Electronic Supplementary Material (ESI) for Biomaterials Science.

Rapid and simple purification of elastin-like polypeptides directly from whole cells and cell lysates by organic solvent extraction

Ross VerHeul, Craig Sweet, David H. Thompson*

*Department of Chemistry, Purdue Center for Cancer Research, Multi-disciplinary Cancer Research Facility, Purdue University, 1203 W State Street, West Lafayette, IN 47907, USA

SUPPLEMENTARY DATA

Materials & Methods

1.1 Materials & Methods

All organic solvents (methanol, ethanol, isopropanol, 1-butanol, 2-butanol, 2-methyl-2-propanol, acetone, acetonitrile, and ethyl acetate) were HPLC grade (Fisher Scientific) and used without additional purification or drying. Unless otherwise noted, water was ultra-filtered. BL21 and 5-alpha *E. coli* cells as well as BseRI, AcuI and DNA T4 ligase were purchased from New England Biolabs. Plasmid miniprep and gel extraction kits were purchased from Zymo Research. Alcohol dehydrogenase, lysozyme and lipopolysaccharide were purchased from Sigma-Aldrich. Blue Stain protein ladder and ampicillin were purchased from Gold Biotechnology. Gel Red was purchased from Biotium. B-PER bacterial lysis reagent was purchased from Thermo Fisher Scientific. All media for bacterial cell culture was autoclaved prior to use and contained an appropriate antibiotic. DNA sequences for gene assembly were purchased from IDT.

1.2 Synthesis of ELP construct genes

Briefly pET-25b(+) plasmid was cut with BseRI or AcuI(NEB) to linearize DNA. A leader DNA sequence (IDT) containing the amino acids GHHHHHHNGW was inserted (DNA T4 ligase, NEB) into V12, V24 and S12. A solubilizing domain containing the amino acid sequence $(GKG)_4$ was then inserted in a similar fashion to the end of V12, V24 and S12. ELP genes were then assembled using plasmid reconstruction recursive

directional ligation (Pre-RDL) techniques[1,2] to iteratively assemble DNA block sequences with themselves. The unmodified V12, V24 and S12 was appended making the final constructs(ie: S12-K4-S12). Final gene constructs were sequence-verified and transformed into BL21(DE3) E. coli for expression.

Protein Sequences

GFP (green fluorescent protein, UV variant with an N-terminal polyhistidine tag)

```
      10        20        30        40        50        60
SKGEELFTGV VPILVELDGD VNGHKFSVSG EGEGDATYGK LTLKFICTTG KLPVPWPTLV

      70        80        90       100       110       120
TTFSYGVQCF SRYPDHMKRH DFFKSAMPEG YVQERTISFK DDGNYKTRAE VKFEGDTLVN

     130       140       150       160       170       180
RIELKGIDFK EDGNILGHKL EYNYNSHNVY ITADKQKNGI KANFKIRHNI EDGSVQLADH

     190       200       210       220       230       240
YQQNTPIGDG PVLLPDNHYL STQSKLSKDP NEKRDHMVLL EFVTAAGITH GMDELYKHHH

HHHHH
```

TEL (human tropoelastin, main chain)

```
      10        20        30        40        50        60
GGVPGAVPGG VPGGVFFPGA GLGGLGVGGL GPGVKPAKPG VGGLVGPGLG AEGSALPGAF

      70        80        90       100       110       120
PGGFFGAGGG AAGAAAAAYKA AAKAGAAGLG VGGIGGVGGL GVSTGAVVPQ LGAGVGAGVK

     130       140       150       160       170       180
PGKVPGVGLP GVYPGGVLPG AGARFPGIGV LPGVPTGAGV KPKAQVGAGA FAGIPGVGPF

     190       200       210       220       230       240
GGQQPGLPLG YPIKAPKLPA GYGLPYKTGK LPYGFGPGGV AGSAGKAGYP TGTGVGPQAA

     250       260       270       280       290       300
AAAAKAAAKL GAGGAGVLPG VGVGGPGIPG APGAIPGIGG IAGVGAPDAA AAAAAAAKAA

     310       320       330       340       350       360
KFGAAGGLPG VGVPGVGVPG VGVPGVGVPG VGVPGVGVPG VGVPGVGVPG VGVPGVGVPG

     370       380       390       400       410       420
VGVPGALSPA ATAKAAAKAA KFGARGAVGI GGIPTFGLGP GGFPGIGDAA AAPAAAAAKA

     430       440       450       460       470       480
AKIGAGGVGA LGGVVPGAPG AIPGLPGVGG VPGVGIPAAA AAKAAAKAAQ FGLGPGVGVA

     490       500       510       520       530       540
PGVGVVPGVG VVPGVGVAPG IGLGPGGVIG AGVPAAAKSA AKAAAKAQFR AAAGLPAGVP

     550       560       570       580       590       600
GLGVGAGVPG LGVGAGVPGL GVGAGVPGPG AVPGTLAAAK AAKFGPGGVG ALGGVGDLGG

     610       620       630       640       650       660
AGIPGGVAGV VPAAAAAAKA AAKAAQFGLG GVGGLGVGGL GAVPGAVGLG GVSPAAAAKA

     670       680       690       700       710       720
AKFGAAGLGG VLGAGQPFPI GGGAGGLGVG GKPPKPFGGA LGALGFPGGA CLGKSCGRKR
```

K

V12

10 20 30 40 50 60
VPGVGVPGVG VPGVGVPGVG VPGVGVPGVG VPGVGVPGVG VPGVGVPGVG VPGVGVPGVG

Y

V24

10 20 30 40 50 60
VPGVGVPGVG VPGVGVPGVG VPGVGVPGVG VPGVGVPGVG VPGVGVPGVG VPGVGVPGVG

70 80 90 100 110 120
VPGVGVPGVG VPGVGVPGVG VPGVGVPGVG VPGVGVPGVG VPGVGVPGVG VPGVGVPGVG

Y

S12-K_4-S12

10 20 30 40 50 60
GHHHHHHNGW GVPGSGVPGS GVPGSGVPGS GVPGSGVPGS GVPGSGVPGS GVPGSGVPGS

70 80 90 100 110 120
GVPGSGVPGS GGGKGGKGGK GGKGGVPGSG VPGSGVPGSG VPGSGVPGSG VPGSGVPGSG

130 140
VPGSGVPGSG VPGSGVPGSG VPGSGY

V12-K_4-V12

10 20 30 40 50 60
GHHHHHHNGW GVPGVGVPGV GVPGVGVPGV GVPGVGVPGV GVPGVGVPGV GVPGVGVPGV

70 80 90 100 110 120
GVPGVGVPGV GGGKGGKGGK GGKGGVPGVG VPGVGVPGVG VPGVGVPGVG VPGVGVPGVG

130 140
VPGVGVPGVG VPGVGVPGVG VPGVGY

V24-K_4-V24

```
      10        20        30        40        50        60
GHHHHHHNGW GVPGVGVPGV GVPGVGVPGV GVPGVGVPGV GVPGVGVPGV GVPGVGVPGV

      70        80        90       100       110       120
GVPGVGVPGV GVPGVGVPGV GVPGVGVPGV GVPGVGVPGV GVPGVGVPGV GVPGVGVPGV

     130       140       150       160       170       180
GVPGVGVPGV GGGKGGKGGK GGKGGVPGVG VPGVGVPGVG VPGVGVPGVG VPGVGVPGVG

     190       200       210       220       230       240
VPGVGVPGVG VPGVGVPGVG VPGVGVPGVG VPGVGVPGVG VPGVGVPGVG VPGVGVPGVG

     250       260
VPGVGVPGVG VPGVGVPGVG VPGVGY
```

CryS96 (αB-crystallin peptide fused with ELP S96)

```
      10        20        30        40        50        60
GDRFSVNLDV KHFSPEELKV KGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG

      70        80        90       100       110       120
SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG

     130       140       150       160       170       180
SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG

     190       200       210       220       230       240
SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG

     250       260       270       280       290       300
SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG

     310       320       330       340       350       360
SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG

     370       380       390       400       410       420
SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG

     430       440       450       460       470       480
SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG SGVPGSGVPG

     490       500
SGVPGSGVPG SGVPGSGVPG SGY
```

1.3 Copper staining SDS-PAGE gels

For samples requiring copper staining for detection, the following procedure was used after electrophoresis. Gels were rinsed in deionized water for 45 sec, stained with 10 % $CuCl_2$ (w/v) for 5 min with gentle shaking, and minimally rinsed again with water. Finally, gels were imaged using a BioRad Chemidoc Touch imaging system.

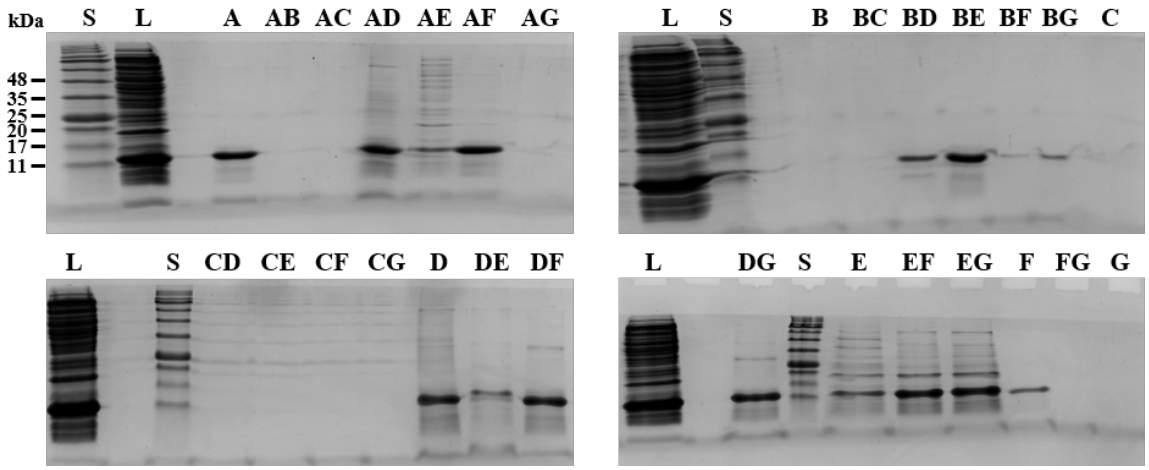

Figure S1. SDS-PAGE analysis of S12-K_4-S12 extraction screen. S = protein MW standards, L = clarified lysate. Extractants used are listed in Table 2 of the paper.

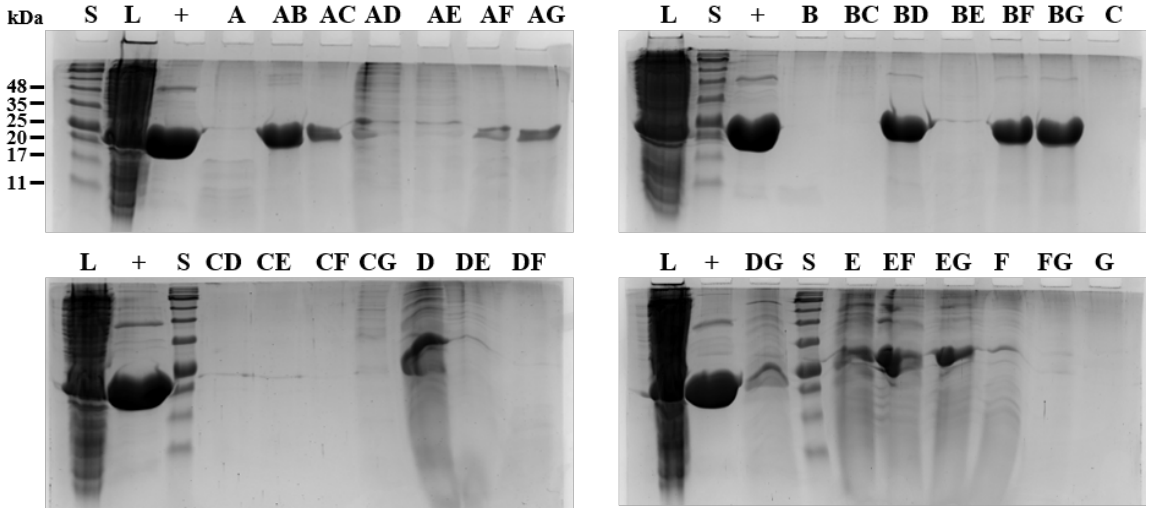

Figure S2. SDS-PAGE analysis of V24-K_4-V24 extraction screen. S = protein MW standards, L = clarified lysate, "+" = 3X ITC-purified V24-K_4-V24. Extractants used are listed in Table 2 of the paper.

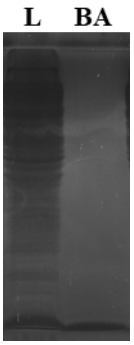

Figure S3. SDS-PAGE analysis of V24 extraction from lysate, visualized by copper staining. L = clarified lysate; BA = 1:1 nBuOH:IPA.

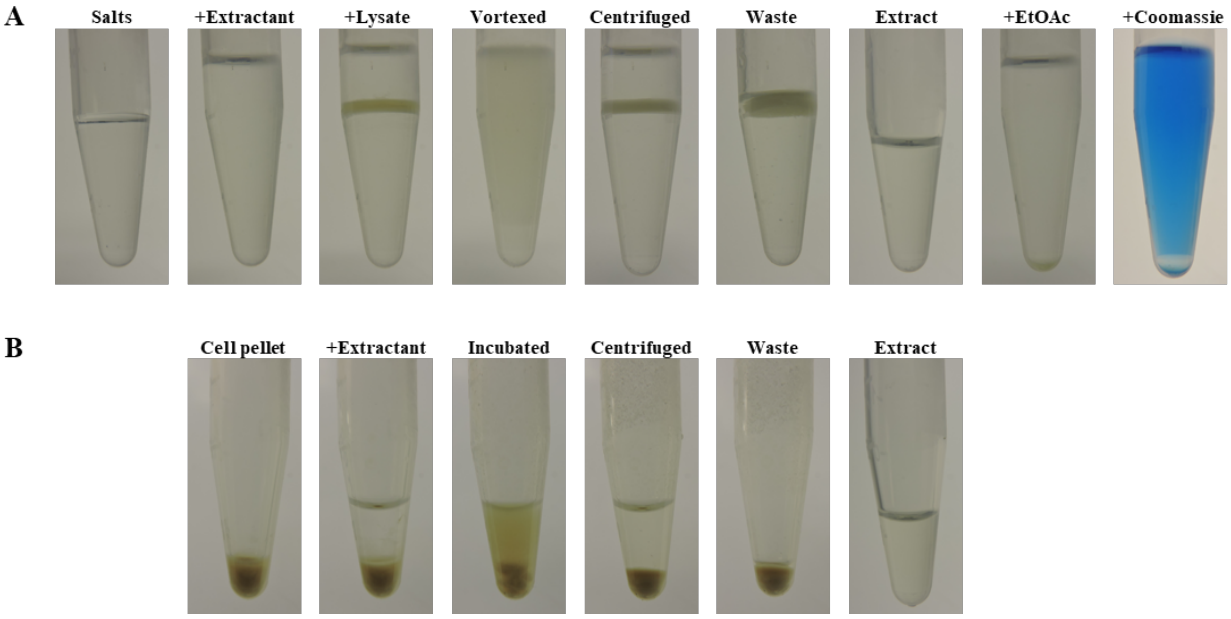

Figure S4. Photos depicting the steps of the non-optimized extraction from cell lysate (A) and whole cells (B). Coomassie Blue was added after back-extraction to enhance visibility of the aqueous subphase.

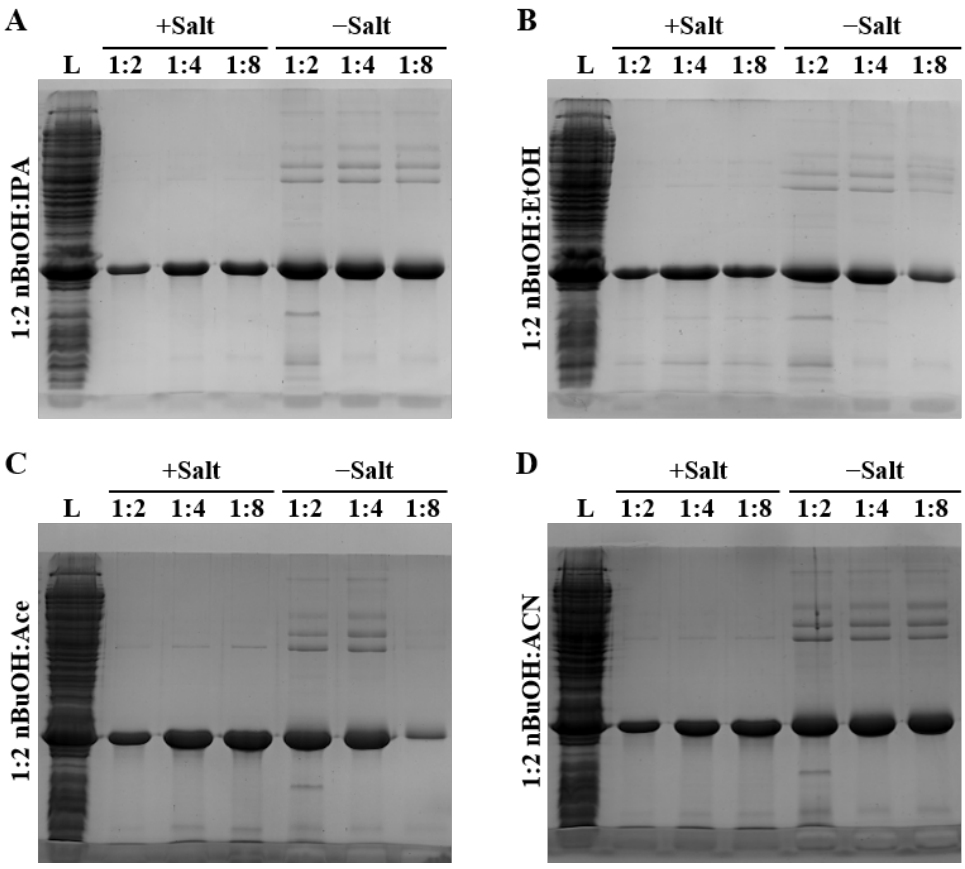

Figure S5. SDS-PAGE analysis showing extraction optimization of V24-K_4-V24 with varied lysate:extractant ratios and salt effects. L = clarified lysate.

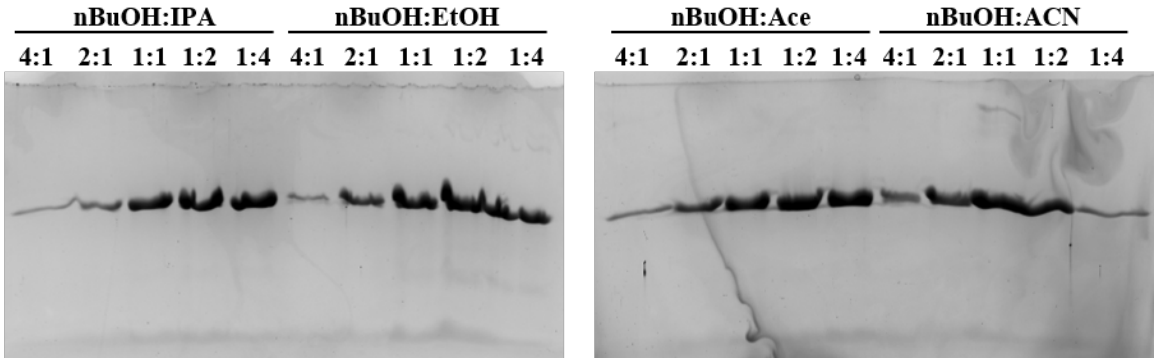

Figure S6. SDS-PAGE analysis of lead extractant blend ratio optimization for V24-K_4-V24. Blend ratios (v/v) are as indicated for each solvent combination.

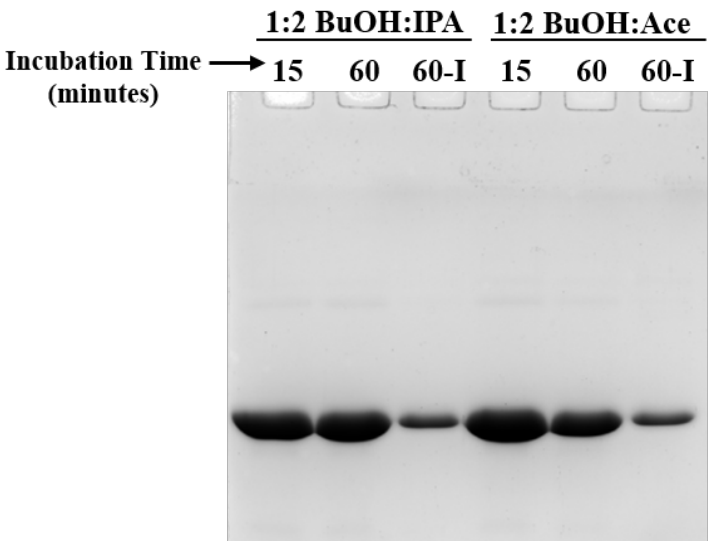

Figure S7. SDS-PAGE analysis of 15 vs 60 min direct extraction of V24-K_4-V24 from whole cells. 60-I = IPTG-induced cells.

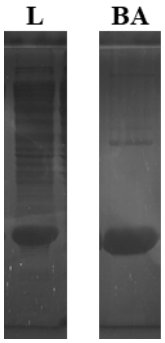

Figure S8. SDS-PAGE analysis of scaled-up V24-K_4-V24 extraction from lysate (35 mL, ~ 8.75 g wet pellet weight), visualized by copper staining. L = clarified lysate; BA = 1:2 nBuOH:IPA.

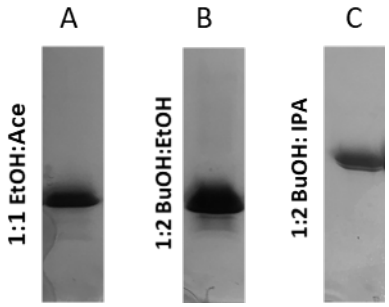

Figure S9. SDS-PAGE analysis of scaled-up (A) S12-K_4-S12, (B) V12-K_4-V12, and (C) V24-K_4-V24 extractions from whole cells with extraction blend ratios (2.5 g of cell pellet).

Table S1. Comparison of ELP Recovery Methods.

Method	Benefits	Limitations
ITC	High recovery potential	Limited to ELP with convenient T_t; requires aqueous solution
Back-extraction	High efficiency; solvent exchange to aqueous solution; target never dried	Care must be taken to avoid target precipitation; process may require optimization
Passive evaporation	High efficiency	Relatively slow process, depending on solvent; may lose structure/function upon drying
Rotary evaporation	Highly scalable for large volumes; high efficiency	May lose structure/function upon drying; may be slow, depending on solvent
Vacuum drying	High throughput; high efficiency	Limited utility with highly volatile solvents; may lose structure/function upon drying
Speed vac	High throughput; high efficiency	May lose structure/function upon drying
Lyophilization	High throughput; high efficiency	Slow; requires transition to aqueous solution; may lose structure/function without the use of cryoprotectants

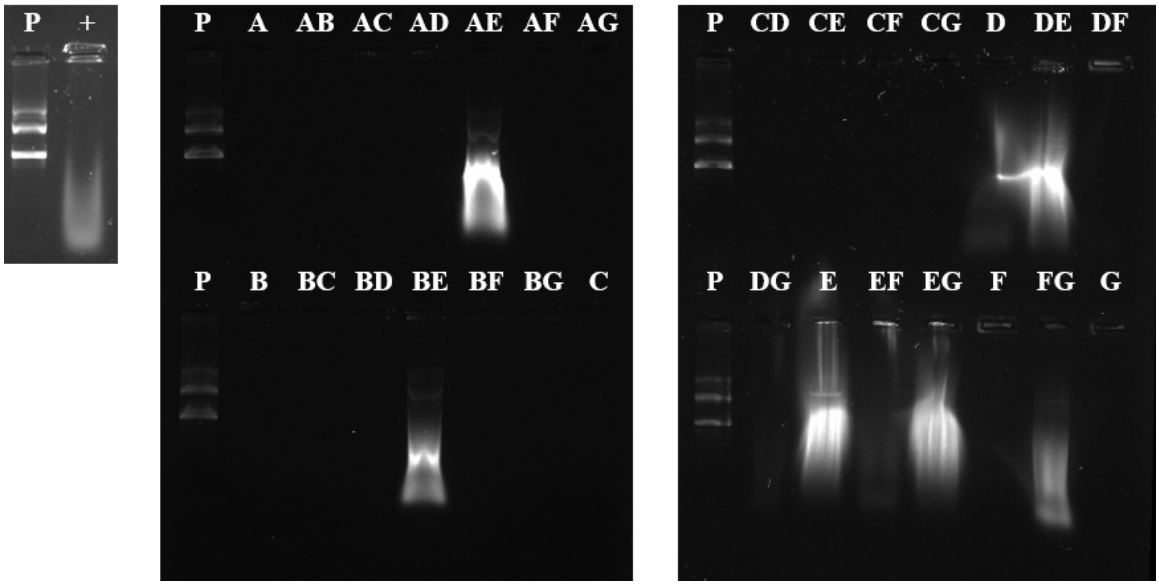

Figure S10. Nucleic acid contamination of V24-K_4-V24 extraction from lysates assessed by agarose gel electrophoresis and stained with GelRed. P = plasmid DNA control; "+" = 3X ITC-purified V24-K_4-V24. Extractants employed are listed in Table 2 of the paper.

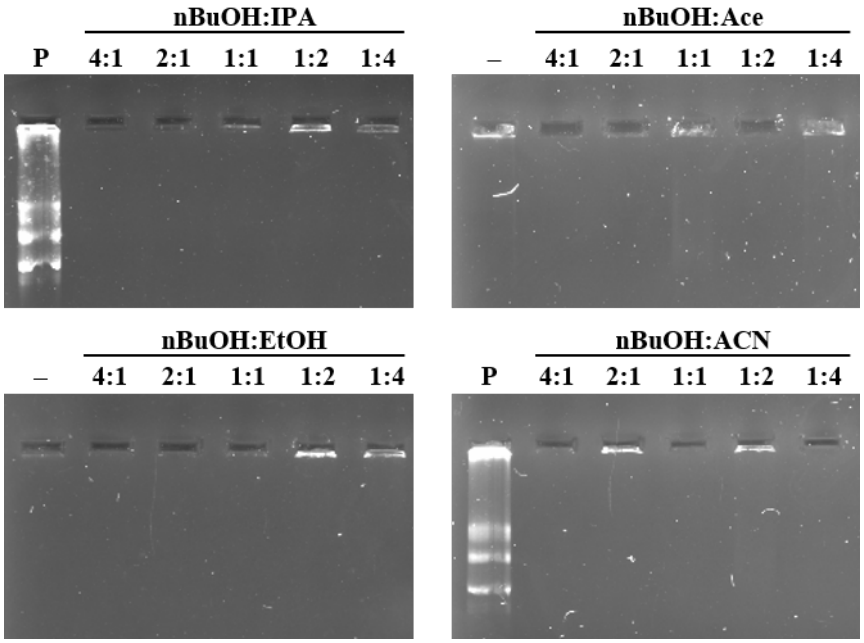

Figure S11. Agarose gel electrophoresis assessment of nucleic acid contamination of optimized extractant blends for V24-K_4-V24 purification from lysates. Blend ratios (v/v) are as indicated for each solvent combination. P = plasmid DNA; "–" = empty lane.

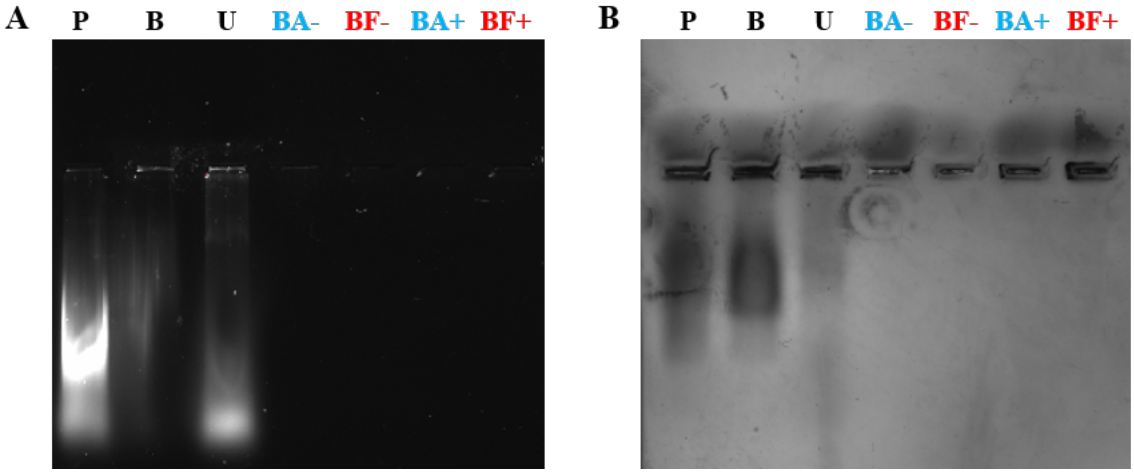

Figure S12. Agarose gel electrophoresis of V24-K_4-V24 extracted directly from whole cells. (A) Nucleic acids are visualized by GelRed staining and (B) proteins by Coomassie R-250 staining. P = probe tip sonication lysate, B = B-PER lysate, U = urea lysate, BA = 1:2 nBuOH:IPA, BF = 1:2 nBuOH:Ace, "–" = without salt, "+" = with salts.

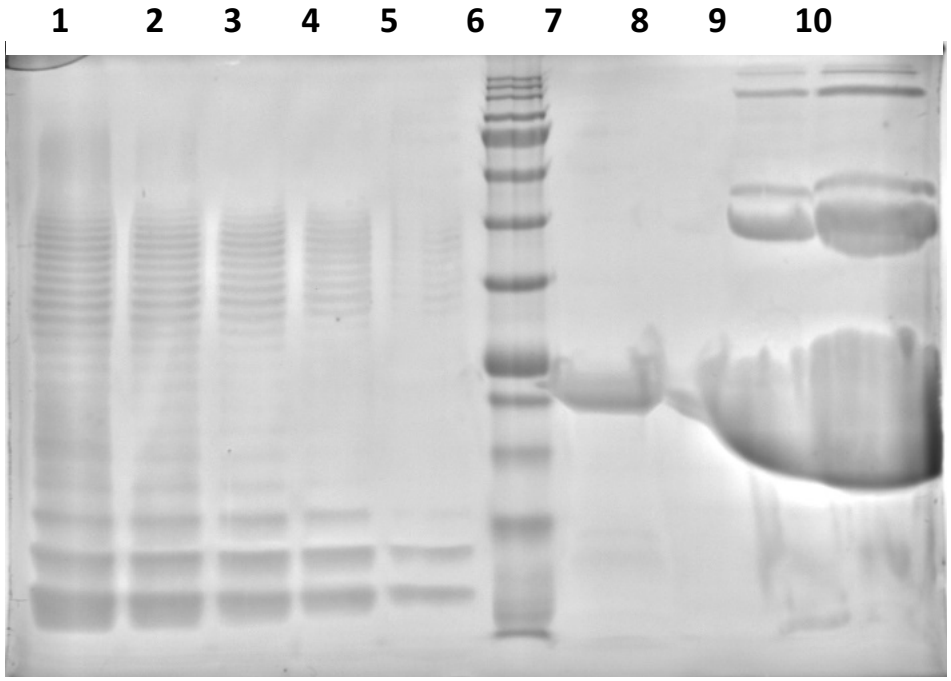

Figure S13. SDS-PAGE analysis of LPS content in V24-K_4-V24 extracted from lysates with 1:2 nBuOH:EtOH. Lanes 1-5 are a serial dilution of commercial LPS from 500 ng-31.25 ng. Lane 7, 9, and 10 contain 5, 30, and 50 μg of V12-K4-V12 extract, respectively.

Table S2. Organic Solvents and Properties.

Solvent	Notation	Formula	MW	Density (g/mL)[a]	Solubility in Water (g/100g)[b]	Dielectric Constant[c]	Relative Solvent Polarity[d]	Polarity Parameter (kcal/mol)[e]	Polarity Index[f]	Log *P*[g]
Acetone	Ace	$(CH_3)_2CO$	58.09	0.784	Miscible	21.0	0.355	42.2	5.1	-0.24
Acetonitrile	ACN	CH_3CN	41.06	0.782	Miscible	36.6	0.460	45.6	5.8	-0.34
n-Butanol	nBuOH	C_4H_9OH	74.14	0.806	6.32	17.8	0.586	49.7	3.9	0.88
sec-Butanol	sBuOH	$CH_3CHOHCH_2C$ H	74.14	0.802	18.1	17.3	0.506	47.1	–	0.61
tert-Butanol	tBuOH	$(CH_3)_3COH$	74.14	0.8	Miscible	12.5	0.389	43.3	3.9	0.35
Diethyl ether	Et_2O	$(C_2H_5)_2O$	74.14	0.708	6	4.27	0.117	34.5	2.8	0.89
Ethanol	EtOH	C_2H_5OH	46.07	0.794	Miscible	25.3	0.654	51.9	5.2	-0.31
Ethyl acetate	EtOAc	$C_4H_8O_2$	88.12	0.895	8	6.08	0.228	38.1	4.4	0.73
MeOH	MeOH	CH_3OH	32.05	0.786	Miscible	33.0	0.762	55.4	5.1	-0.77
Isopropanol	IPA	$CH_3CHOHCH_3$	60.11	0.786	Miscible	20.2	0.546	48.4	3.9	0.05
Water	H_2O	H_2O	18.01	1	–	80.1	1.00	63.1	10.2	-1.38

[a] Density (specific gravity). [3,3,3,3,3]
[b] Solvent solubility in water. [3,3,3,3,3]
[c] Dielectric constant (relative permittivity, ε). [4,4,4,4,4]
[d] Relative solvent polarity (E_T^N) measured at 25 °C, except for tBuOH at 30 °C. [5,5,5,5,5]
[e] Polarity parameter (E_T). [3,3,3,3,3]
[f] Polarity Index (P') measure at 25 °C. [6,6,6,6,6]
[g] Partition coefficient of octanol/water (Log P_{ow}). [7,7,7,7,7]

1 J. R. McDaniel, J. A. MacKay, F. G. Quiroz and A. Chilkoti, *Biomacromolecules*, 2010, **11**, 944-952.
2 D. E. Meyer and A. Chilkoti, *Biomacromolecules*, 2002, **3**, 357-367.
3 G. Wypych, ed., *Knovel Solvents - A Properties Database*, ChemTec Publishing, 2008; 2012.
4 *CRC Handbook of Chemistry and Physics*, CRC Press/Taylor & Francis, Boca Raton, FL, 2017.
5 C. Reichardt and T. Welton, eds., *Solvents and Solvent Effects in Organic Chemistry*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2010.
6 L. R. Snyder, *Journal of chromatography. A*, 1974, **92**, 223-230.
7 C. L. Yaws, ed., *Yaws' Handbook of Thermodynamic and Physical Properties of Chemical Compounds*, Knovel, 2003.