

Supporting Information

**Insight into the Programmed Ketoreduction of Partially Reducing
Polyketide Synthases: Stereo- and Substrate-Specificity of Ketoreductase
Domain**

Ishin Soehano, Lifeng Yang, Feiqing Ding, Huihua Sun, Zhen Jie Low, Xuewei Liu, Zhao-
Xun Liang*

^a Division of Structural Biology & Biochemistry, School of Biological Sciences,
Nanyang Technological University, 60 Nanyang Drive, Singapore 637551; Tel: (+) 65
6592 7736, Fax: (+) 65 6791 3856, E-mail: zxliang@ntu.edu.sg.

^b Division of Chemistry and Biological Chemistry, School of Mathematics and Physics,
Nanyang Technological University, 21 Nanyang Link, Singapore 637371,

Table S1. SACE3352 and homologous bacterial iPKSs.

Gene	Access no.	Organism	Identity%/Homology%
NcsB	AAM77986	<i>Streptomyces carzinostaticus</i> subsp. <i>neocarzinostaticus</i>	51/66
AziB	ABY83164	<i>Streptomyces sahachiroi</i>	48/64
PokM1	ACN64831	<i>Streptomyces diastatochromogenes</i>	49/63
ChlB1	AAZ77673	<i>Streptomyces antibioticus</i>	46/61
MdpB	ABY66019	<i>Actinomadura madurae</i>	47/60
CalO5	AAM70355	<i>Micromonospora echinospora</i>	48/62
AviM	AAK83194	<i>Streptomyces viridochromogenes</i>	48/61

TableS2. PCR primers used to generate double mutants KR_{SACE5532}KR_{I2A} and KR_{I2B}, as well as quadruple mutant KR_{SACE5532}KR_{I4}. F, forward primer, R, reverse primer

KR _{SACE5532} variants	Amino acid replacements	Primer sequences
KR _{I2A}	T1530→LT1 531→S	F: 5'-CCAGTTTGCCCGCCTGTCTGGCCAAGTCACCTATG -3' R: 5'-CATAGGTGACTTGGCCAGACAGGCGGGCAAACCTGG -3'
KR _{I2B}	V1534→AT1 535→S	F: 5'- ACGGGCCAAGCGAGCTATGCGGGCGGCGAACAGCTTTCTGGATG-3' R: 5'-TTCGCCCGCCGCATAGCTCGCTTGGCCCCGTGGTGGGGCAAAC- 3'
KR _{I4}	T1530→LT1 531→S V1534→AT1 535→S	F: 5'- CTGTCTGGCCAAGCAAGTTATGCGGGCGGCGAACAG-3' R: 5'- CTGTTCCGCCCGCCGCATAACTTGCTTGGCCAGACAG-3'

TableS3. PCR Primers used to generate KR_{SACE5532}KR_{III}and KR_{SACE5532}KR_{I4/II/III}mutants.

Primer names	Primer sequences
Forward primer A	5'-ACGTTTCATATGCGCGATCTGGCGTATGAAATCATTGGAAACCG-3'
Reverse primer B	5'-ACGCGCTTTCGGAGCCAGGACTTCGCGCAGACCTTCCAGATCAA CTTTATCAACCAGCGCATTGTTACACCAGCCCCGCGGCATGAACC-3'
Forward primer C	5'-TGGTGAACAATGCGCTGGTTGATAAAGTTGATCTGGAAGGTCT GCGCGAAGTCCTGGCTCCGAAAGCGCGTGGCGCGATGGC-3'
Reverse primer D	5'-AACAAACTCGAGGCCGGTATCGCCGCTCGCGGTCAGTTC-3'

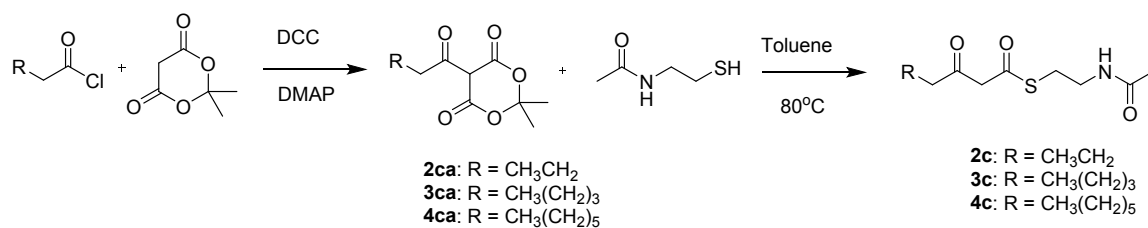
Table S4. PCR Primers used to generate the SACE5532 variant KR_{I/II/III/IV}.

Primer names	Primer sequences
Forward primer A	5'- TGCCGCGCGGCAGCCATATGAGCACCTATGCCGCGGTAAAG -3'
Reverse primer B	5'- AGGTCACGCGGGCGCAGCCATGGTGCCCGGAGCACCATCCAGAAC -3'
Forward primer C	5'- TGGCTGCGCCGCGTGACCTGGCTTACGAAATCATCTG -3'
Reverse primer D	5'- TCATCCGGACCCGTATCACCCGATGCGGTCAGTTC -3'
Forward primer E	5'- TGATACGGGTCCGGATGAAGGCCCGTCAATCGTGGAAG -3'
Reverse primer F	5'- TCAGTGGTGGTGGTGGTGGTGCTCGAGGCTCACTTTCGCTTTCAGCAG -3'

TableS5. Standalone ketoreductase (KR) domains and the amino acid replacements in KR_{SACE5532} tested in this work.

KR names	KR _{SACE5532} amino acid replacement			
	Motif I (red)	Motif II (green)	Motif III (blue)	Motif IV (magenta)
KR _{SACE5532}	Nil	nil	nil	nil
KR _{I2A}	T1530→L, T1531→S	nil	nil	nil
KR _{I2B}	V1534→A, T1535→S	nil	nil	nil
KR _{I4}	T1530→L, T1531→S V1534→A, T1535→S	nil	nil	nil
KR _{II}	Nil	M1564→T, L1567→R N1568→G, T1569→L A1572→S, A1573→S S1574→N, E1576→D S1577→T, S1578→T A1580→F, R1583→N A1584→S, N1585→R I1587→L, D1588→E G1589→A, F1590→V T1591→S	nil	nil
KR _{III}	Nil	nil	D1481→N, A1484→V E1485→D, N1486→K, M1487→V, T1489→L, D1490→E, A1493→R, R1494→E, T1495→V M1496→L, S1497→A	nil
KR _{I2A/II}	T1530→L, T1531→S	M1564→T, L1567→R N1568→G, T1569→L A1572→S, A1573→S S1574→N, E1576→D S1577→T, S1578→T A1580→F, R1583→N A1584→S, N1585→R I1587→L, D1588→E G1589→A, F1590→V T1591→S	nil	nil
KR _{I2B/II}	V1534→A, T1535→S	M1564→T, L1567→R N1568→G, T1569→L A1572→S, A1573→S S1574→N, E1576→D S1577→T, S1578→T A1580→F, R1583→N A1584→S, N1585→R I1587→L, D1588→E G1589→A, F1590→V T1591→S	nil	nil
KR _{I4/II}	T1530→L, T1531→S V1534→A, T1535→S	M1564→T, L1567→R N1568→G, T1569→L A1572→S, A1573→S S1574→N, E1576→D S1577→T, S1578→T A1580→F, R1583→N A1584→S, N1585→R I1587→L, D1588→E G1589→A, F1590→V	nil	nil

		T1591→S		
KR _{I4/II/III}	T1530→L, T1531→S V1534→A, T1535→S	M1564→T, L1567→R N1568→G, T1569→L A1572→S, A1573→S S1574→N, E1576→D S1577→T, S1578→T A1580→F, R1583→N A1584→S, N1585→R I1587→L, D1588→E G1589→A, F1590→V T1591→S	D1481→N, A1484→V E1485→D, N1486→K, M1487→V, T1489→L, D1490→E, A1493→R, R1494→E, T1495→V M1496→L, S1497→A	nil
KR _{VI/III/IV}	S1524→C, T1530→L, T1531→S V1534→A, T1535→S	M1564→T, L1567→R N1568→G, T1569→L A1572→S, A1573→S S1574→N, E1576→D S1577→T, S1578→T A1580→F, R1583→N A1584→S, N1585→R I1587→L, D1588→E G1589→A, F1590→V T1591→S	D1481→N, A1484→V E1485→D, N1486→K, M1487→V, T1489→L, D1490→E, A1493→R, R1494→E, T1495→V M1496→L, S1497→A	I1387→L, Q1390→E G1397→D, N1409→R, G1410→A P1413→N, H1415→T D1418→P, T1424→E T1427→R, E1430→D A1431→G, R1433→L S1434→A, Y1437→A
KR _{NcsB}	Nil	nil	nil	nil



Scheme S1. General Procedures for synthesis of compounds **2c**, **3c** and **4c**.

Figure S1. Sequence alignment of KR_{SACE5532} and KR_{NcsB} with a sequence identity of 45% and sequence homology of 63%. The four motifs discussed in the main text are highlighted.

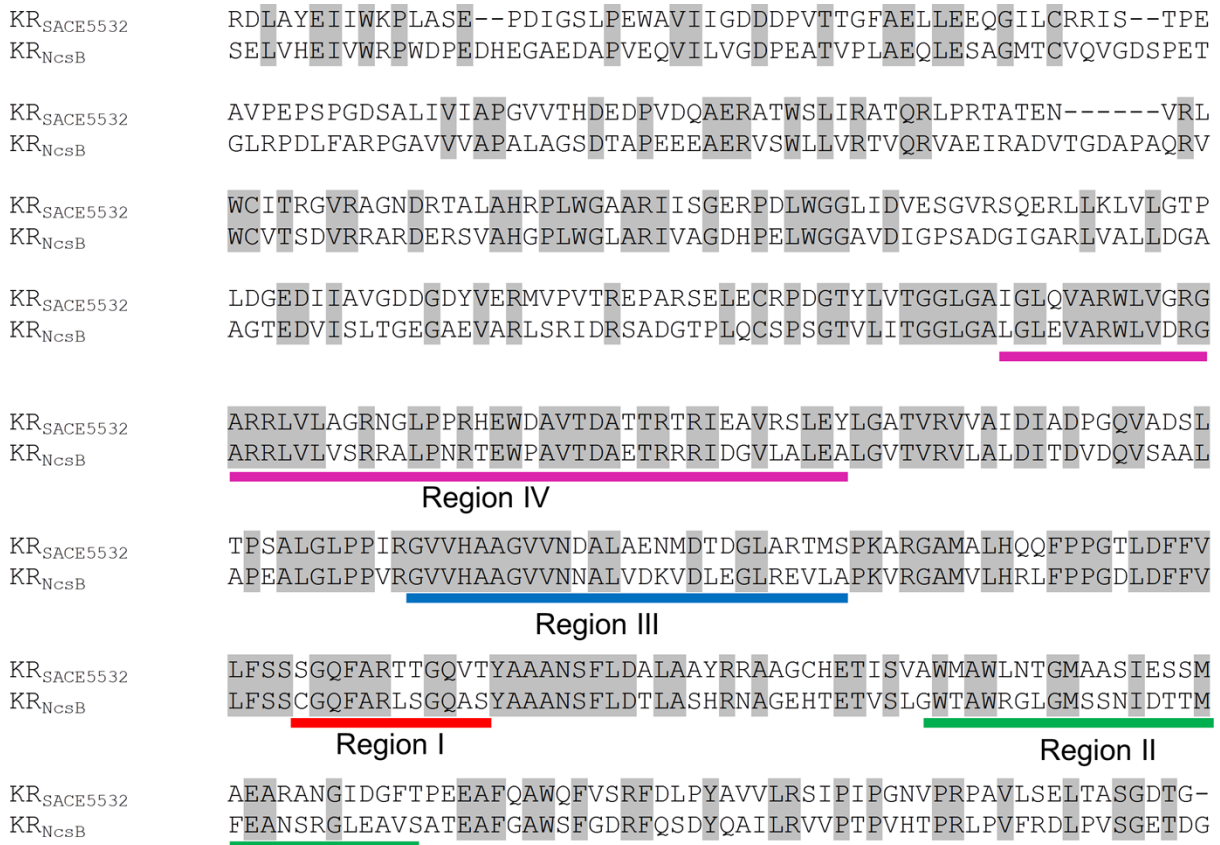
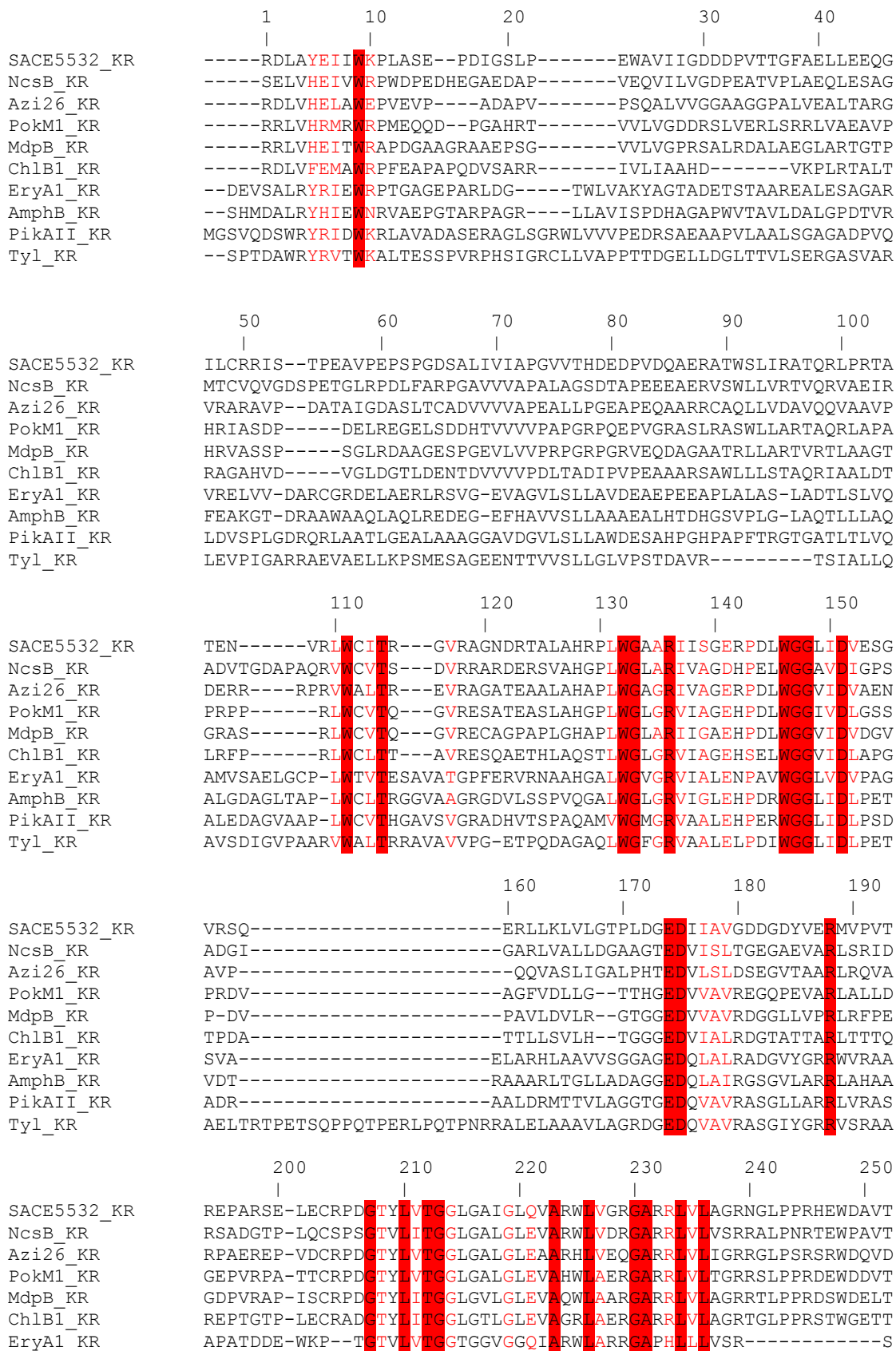


Figure S2. Multiple sequence alignment of the KR domains of selected type I partially reducing polyketide synthases and modular polyketide synthases. The conserved residues were highlighted with red color.

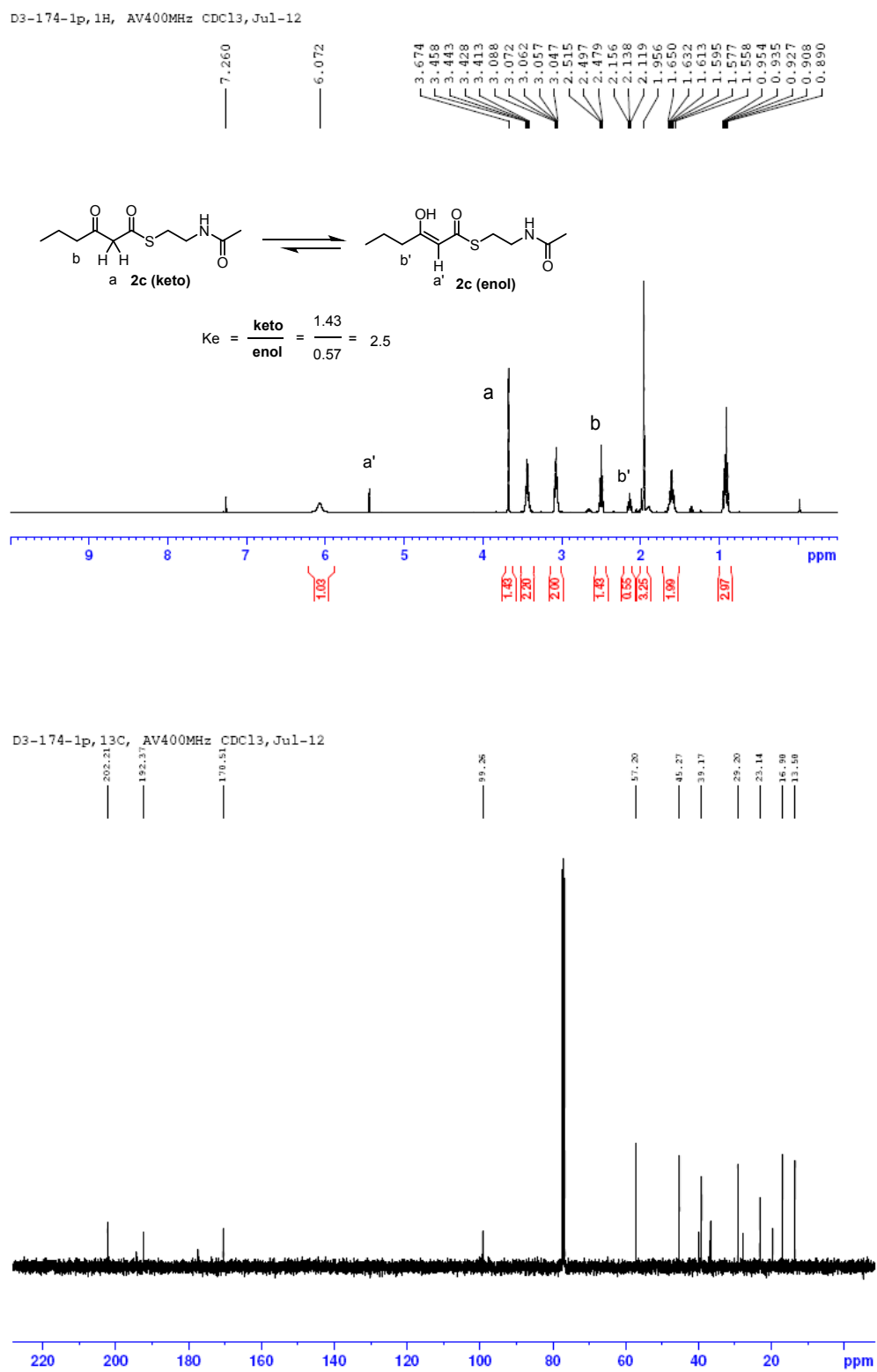


```

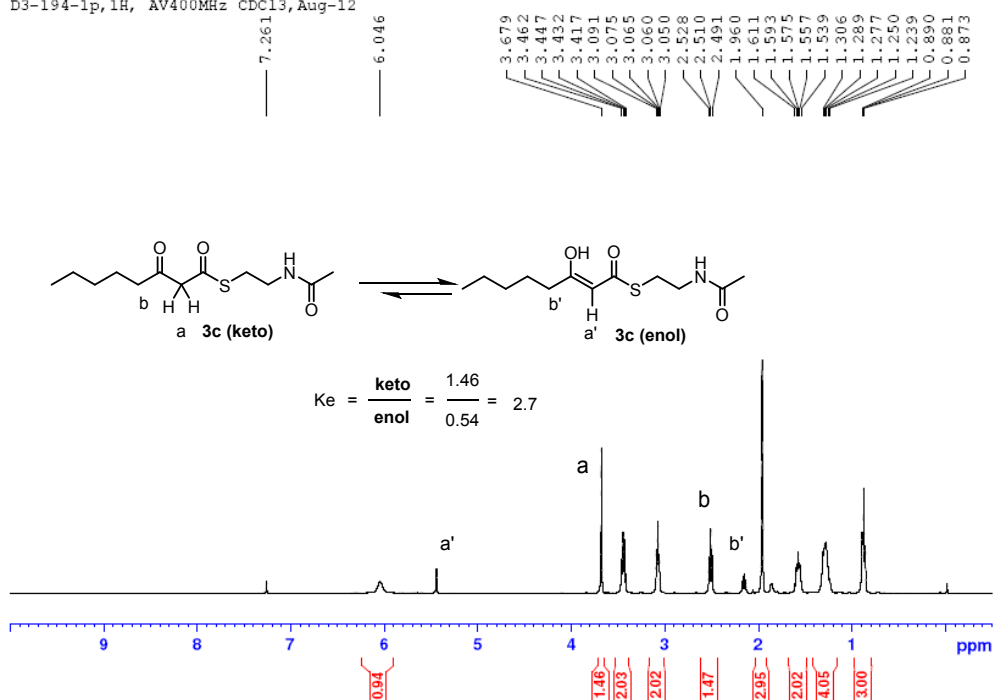
AmphB_KR      PAVPGSG-KRPPVHGSVIVTGGTGGIGGRVARRAEQGAHVITSR-----R
PikAII_KR     LPAHGTSPPWQADGTVLVTGAEPPAAEAARRLARDGAGHLLHTTPSG-SEGAEGTSG
Tyl_KR        AAG----AASWQPSGTVLITGGMGAIGRRLARRLAAEGAERLVLTSR-----R
                260       270       280       290       300       310
                |         |         |         |         |         |
SACE5532_KR   DATTRTRIEAVRSLEYLGGATVRVVAIDIDADPGQVADSLTPSALGLFPPIRGVVHAAGVVND
NcsB_KR       DAETRRRIDGVLALAEALGVTVRVVALDITDQVQVSAALAPEALGLFPVVRGVVHAAGVVNN
Azi26_KR      DPAAVAQIAEVVAIEAAGATVRVLSLDISDAEATARALDPGALDMPVVRGIVHCAGVSD
PokM1_KR      DPLVRSRVEAVRSIERLGGSTVVTVALDLADADAAEKLSPALGLFPVVRGVVHAAGVLD
MdpB_KR       DPADVAVRAIRSLERLGGTVVVVTVDVADAETAGKLLSPAALGLFPPIRGVVHAAGVVND
ChlB1_KR      DTHTRQRIEAVKALDQGVTVRVIPLDITDTAKAAEQQLTPDALGLFPPIRGIVHLAGVLDN
EryA1_KR     GPDADGAGELVAELAEALGARTVAACDVTDRESVRELLGGIGDDVLSAVFHAA-ATLDD
AmphB_KR     GADAPGAAELRAELEQLGVRVTIACDAADREALAALLAELPEDAPLTAVFHSAGVAHDD
PikAII_KR    AAEDSGLAGLVAELADLGATATVVTCDLTDAAAARLLAGVSDAHELSAVLHLP-PTVDS
Tyl_KR       GPEAPGAAELAEELRGHGCEVVHAACDVAERDALAALVT---AYEPNAVFTA-GILDD
                320       330       340       350       360
                |         |         |         |         |
SACE5532_KR   ALAENMDTDGLARTMSPKARGAMALHQQFPPGTLD-----FVLFSSSGQFARTTGQVTY
NcsB_KR       ALVDKVDLEGLREVLPKVRGAMVLRHRLFPFGDL-----FVLFSSCGQFARLSGQASY
Azi26_KR      ALVEKTGAANLDTTMGPKADGAMVLRHRLFPAGTL-----FVTFSSCGQLARLTGQVSY
PokM1_KR      RPLRALDEDSLRAVLRPKAEGAVVLRHQLFPPGSLD-----FVLFSSCGQLLGLPGQSSY
MdpB_KR       RLLRDLDGESLRAVQRPKIGGALVLRHSLFPPGEPD-----FVLFSSCGQLLGLPGQAY
ChlB1_KR      RMVTAVDETSLRTVLRPKADGAWTLHHTLFPFGTID-----FLILFSSCGQLLGLPGQAY
EryA1_KR     GTVDTLTGERIERASRAKVLGARNLHELTRDLT-----AFVLFSSFASAFGAPGLGGY
AmphB_KR     APVADLTGQLDALMRAKLTAARHLHELTDLDD-----AFVLFSSGAAVWGSQPGY
PikAII_KR    EPLAATDADALARVVTAKATAALHLDRLLREAAAAGRPPVLVLFSSVAAIWGGAGQGY
Tyl_KR       AVIDTLSPESFETVRGAKVCGAELHQLTADIKGLD----AFVLFSSVTGTWGNAGQGY
                370       380       390       400       410       420
                |         |         |         |         |         |
SACE5532_KR   AAANSFLDALAYRRAAG--CHETISVAMMNLNTGMAAS-IESSMAEARANGIDGFTPE
NcsB_KR       AAANSFLDTLASHRNAGE--HTEVSLGWTAWRGLGMSSN-IDTTFEANSRGLAVSAT
Azi26_KR      ASANSFLDALAALRRSRG--ETGTSFAMAQWIRGRMGETTGRATILEAESRGLGGITVS
PokM1_KR      AAGNAFLDALAAHGAAG--DTGTVSFGWTSWRGLGMSTSSSEVIDAELA-AHETADISAL
MdpB_KR       AASNAFLDGLAAHRNAG--DTGTISFGWTSWRGLGMSRSSAVIDEELA-ARGTADISRA
ChlB1_KR     GSANAFLDALAVHRNTPPTAADDTSFGWTSWRGQGMVN-DVVDALR-ARGVDTITQ
EryA1_KR     APGNAYLDGLAQRRSDG---LPATAVAVGTWAGSGMAEG-AVADRFR--HCVIEMPE
AmphB_KR     AAANAYLDALAEHRRLG---LTASSVAVGTWGEVGMATDPEVHDLRVR--QVFLAMEPE
PikAII_KR    AAGTAFLDALAGQHRADG---PTVTSVAWSPWEGSRVTEG-ATGERLRR--LGLRPLAPA
Tyl_KR       AAANAALDALAERRRAAG---LPATSVAVGLWGGGMAAG-AGEESLSR--RCLRAMDPD
                430       440       450       460
                |         |         |         |
SACE5532_KR   EAFQAWQFVSRFD-----LPYAVVLRSIPIPGNVPRPAVLSLTLASGDTG-----
NcsB_KR       EAFGAWSFGRFQ-----SDYQAILRVVPTPVHTPRLPVFRDLPVSGETDG-----
Azi26_KR     EALRSWAYADRFA-----LPYAAVMRVMPD----HTLPVFSHLSVTDAG-----
PokM1_KR     EAFASWELADRYG-----LGAAVLRITLPPPEPQRRLLPGLGVPPTDAPH-----
MdpB_KR      EAFQAWELAERYD-----LGAAVLRITLPLDPGRRRPLFDDLPAPAA-----
ChlB1_KR     EAFAAWDFAAQH-----PGNYVLRRLPHEPMDQLLSEIHHTGPT-----
EryA1_KR     TACRALQNALDRAEVCPIVIDVRWDRFLAYTAQRPTRLFDEIDAR-----
AmphB_KR     HALGALDQMLENDDTAAAITLMDWEMFAPAFATANRPSALLSTVPEAVSALS-----
PikAII_KR    TALTALDTALGHGDTAVTIADVWSSFAPGFTTARPGTLLADLPEARRALDEQQST
Tyl_KR       AAVDALLGAMGRNDVCVTVVDVWDFAPATNAIRPGRLFDTVPEAREALT-----
                |         |         |         |
SACE5532_KR   DIADPGQVADSLTPSALGLFPPIRGVVHAAGVVND.....FSSSGQFARTTGQVTY
NcsB_KR       DITDQVQVSAALAPEALGLFPVVRGVVHAAGVVNN.....FSSCGQFARLSGQASY
Azi26_KR     DISDAEATARALDPGALDMPVVRGIVHCAGVSD.....FSSCGQLARLTGQVSY
PokM1_KR     DLADADAAEKLSPALGLFPVVRGVVHAAGVLD.....FSSCGQLLGLPGQSSY
MdpB_KR      DVADAETAGKLLSPAALGLFPPIRGVVHAAGVVND.....FSSCGQLLGLPGQAY
ChlB1_KR     DITDTAKAAEQQLTPDALGLFPPIRGIVHLAGVLDN.....FSSCGQLLGLPGQAY
Ery_KR1     DVTDRESVRELLGGIGDDVLSAVFHAA-ATLDD.....FSSFASAFGAPGLGGY
Amp_KR1     DAADREALAALLAELPEDAPLTAVFHSAGVAHDD.....FSSGAAVWGSQPGY
Pik_KR1     DLTDAAAARLLAGVSDAHELSAVLHLP-PTVDS.....FSSVAAIWGGAGQGY
Tyl_KR1     DVAERDALAALVT---AYEPNAVFTA-GILDD.....FSSVTGTWGNAGQGY

```

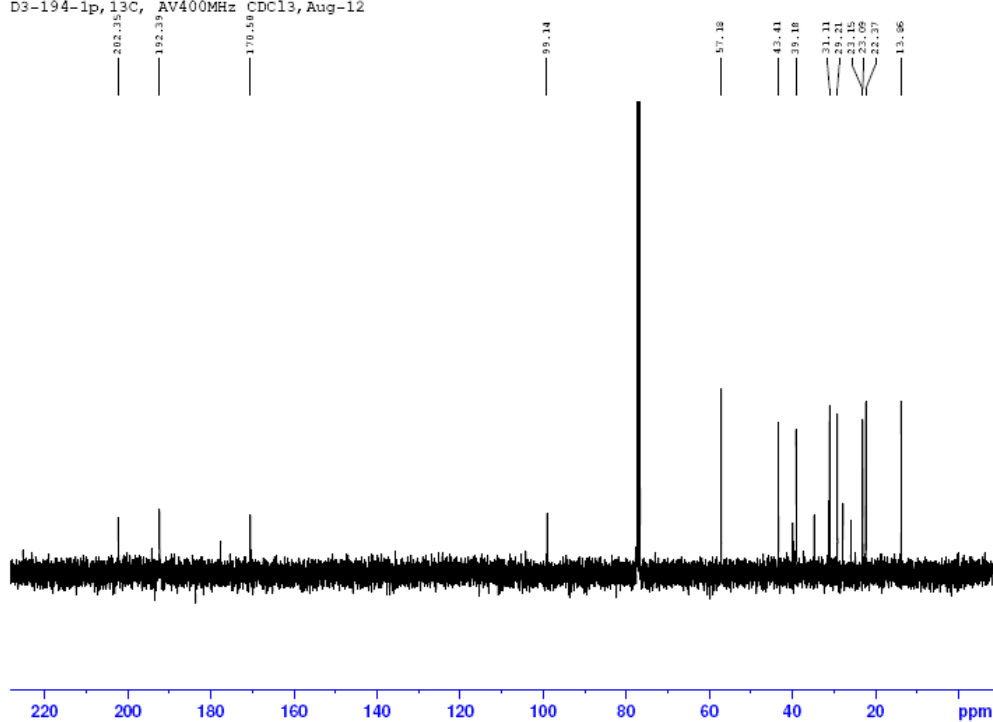

Fig. S3. ^1H and ^{13}C NMR spectra of SNAC substrates **2c**, **3c** and **4c**.



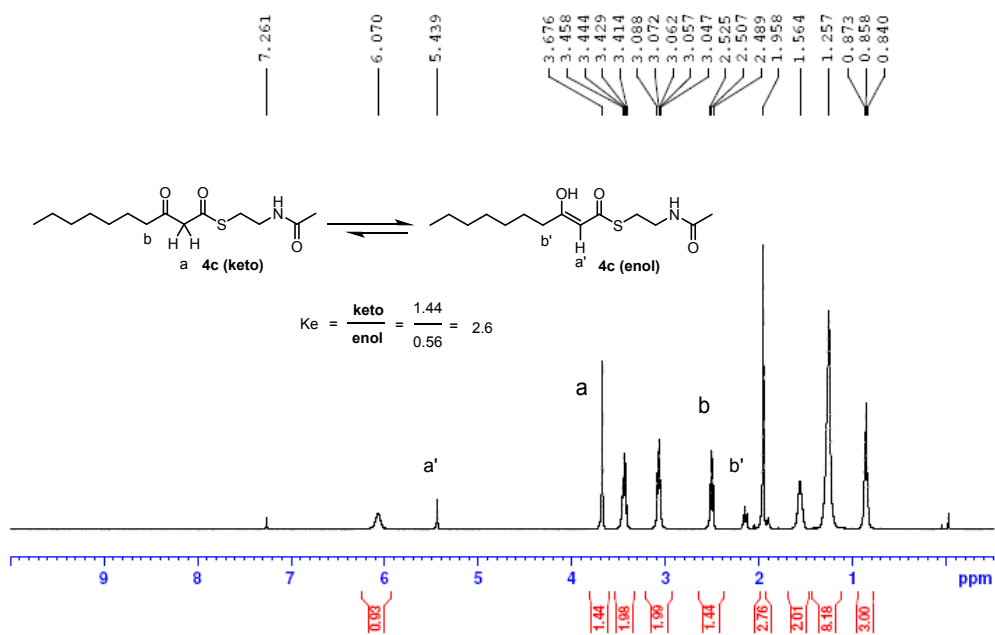
D3-194-1p, 1H, AV400MHz CDCl3, Aug-12



D3-194-1p, 13C, AV400MHz CDCl3, Aug-12



D3-198-1p, 1H, AV400MHz CDCl3, Aug-12



D3-198-1p, 13C, AV400MHz CDCl3, Aug-12

