30

```
Ala Asp Glu Leu Gly Glu Ile Phe Lys Phe Glu Ala Pro Gly Arg Val
        35                  40                  45

Thr Arg Tyr Leu Ser Ser Gln Arg Leu Ile Lys Glu Ala Cys Asp Glu
    50                  55                  60

Ser Arg Phe Asp Lys Asn Leu Ser Gln Ala Leu Lys Phe Val Arg Asp
65                  70                  75                  80

Phe Ala Gly Asp Gly Leu Phe Thr Ser Trp Thr His Glu Lys Asn Trp
                85                  90                  95

Lys Lys Ala His Asn Ile Leu Leu Pro Ser Phe Ser Gln Gln Ala Met
            100                 105                 110

Lys Gly Tyr His Ala Met Met Val Asp Ile Ala Val Gln Leu Val Gln
        115                 120                 125

Lys Trp Glu Arg Leu Asn Ala Asp Glu His Ile Glu Val Pro Glu Asp
    130                 135                 140

Met Thr Arg Leu Thr Leu Asp Thr Ile Gly Leu Cys Gly Phe Asn Tyr
145                 150                 155                 160

Arg Phe Asn Ser Phe Tyr Arg Asp Gln Pro His Pro Phe Ile Thr Ser
                165                 170                 175

Met Val Arg Ala Leu Asp Glu Ala Met Asn Lys Leu Gln Arg Ala Asn
            180                 185                 190

Pro Asp Asp Pro Ala Tyr Asp Glu Asn Lys Arg Gln Phe Gln Glu Asp
        195                 200                 205

Ile Lys Val Met Asn Asp Leu Val Asp Lys Ile Ile Ala Asp Arg Lys
    210                 215                 220

Ala Ser Gly Glu Gln Ser Asp Asp Leu Leu Thr His Met Leu Asn Gly
225                 230                 235                 240
```

```
Lys Asp Pro Glu Thr Gly Glu Pro Leu Asp Asp Glu Asn Ile Arg Tyr
                245                 250                 255

Gln Ile Ile Thr Phe Leu Ile Ala Gly His Glu Thr Thr Ser Gly Leu
            260                 265                 270

Leu Ser Phe Ala Leu Tyr Phe Leu Val Lys Asn Pro His Val Leu Gln
        275                 280                 285

Lys Ala Ala Glu Glu Ala Ala Arg Val Leu Val Asp Pro Val Pro Ser
    290                 295                 300

Tyr Lys Gln Val Lys Gln Leu Lys Tyr Val Gly Met Val Leu Asn Glu
305                 310                 315                 320

Ala Leu Arg Leu Trp Pro Thr Ala Pro Ala Phe Ser Leu Tyr Ala Lys
                325                 330                 335

Glu Asp Thr Val Leu Gly Gly Glu Tyr Pro Leu Glu Lys Gly Asp Glu
            340                 345                 350

Leu Met Val Leu Ile Pro Gln Leu His Arg Asp Lys Thr Ile Trp Gly
        355                 360                 365

Asp Asp Val Glu Glu Phe Arg Pro Glu Arg Phe Glu Asn Pro Ser Ala
    370                 375                 380

Ile Pro Gln His Ala Phe Lys Pro Phe Gly Asn Gly Gln Arg Ala Cys
385                 390                 395                 400

Ile Gly Gln Gln Phe Ala Leu His Glu Ala Thr Leu Val Leu Gly Met
                405                 410                 415

Met Leu Lys His Phe Asp Phe Glu Asp His Thr Asn Tyr Glu Leu Asp
            420                 425                 430
```

```
Ile Lys Glu Thr Leu Thr Leu Lys Pro Glu Gly Phe Val Val Lys Ala
        435                 440                 445

Lys Ser Lys Lys Ile Pro Leu Gly Gly Ile Pro Ser Pro Ser Thr Glu
    450                 455                 460

Gln Ser Ala Lys Lys Val Arg Lys Lys Ala Glu Asn Ala His Asn Thr
465                 470                 475                 480

Pro Leu Leu Val Leu Tyr Gly Ser Asn Met Gly Thr Ala Glu Gly Thr
                485                 490                 495

Ala Arg Asp Leu Ala Asp Ile Ala Met Ser Lys Gly Phe Ala Pro Gln
            500                 505                 510

Val Ala Thr Leu Asp Ser His Ala Gly Asn Leu Pro Arg Glu Gly Ala
        515                 520                 525

Val Leu Ile Val Thr Ala Ser Tyr Asn Gly His Pro Pro Asp Asn Ala
    530                 535                 540

Lys Gln Phe Val Asp Trp Leu Asp Gln Ala Ser Ala Asp Glu Val Lys
545                 550                 555                 560

Gly Val Arg Tyr Ser Val Phe Gly Cys Gly Asp Lys Asn Trp Ala Thr
                565                 570                 575

Thr Tyr Gln Lys Val Pro Ala Phe Ile Asp Glu Thr Leu Ala Ala Lys
            580                 585                 590

Gly Ala Glu Asn Ile Ala Asp Arg Gly Glu Ala Asp Ala Ser Asp Asp
        595                 600                 605

Phe Glu Gly Thr Tyr Glu Glu Trp Arg Glu His Met Trp Ser Asp Val
    610                 615                 620

Ala Ala Tyr Phe Asn Leu Asp Ile Glu Asn Ser Glu Asp Asn Lys Ser
625                 630                 635                 640
```

```
Thr Leu Ser Leu Gln Phe Val Asp Ser Ala Ala Asp Met Pro Leu Ala
                645                 650                 655

Lys Met His Gly Ala Phe Ser Thr Asn Val Val Ala Ser Lys Glu Leu
            660                 665                 670

Gln Gln Pro Gly Ser Ala Arg Ser Thr Arg His Leu Glu Ile Glu Leu
        675                 680                 685

Pro Lys Glu Ala Ser Tyr Gln Glu Gly Asp His Leu Gly Val Ile Pro
    690                 695                 700

Arg Asn Tyr Glu Gly Ile Val Asn Arg Val Thr Ala Arg Phe Gly Leu
705                 710                 715                 720

Asp Ala Ser Gln Gln Ile Arg Leu Glu Ala Glu Glu Glu Lys Leu Ala
                725                 730                 735

His Leu Pro Leu Ala Lys Thr Val Ser Val Glu Glu Leu Leu Gln Tyr
            740                 745                 750

Val Glu Leu Gln Asp Pro Val Thr Arg Thr Gln Leu Arg Ala Met Ala
        755                 760                 765

Ala Lys Thr Val Cys Pro Pro His Lys Val Glu Leu Glu Ala Leu Leu
    770                 775                 780

Glu Lys Gln Ala Tyr Lys Glu Gln Val Leu Ala Lys Arg Leu Thr Met
785                 790                 795                 800

Leu Glu Leu Leu Glu Lys Tyr Pro Ala Cys Glu Met Lys Phe Ser Glu
                805                 810                 815

Phe Ile Ala Leu Leu Pro Ser Ile Arg Pro Arg Tyr Tyr Ser Ile Ser
            820                 825                 830
```

```
Ser Ser Pro Arg Val Asp Glu Lys Gln Ala Ser Ile Thr Val Ser Val
        835                 840                 845

Val Ser Gly Glu Ala Trp Ser Gly Tyr Gly Glu Tyr Lys Gly Ile Ala
    850                 855                 860

Ser Asn Tyr Leu Ala Glu Leu Gln Glu Gly Asp Thr Ile Thr Cys Phe
865                 870                 875                 880

Ile Ser Thr Pro Gln Ser Glu Phe Thr Leu Pro Lys Asp Pro Glu Thr
                885                 890                 895

Pro Leu Ile Met Val Gly Pro Gly Thr Gly Val Ala Pro Phe Arg Gly
            900                 905                 910

Phe Val Gln Ala Arg Lys Gln Leu Lys Glu Gln Gly Gln Ser Leu Gly
        915                 920                 925

Glu Ala His Leu Tyr Phe Gly Cys Arg Ser Pro His Glu Asp Tyr Leu
    930                 935                 940

Tyr Gln Glu Glu Leu Glu Asn Ala Gln Ser Glu Gly Ile Ile Thr Leu
945                 950                 955                 960

His Thr Ala Phe Ser Arg Met Pro Asn Gln Pro Lys Thr Tyr Val Gln
                965                 970                 975

His Val Met Glu Gln Asp Gly Lys Lys Leu Ile Glu Leu Leu Asp Gln
            980                 985                 990

Gly Ala His Phe Tyr Ile Cys Gly  Asp Gly Ser Gln Met  Ala Pro Ala
        995                 1000                1005

Val Glu  Ala Thr Leu Met Lys  Ser Tyr Ala Asp Val  His Gln Val
    1010                1015                1020

Ser Glu  Ala Asp Ala Arg Leu  Trp Leu Gln Gln Leu  Glu Glu Lys
    1025                1030                1035
```

Gly Arg Tyr Ala Lys Asp Val Trp Ala Gly
    1040   1045

<210> 3
<211> 3162
<212> DNA
<213> artificial sequence

<220>
<223> synthetic

<400> 3
atgcaccatc atcatcatca tattaaggag atgccgcagc caaaaacatt cggcgaactc   60

aaaaacttac cattactgaa taccgacaaa ccggtccaag cactgatgaa aattgcggac   120

gaattaggtg aaatcttcaa attcgaggcg cccggtcgcg taacacgtta tttatccagt   180

cagcgcctta tcaaagaagc gtgtgatgaa agtcgttttg ataaaaatct gtcccaggca   240

cttaaatttg ttcgtgactt tttcggtgat ggcctgttta cctcttggac tcatgaaaaa   300

aactggaaaa aagcgcataa tatcttgctt ccgtcgtttt cgcagcaggc aatgaaaggt   360

taccatgcca tgatggtcga tattgccgtc cagctggtgc aaaaatggga acgtcttaac   420

gctgatgaac atattgaagt gcccgaagac atgacccgtc tgacgctgga tactattgga   480

ctgtgcgggt tcaactatcg tttcaactcc ttctaccgtg atcagccaca tccgtttatt   540

acttctatgg tccgcgcctt agacgaagcc atgaacaaac tgcagcgcgc caacccagac   600

gacccagctt atgatgagaa taaacgtcag tttcaagaag acatcaaagt catgaacgac   660

ttagtggata aaattattgc agaccgtaaa gcgagcggcg aacagagtga tgacctgctt   720

acccacatgc tgaatggtaa agatccagag accggcgagc cgttagatga tgaaaatatt   780

cgctaccaga tcattacctt tttaatcgca ggacacgaaa caacaagtgg actgctcagc   840

tttgcactct actttctggt taaaaacccg catgttctgc aaaaagcagc ggaagaggcc   900

gcccgtgtgc tggtcgatcc ggttcccagc tataaacagg tcaaacagtt aaaatacgtg   960

ggcatggtct taaacgaggc tctgcgctta tggccaacag caccagcatt ttcgttatat   1020

gcaaaagaag ataccgttct gggaggagaa tacccgttag aaaaaggcga cgagcttatg   1080

gtgctgatcc cacagttaca ccgtgataaa accatttggg gcgacgatgt ggaagaattt 1140

cgcccagaac gtttcgagaa ccctagcgca attccacagc atgccttcaa acccttcggg 1200

aacggtcagc gcgcgtgcat tgggcagcag ttcgcgctgc atgaagcaac tttggtgtta 1260

ggcatgatgc tgaaacactt tgattttgaa gaccacacga attatgaact ggatattaaa 1320

gaaaccctga cactgaaacc agaaggattc gtagttaaag cgaaaagcaa aaagattccg 1380

ctgggtggca ttccgagccc atccaccgaa cagagcgcga aaaaagttcg gaaaaaggcg 1440

gaaaatgcgc acaatacccc cttgttagtc ctttacggct caaatatggg cacagcagaa 1500

ggcaccgcac gtgacttagc cgatattgca atgagcaagg gtttcgcgcc ccaagtcgcg 1560

accttggatt cacacgctgg aaacctgccg cgggaaggcg ccgtccttat cgttactgcc 1620

tcatataacg gtcaccctcc ggacaatgcg aaacaatttg tggactggtt agatcaagcc 1680

tcggccgacg aagtgaaagg cgttcgttat tctgtttttg gatgtgggga taaaaactgg 1740

gcgacgacgt accaaaaagt ccctgctttt attgatgaaa cgttggctgc aaaaggtgca 1800

gaaaacattg cagaccgtgg cgaagcagac gcgagcgacg actttgaagg tacctatgag 1860

gaatggcgtg aacacatgtg gagtgatgtc gccgcttact tcaacttaga tattgaaaat 1920

tccgaagata ataaaagtac cctgagcttg caattcgtgg actcggctgc cgacatgccg 1980

ctcgctaaaa tgcacggggc gtttagtacg aatgtagtgg cttccaaaga gttgcaacaa 2040

cccggtagcg cacgctcgac ccggcacctg gaaattgaat taccgaagga agcgtcttat 2100

caggaaggag atcatctggg tgtaatccca cgcaattacg aaggtattgt taatcgcgtt 2160

accgcgcgtt ttggtttaga tgcctcccaa caaatccgtt tagaagcaga agaagaaaaa 2220

ctcgcgcatt tacccttagc caaaaccgtt tcggtcgaag aactgctgca atatgttgaa 2280

cttcaggacc ctgtgacccg tacccagctc cgtgccatgg ccgcgaaaac agtatgccca 2340

ccccacaaag ttgaattaga ggcgctgtta gagaaacaag catacaaaga acaagtgtta 2400

gctaagcgtc tgaccatgtt agagttactg gagaaatatc cggcgtgcga gatgaaattc 2460

tcagaattca ttgcattgtt gccgagcatt cgtccgcggt attacagtat ctcgagctca 2520

ccgcgcgttg atgaaaaaca ggcctctatt acggtctccg tagtttccgg cgaagcctgg 2580

agcgggtatg gagaatataa aggaattgct agcaactatc tcgcggagct gcaagagggc 2640

gacactatta catgcttcat ttctacgccg caatccgaat ttacactgcc gaaagacccg 2700

gaaacgccac tcattatggt aggcccaggt actggcgtag cgccatttcg cggattcgtt 2760

caggctcgta aacagttgaa agaacaaggt caaagtcttg gcgaagcaca tttatacttc 2820

ggctgccgct cgccgcatga ggactatctc tatcaggaag aattggagaa cgcacagagt 2880

gagggcatta tcaccttgca tacggctttt tctcgcatgc ctaatcaacc taaaacctat 2940

gtccaacatg tgatggagca ggatggaaaa aaattgatcg agctgttgga tcagggcgcg 3000

cattttaca tttgcgggga tggttcgcag atggcacccg ccgtggaggc cacccttatg 3060

aaaagctatg cagatgtgca ccaggtaagc gaagcggatg cccgtctgtg gctgcaacag 3120

ttggaagaaa aaggtcgcta tgcaaaagac gtgtgggcag gt 3162