The 2.1Å Crystal Structure of an Acyl-CoA Synthetase from *Methanosarcina acetivorans* reveals an alternate acyl binding pocket for small branched acyl substrates

Manish B. Shah, Cheryl Ingram-Smith, Leroy L. Cooper, Jun Qu,
Kerry S. Smith, and Andrew M. Gulick

Supplemental Material

The purified acyl-CoA synthetase from *M. acetivorans* was precipitated and digested with both trypsin and Glu-C. The digests were analyzed using nano-LC/MS. A high-resolution nano LC system, composed of four Eksigent direct-flow capillary/nano LC pumps (Dublin, CA), a Spark Endurance autosampler (Emmen, Holland), a lab-made valveless reversed phase-trap and nano-LC flow path, and two Vici10-port low-dead-volume valves, was used for separation of AAE-derived peptides. A lab-made dynamic flow nano-spray interface was used to couple the nano-LC system to a Thermofisher LTQ/Orbitrap high resolution mass spectrometer (San Jose, CA). The Eksigent LC pump system included two micro-flow pumps and two nano-flow pumps powered by pressurized nitrogen (100 p.s.i.) with a fast-response active flow-rate control system. A double transport liquid wash, with mobile phases A (3% acetonitrile in 0.1% formic acid) and B (84% acetonitrile in 0.1% formic acid) sequentially, was employed to eliminate carryover by the autosampler.

The MS operates under data dependent mode; one micro cycle is composed of a MS1 survey scan at a resolution of 100,000 by Orbitrap, followed by seven sequential dependent MS2 scans by LTQ. Digested reaction mixtures were loaded on a reversed-phase peptide trap (5 mm x 300 μm I.D.) at a flow rate of 10 μL/min and washed with 3% acetonitrile in 0.1% formic acid for 3 minutes to remove salts and other hydrophilic buffer components. The trap was then switched online with a lab-packed reversed-phase

nano C-18 column, (3 μm reversed-phase particle packed in a 40 cm long, 360 μm O.D. and, 75μm I.D. fused silica capillary ended with a non-coated 2 μm tampered tip). A shallow multi-steps gradient was used to resolve the samples and the on-column flow rate were 250 nL/min. The raw data was converted to dta files with ZSA and combolon filter algorithms to remove low quality DTA files; then the group of dta files was searched against the pre-indexed tryptic peptide database that contains the sequence of the target protein. A mass error threshold of 5 ppm was used for the precursor ions and 1.0 mass unit fragment ions. The result data was grouped into a srf file, and the peptide probability scores were calculated and probability filter (<0.01) was used to remove false-positive identification.

To ensure the credibility of the result, a secondary group of filters was applied to the result: Xcorr>2 for z1, >2.5 for z=2, >3 for z=3, and >4 for z=4; For each of the identified peptides containing either the Lys-256 and Cys-298, a manual examination of the fragment pattern is performed to eliminate database searching artifacts.

MTSLLSQFVS KTDFESYEDF QENFKILVPE NFNFAYDVVD VYARDSPEKL
AMIWCDDYGN EKIFTFKDLK YYSDKAANFF VKHGIGKGDY VMLTLKSRYD
FWYCMLGLHK LGAIAVPATH MLKTRDIVYR IEKAGLKMIV CIAEDDVPEQ
VDEAHAECGD IPLKKAKVGG DVLEGWIDFR KELEESSPIF ERPTGEVSTK
NEDICLVYFS SGTAGFPKMV EHDNTYPLGH ILTAKYWQNV EDDGLHYTVA
DSGWGKCVWG KLYGQWIAGC AVFVYDYDRF EAKNMLEKAS KYGVTTFCAP
PTIYRFLIKE DLSHYNFSTL KYAVVAGEPL NPEVFNRFLE FTGIKLMEGF
GQTETVVTIA TFPWMEPKPG SIGKPTPGYK IELMDRDGRL CEVGEEGEIV
INTMEGKPVG LFVHYGKDPE RTEETWHDGY YHTGDMAWMD EDGYLWFVGR
ADDIIKTSGY KVGPFEVESA LIQHPAVLEC AITGVPDPVR GQVIKATIVL
TKDYTPSDSL KNELQDHVKN VTAPYKYPRI IEFVPELPKT ISGKIRRVEI
RDKDQSQ

Fig 1: The coverage of AAE achieved by nano-LC/Orbitrap by using two enzymes (trypsin and Glu-c) for proteolysis. The amino acid residues in red were these confidently identified, and these in black font were residues not covered. Lys-256 and Cys-298, the positions of the chemical cross-link in the crystal structure, are boxed.

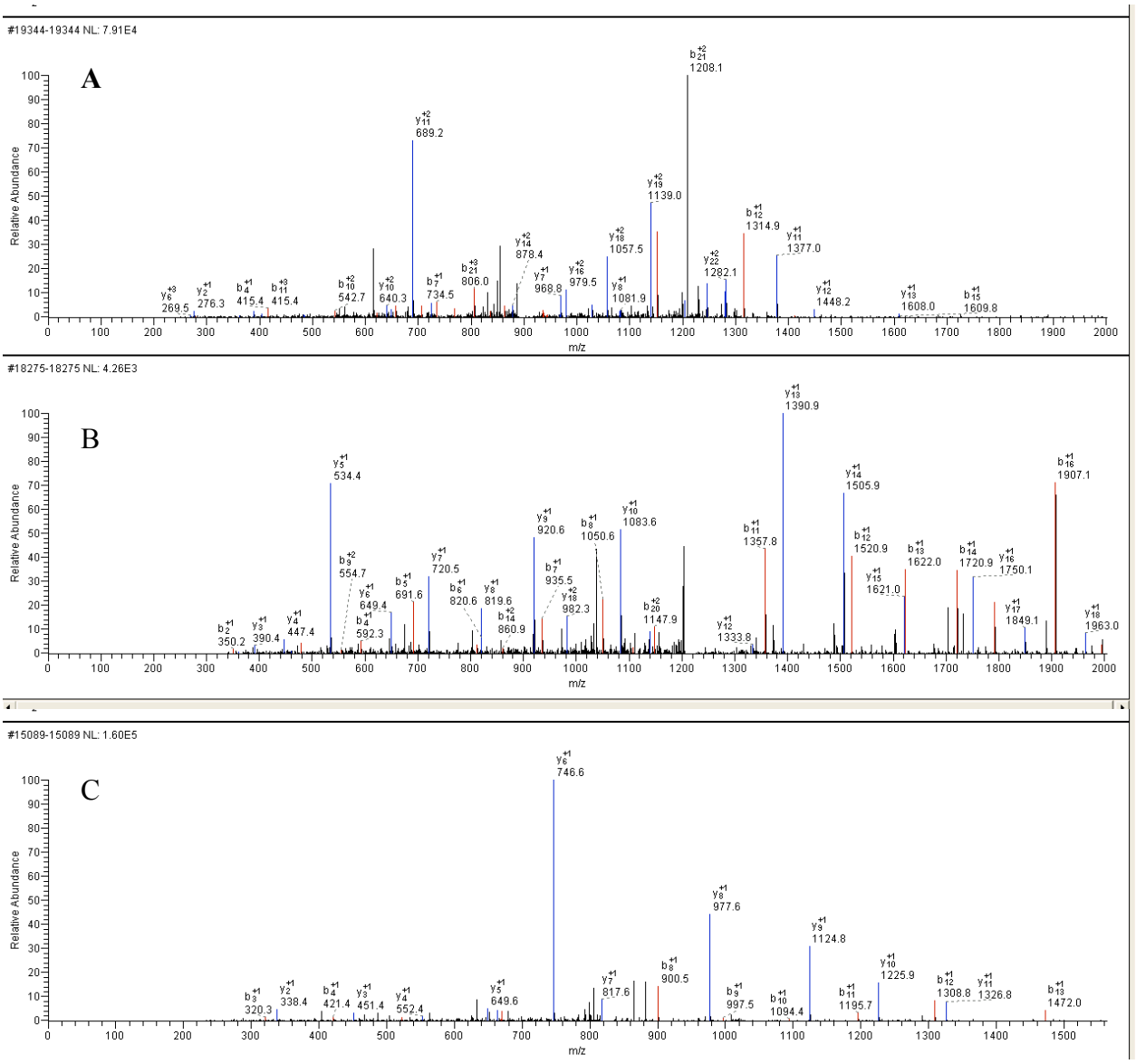

Figure S2: The identification of Lys-256- and Cys-298-containing peptides derived from AAE.

(A) The MS/MS sequencing of KASKYGVTTF**C(298)**APPTIYRFLIKE generated by Glu-c; mass error=0.2 ppm, Xcorr=4.11; z=3 and p=1.33E-15

(B) The MS/MS sequencing of YWQNVEDDGLHYTVADSGWGK(256) generated by trypsin; mass error=0.9 ppm, Xcorr=3.0; z=2 and p=1.00E-30

(C) The MS/MS sequencing of YGVTTFC(298)APPTIYR generated by trypsin; mass error=1.3 ppm, Xcorr=2.8; z=2 and p=3.00E-10;