Supplementary Methods for Clark, et. al. "Design, Synthesis and Selection of DNA-Encoded Small Molecule Libraries"

Table of Contents

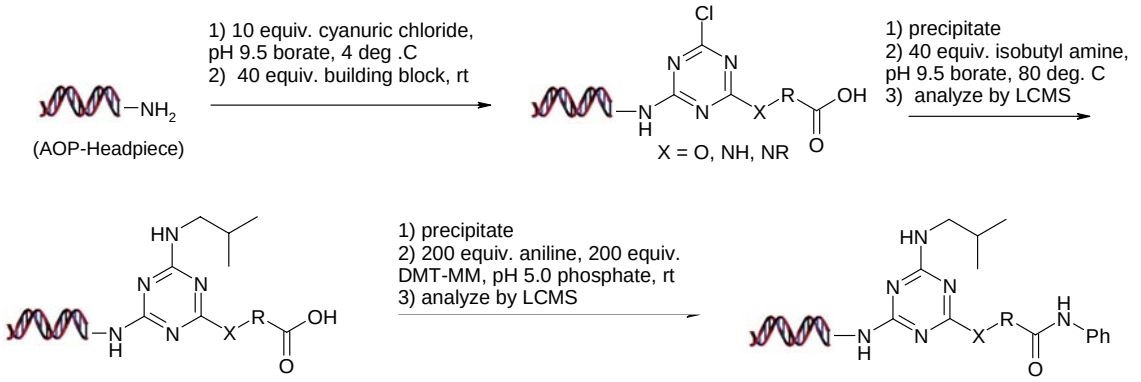
Analytical Methods

On-DNA reactions conducted during validation, library synthesis, or single compound synthesis were analyzed by LCMS. Samples (ca. 100 pmol) were injected onto a reverse-phase chromatography column (Targa C18, 5μ, 2.1 x 40 mm) and eluted (15 – 70% solvent B over 7 minutes, 0.36 mL/min flow rate; Solvent A: 0.75% hexafluoroisopropanol / 0.38% triethylammonium acetate /10 μM EDTA in deionized water; Solvent B: 0.75% HFIP/0.38% TEAA/10 μM EDTA in 90/10 methanol/water) with monitoring at 260 nm. Effluent was analyzed on a ThermoFinnigan Advantage electrospray mass spectrometer in negative ion mode. When necessary, mass deconvolution was achieved using ProMass software (Novatia). Chromatographic purification was likewise achieved using reverse-phase liquid chromatography (Gemini C18 5μ, 30 x 100 mm; Solvent A: pH 7.5 50 mM triethylammonium acetate; Solvent B: 1% water in acetonitrile).

Nature Chemical Biology: doi:10.1038/nchembio.211

Synthon Validation

All synthons were validated in 96-well plates using library-like conditions. Analysis was performed using LCMS as described above. The validation of DEL-B cycle 2 synthons will be used as an example. According to the scheme for DEL-B, the synthons in Cycle 2 needed to participate in two reactions: triazine substitution and amide bond formation. Therefore, a validation scheme that would assess both modes of reactivity was designed:

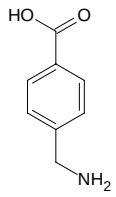
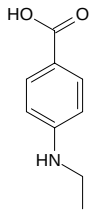
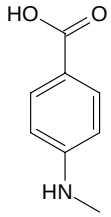
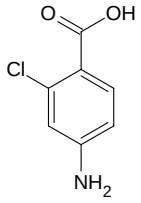
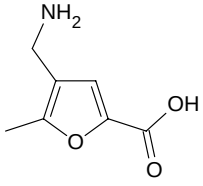


The validation process ran as follows: A 1 mM solution of AOP-HP in 250 mM borate pH 9.5 was arrayed into a 96 well plate (20 nmol/well, 20 µL) and cooled to 4 degrees C. To each well was then added cyanuric chloride (10 equiv of a 200 mM MeCN stock). After one hour, each well received 50 equivs of the candidate amino acid or phenolic acid synthons (200 mM stocks in MeCN or DMA). The substitution reaction was allowed to proceed for 16 h at 4 degrees C, at which point the DNA was precipitated by addition of 5 M NaCl (10% by volume) and 2.5 volumes of cold ethanol. The DNA pellets were redissolved in 20 µL 250 mM borate pH 9.5. Isobutyl amine (50 equiv from a 200 mM MeCN stock) was added and the plate heated to 80 deg. C for 6 h. An aliquot was removed from each well for LCMS analysis. Plates were then precipitated a final time, and the DNA pellets redissolved in 20 µL of 250 mM phosphate pH 5.5. To each well was added 50 equiv of aniline (200 mM in 1:1 MeCN/H_2O), 100 equiv. of HCl (as 200 mM stock) and 100 equiv. of DMT-MM (200 mM in water). At 4, 18, and 30 h intervals, additional 50 equiv. aliquots of DMT-MM were added to each well. After 48 h total reaction time, an additional aliquot was removed for LCMS analysis.

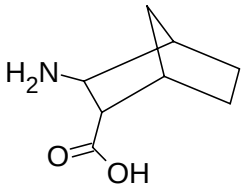
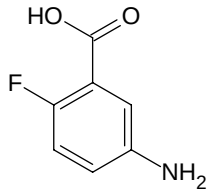
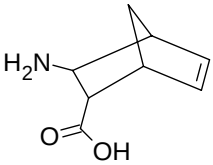
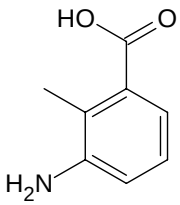
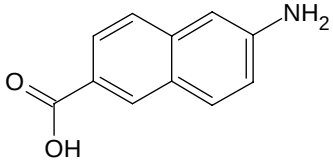
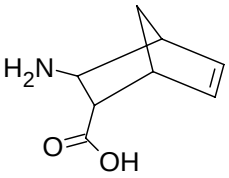
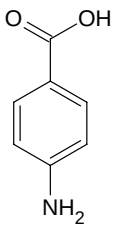
Yields were calculated by examination of the UV and TIC traces of the LCMS chromatograms. Side products like double substitution and chloride hydrolysis were identified and excluded from the product percentage. In addition, any products in which the DNA was damaged by gain or loss of mass, such as depurination, phosphate hydrolysis, or covalent modification, were likewise excluded from the product percentage (such products were seen very rarely). Generally, only synthons which gave >70% yield in both the substitution and amidation reactions (aggregate yield ≥50%) were allowed into the library. A total of 75 amino, phenolic, and heterocyclic acids were validated to yield the 32 used in the library. It should be noted that, because mass is the only means of

Nature Chemical Biology: doi:10.1038/nchembio.211

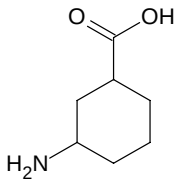
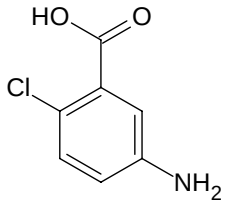
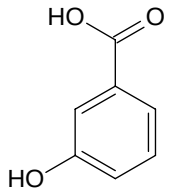
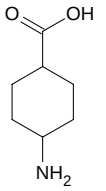
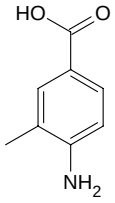
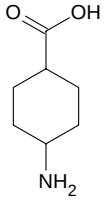
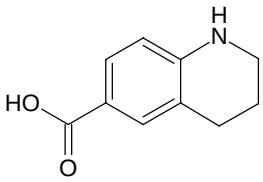
analysis, the validation procedure provided no information about stereo- or regioselectivity when applicable. Indeed, as seen in the p38 DEL-B selection, regioisomerism was observed in the off-DNA reactivity of benzimidazole-5-carboxylic acid. While isomerism of this type complicated the post-selection chemistry campaign, it did not compromise the ligand-generating ability of the library.

Structure	Yield for Triazine Substitution	Yield for Amidation	2-Step Yield	In DEL-B?
	98	97	95	y
	94	99	93	y
	98	93	91	y
	97	93	90	y
	89	99	88	y

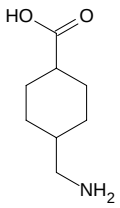
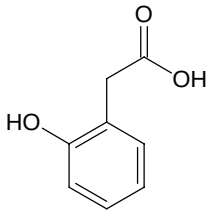
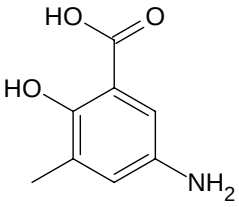
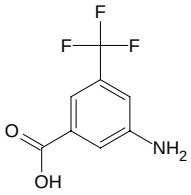
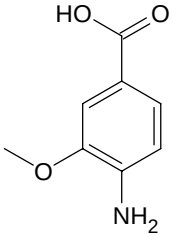
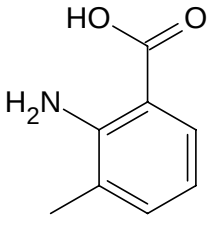
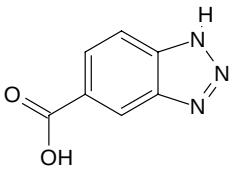
Nature Chemical Biology: doi:10.1038/nchembio.211

	98	90	88	y
	97	90	87	y
	98	88	86	y
	94	90	85	y
	93	89	83	y
	97	85	83	y
	98	84	82	y

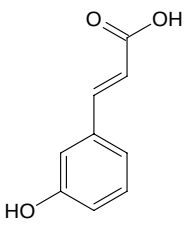
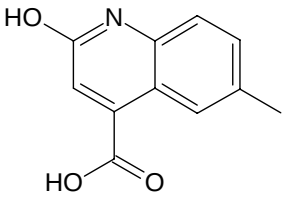
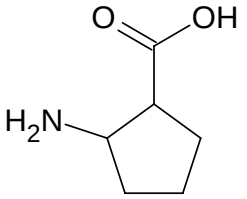
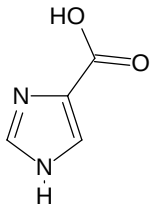
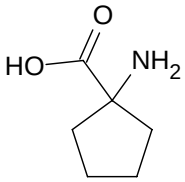
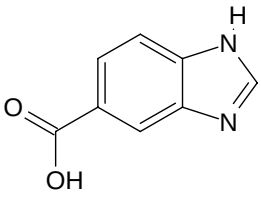
Nature Chemical Biology: doi:10.1038/nchembio.211

	98	83	81	y
	94	85	80	y
	85	94	80	y
	98	81	80	y
	93	85	79	y
	90	87	78	y
	83	91	75	y

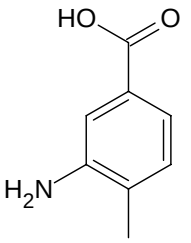
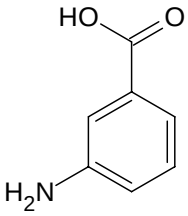
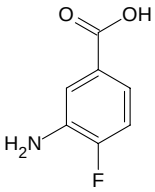
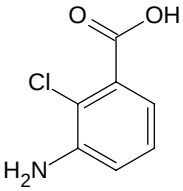
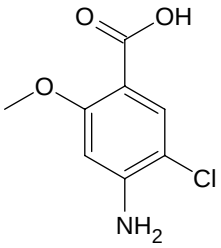
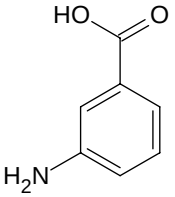
Nature Chemical Biology: doi:10.1038/nchembio.211

	99	76	75	y
	85	85	72	y
	89	78	69	y
	78	87	68	y
	95	71	67	y
	85	70	60	y
	85	70	60	y

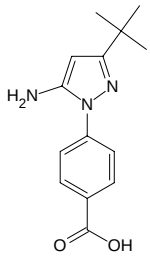
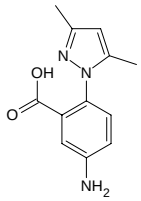
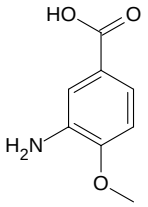
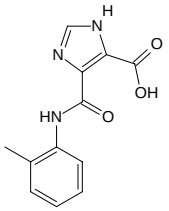
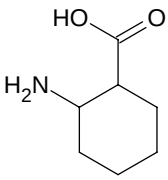
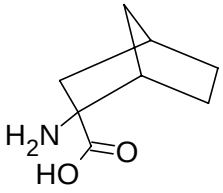
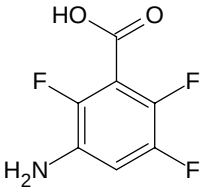
Nature Chemical Biology: doi:10.1038/nchembio.211

	77	77	59	y
	93	64	59	y
	96	60	58	y
	75	70	53	y
	84	60	50	y
	59	75	45	y

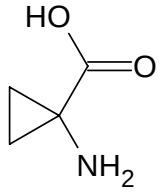
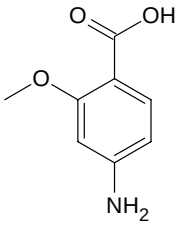
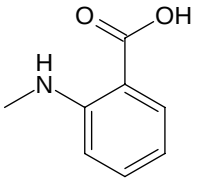
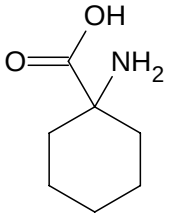
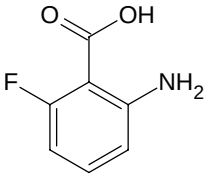
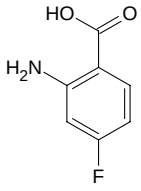
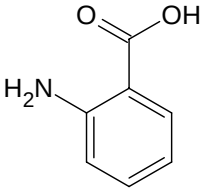
Nature Chemical Biology: doi:10.1038/nchembio.211

	97	86	83	n
	99	80	79	n
	96	60	58	n
	67	78	53	n
	58	83	48	n
	56	83	47	n

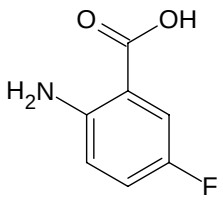
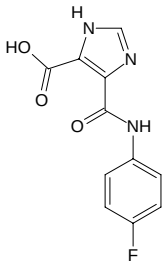
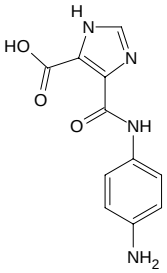
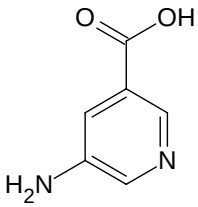
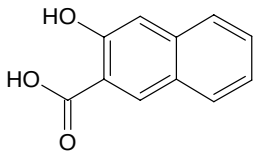
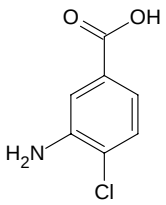
Nature Chemical Biology: doi:10.1038/nchembio.211

	84	1	1	n
	92	1	1	n
	94	45	43	n
	77	50	39	n
	97	39	38	n
	81	42	34	n
	77	41	32	n

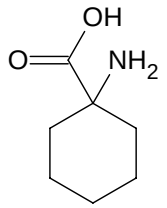
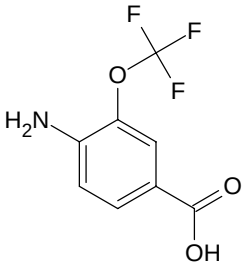
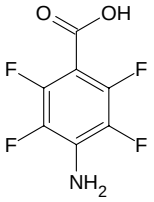
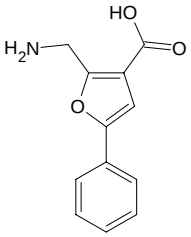
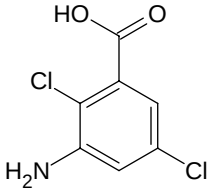
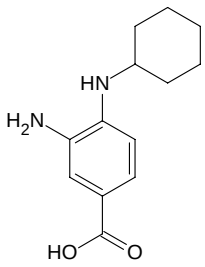
Nature Chemical Biology: doi:10.1038/nchembio.211

	35	85	30	n
	85	35	30	n
	99	30	30	n
	53	50	27	n
	95	13	12	n
	95	7	7	n
	99	6	6	n

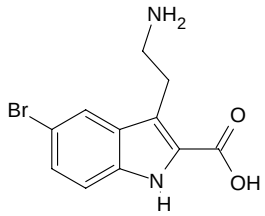
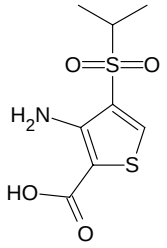
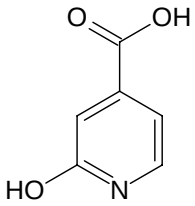
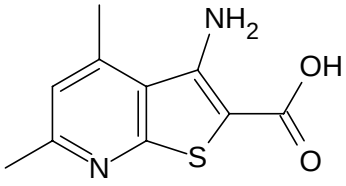
Nature Chemical Biology: doi:10.1038/nchembio.211

	61	7	4	n
	94	4	4	n
	73	4	3	n
	61		0	n
	55		0	n
	53		0	n

Nature Chemical Biology: doi:10.1038/nchembio.211

	49		0	n
	33		0	n
	6		0	n
	6		0	n
	6		0	n
	4		0	n

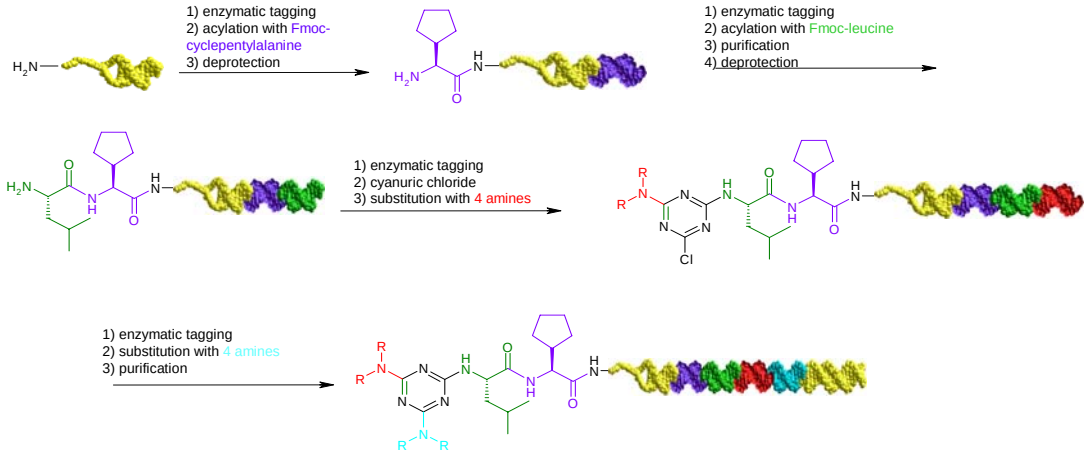
Nature Chemical Biology: doi:10.1038/nchembio.211

	3		0	n
	1		0	n
	1		0	n
	1		0	n

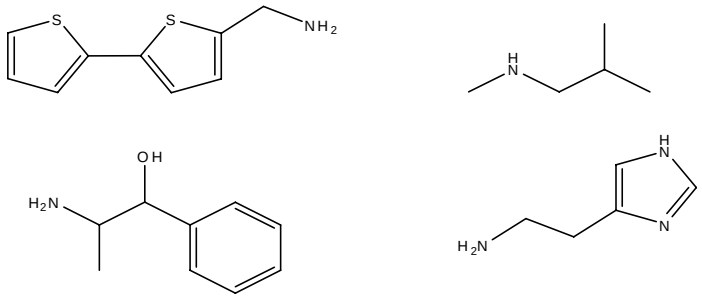
Nature Chemical Biology: doi:10.1038/nchembio.211

Test Library

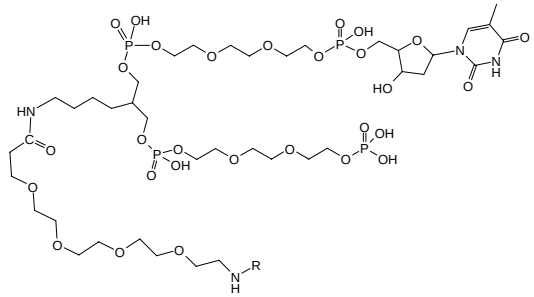
Further validation of our library synthetic plan was achieved through the synthesis of a small test library. A library similar to DEL-A, in this case incorporating two amino acid residues before the triazine rather than the one residue of DEL-A, was synthesized with a total diversity of 10 compounds. The same 4 amines were used in the last two cycles of the synthesis, generating the diversity.



<u>Amine set for Cycles 3 & 4:</u>

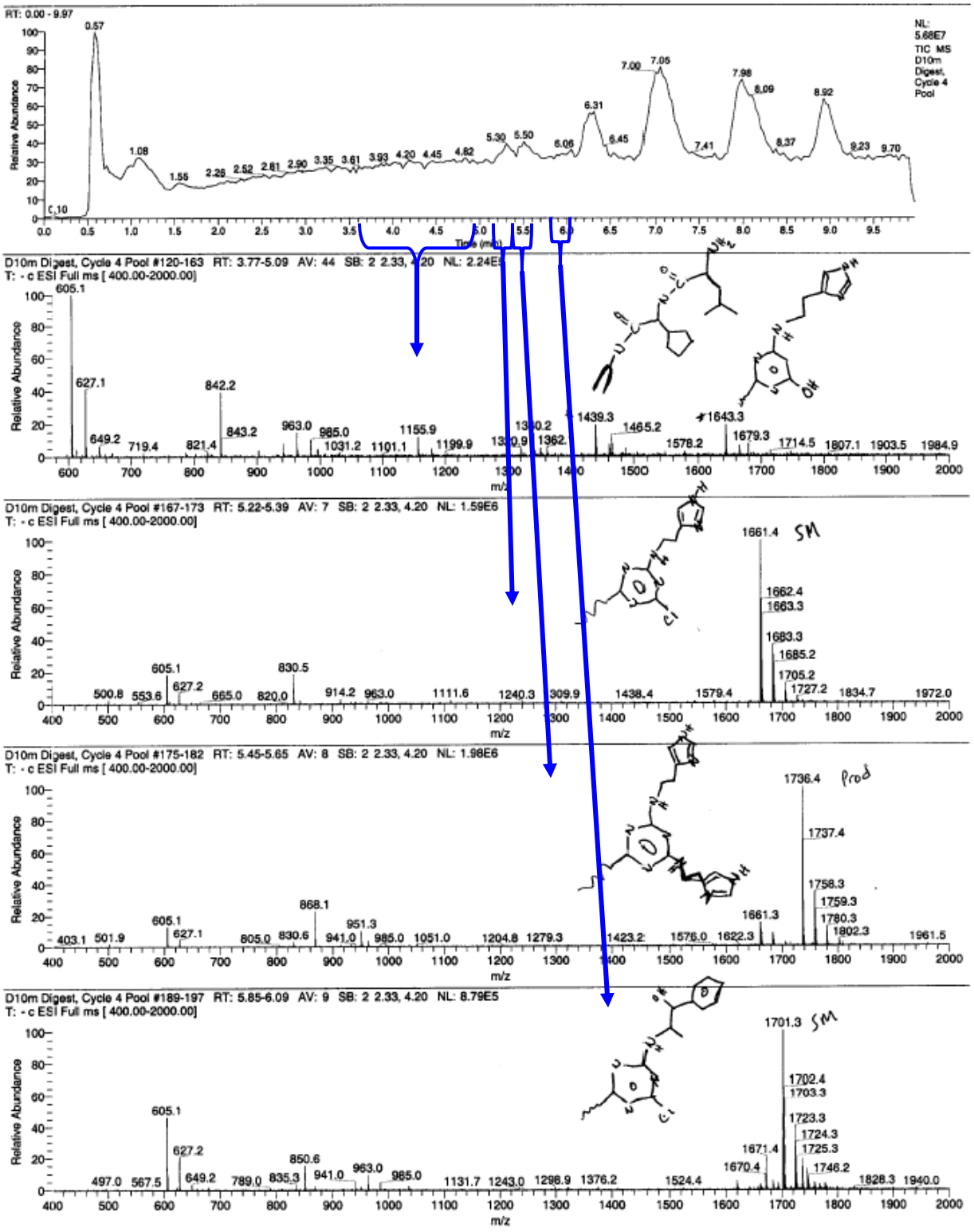


After completion of the test library synthesis, the DNA portion of the library was digested using a combination of DNase and S1 Nuclease (Invitrogen) as per the manufacturer's instructions. This process left the library products attached to the linker and a single nucleotide:

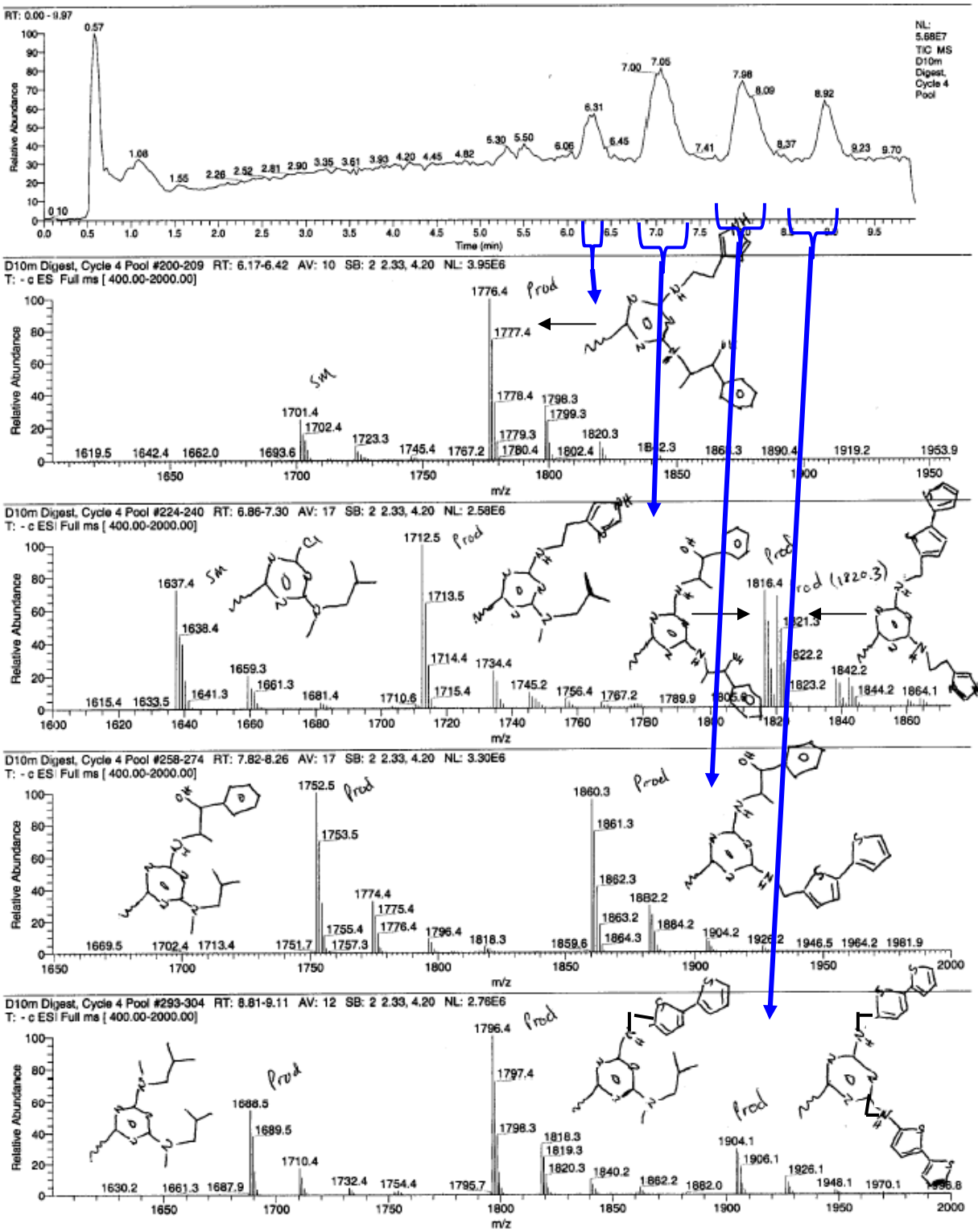


Nature Chemical Biology: doi:10.1038/nchembio.211

The digested library sample was analyzed by LCMS. All ten of the expected masses (and their Na adducts) were observed, as well as several expected byproducts:



Nature Chemical Biology: doi:10.1038/nchembio.211



Nature Chemical Biology: doi:10.1038/nchembio.211

Library Synthesis Starting Materials

Materials. Chemical building blocks (Fmoc-protected and unprotected amino acids; amines) were obtained from commercial sources, except for AMMB, which was synthesized in-house. Building blocks were dissolved in an appropriate solvent before use; amino acids in DMF, amines in either DMA or 1:1 MeCN/water, depending on solubility. DNA headpiece (5′-/5Phos/GAGTCA/iSp9/iUniAmM/iSp9/TGACTCCC-3′) and the various DNA tags were obtained from IDT, Inc., Coralville, IA and Biosearch Technologies, Novato, CA. T4 DNA Ligase was obtained from Fermentas (30 U/ul). The 10X ligation buffer stock used in ligation reactions was composed as follows: 500 mM Tris pH 7.5, 500 mM NaCl, 100 mM $MgCl_2$, 100 mM DTT, 25 mM ATP.

"Headpiece." Sequence: 5′-/5Phos/GAGTCA/iSp9/iUniAmM/iSp9/TGACTCCC-3′

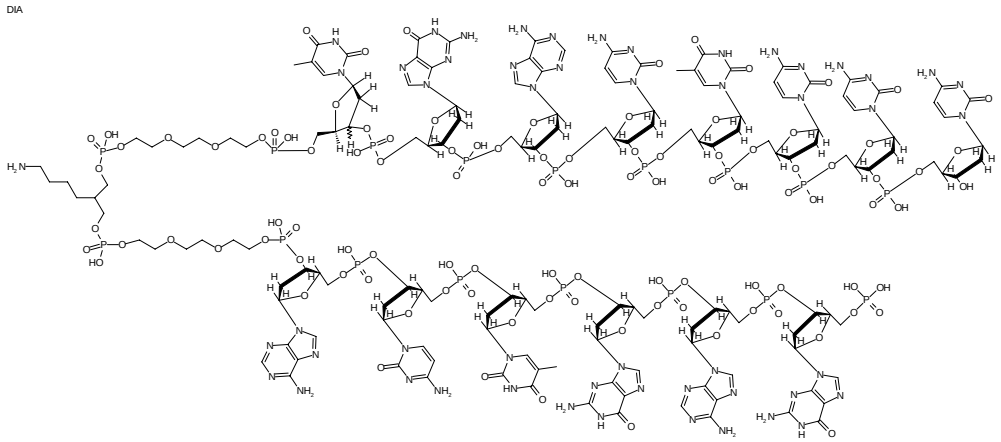


Figure 1. Sequence and structure of the "headpiece." MW = 4937 D

Installation of chemical spacer. A solution of headpiece DNA (43 μmol in 43 mL 250 mM pH 9.5 borate buffer) was treated with 40 equivalents of Fmoc-15-amino-4,7,10,13-tetraoxapentadecanoic acid ("AOP") (8.6 mL of a 200 mM DMF solution) followed by 40 equivalents of 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride (DMT-MM, Acros) (8.6 mL of a 200 mM aqueous solution). The acylation reaction was allowed to proceed for 18 hours at room temperature. After completion, the reaction was precipitated by addition of 5 M NaCl (10% by volume) and 2.5 volumes of cold ethanol, followed by centrifugation. The lyophilized pellet was then deprotected by exposure to 20 mL of 10% piperidine in water. The deprotected product was precipitated with ethanol and purified by reverse-phase HPLC.

Nature Chemical Biology: doi:10.1038/nchembio.211

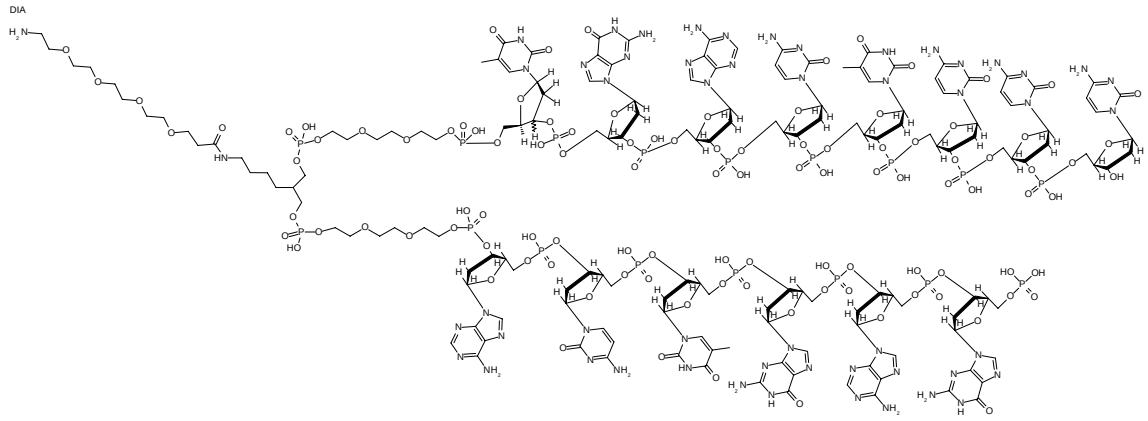


Figure 2. Structure of AOP-Headpiece. MW = 5184 D

Synthesis of DEL-A

DEL A Primer Ligation. The AOP-headpiece (20 μmol) was dissolved in 18 mL water. Primer duplex was added (28 mL of a 1mM solution in water, 1.4 equivalents), followed by water (25 mL) and 8 mL 10X ligation buffer. The solution was heated to 95° C for 1 min, then cooled to 16° C over 10 min. A solution of T4 DNA ligase (800 μL, 30U/μL) was then added and the ligation incubated at 16° C for 16 h. The DNA product was precipitated with NaCl and ethanol and taken on to the first cycle of library synthesis without further purification.

Sequence of primer (MW = 12458 D):

5’ AAATCGATGTGGTCAGGAAG 3′
3′ GGTTTAGCTACACCAGTCCT 5′

Sequence of the extended headpiece (MW = 17606 D):

TGACTCCCAAATCGATGTGGTCAGGAAG 3′
ACTGAGGGTTTAGCTACACCAGTCCT 5′

Tags. The DNA tags contained a 7bp coding region, flanked by two 2-base 3′ overhangs: All 5′-ends were phosphorylated.

	Cycle 1	Cycle 2	Cycle 3	
5′	XXXXXXXGT	XXXXXXXGA	XXXXXXXTT	3′
3′	TCXXXXXXX	CAXXXXXXX	CTXXXXXXX	5′

Nature Chemical Biology: doi:10.1038/nchembio.211

DEL A Cycle 1. A 1 mM solution of the primer-elongated AOP headpiece (19.2 μmol, 19.2 mL) was split into 384 wells (50 nmol/well). To each well was then added 20 μL of 10x ligation buffer, 2.0 μL T4 DNA ligase (Fermentas), and 25 μL water. An aliquot of 1 of 384 tag solutions (100 μL of 1 mM stock solutions in water) was added to each well, and ligation allowed to proceed at 16 degrees C for 16 h. After ligation, 5 M NaCl (10% by volume) and 2.5 volumes of cold ethanol were added to each vessel to precipitate the DNA. The DNA pellets recovered after centrifugation were each dissolved in 50 μL pH 9.5 150 mM borate buffer. The plates were cooled to 4 degrees C, and to each well was added 40 equivalents of 1 of 192 Fmoc-protected amino acids (12.6 μL of a 150 mM DMF solution), followed by 40 equivalents of DMT-MM (7.6 μL of a 250 mM water solution). The acylation reactions were allowed to proceed for 18 hours at 4 degrees C. After completion, the reactions were pooled, precipitated with ethanol, and purified by reverse-phase HPLC, yield = 7.4 μmol . The lyophilized product was then deprotected by exposure to 20 mL of 10% piperidine in water for 4 h. The deprotected product was precipitated. Since 384 tags were used, but only 192 chemical building blocks, each building block was encoded by two different tags.

DEL A Cycle 2. The Cycle 1 product (7.4 μmol) was dissolved in 7.4 mL water and split into 384 wells (19.2 μL/well). To each well was then added 7.8 μL of 10X ligation buffer, 0.77 μL T4 DNA ligase, and 10.9 μL water. DNA tags were then added (38.4 μL of 1 mM stocks in water) and the ligations were allowed to proceed at 16 degrees C for 16 h. The DNA was precipitated as above, and pellets dissolved in 19.2 μL of 150 mM pH 9.5 borate buffer per well. The library was cooled to 4 degrees C. To each well was added 10 equivalents of cyanuric chloride (1 μL of a 200 mM stock in acetonitrile). After 1 h, each well received 50 equivalents of an amine (4.8 μL of a 200 mM stock in acetonitrile or dimethylacetamide). The substitution was allowed to proceed for 16 h at 4 degrees C. The library was then pooled and precipitated, and taken into the next cycle without purification. A total of 192 amines and 384 tags were used, so that each synthon corresponded to 2 different DNA tags.

DEL A Cycle 3. The Cycle 2 product (7.4 μmol) was dissolved in 7.4 mL water and split into 192 wells (38.4 μL/well). To each well was then added 15 μL of 10X ligation buffer, 1.5 μL T4 DNA ligase, and 21 μL water. DNA tags were then added (95 μL of 1 mM stocks in water, 2.5 eq's of tags) and the ligations were allowed to proceed at 16 degrees C for 16 h. The DNA was precipitated as above and dissolved in 38.4 μL of 150 mM pH 9.5 borate buffer per well. To each well was added 45 equivalents of amine (9 μL of a 200 mM stock in acetonitrile or dimethyl acetamide). A total of 192 amines and 192 tags were used. The substitution was allowed to proceed for 6 h at 80 degrees C. The library was then pooled, precipitated, and purified to give 3.9 μmol of product (19% final yield). This quantity is sufficient for 780 selection experiments @ 5 nmol/selection.

Nature Chemical Biology: doi:10.1038/nchembio.211

Synthesis of DEL-B

Tags. Tags contained a 7bp coding region, flanked by two 2-base 3′ overhangs: All 5′-ends were phosphorylated. Cycle 1 tags contained an invariant primer region in addition to the 7-base coding region.

Cycle 1

5′ AAATCGATGTGXXXXXXXAG 3′
3′ GGTTTAGCTACACXXXXXXX 5′

Cycle 2	Cycle 3	Cycle 4	
5′ XXXXXXXGT	XXXXXXXGA	XXXXXXXTT	3′
3′ TCXXXXXXX	CAXXXXXXX	CTXXXXXXX	5′

Cycle 1. A 1 mM solution of the AOP headpiece (43 μmol, 43 mL) in water was split into 768 wells (54 nmol/well). To each well was then added 20 μL of 10x ligation buffer, 2.0 μL T4 DNA ligase (Fermentas), and 28 μL water. An aliquot of 1 of 384 tag solutions (110 μL of 1 mM stock solutions in water) was added to each well, and ligation allowed to proceed at 16 degrees C for 16 h. After ligation, 5 M NaCl (10% by volume) and 2.5 volumes of cold ethanol were added to each vessel to precipitate the DNA. The DNA pellets recovered after centrifugation were each dissolved in 50 μL 150 mM borate pH 9.5. The plates were cooled to 4 degrees C, and to each well was added 40 equivalents of 1 of 192 Fmoc-protected amino acids (13.3 μL of a 150 mM DMF solution), followed by 40 equivalents of DMT-MM (8 μL of a 250 mM water solution). The acylation reactions were allowed to proceed for 18 hours at 4 degrees C. After completion, the reactions were pooled, precipitated with ethanol, and purified by reverse-phase HPLC. The lyophilized product was then deprotected by exposure to 20 mL of 10% piperidine in water. The deprotected product was precipitated. Since 384 tags were used, but only 192 chemical building blocks, each building block was encoded by two different tags.

Cycle 2. The Cycle 1 product (23.3 μmol) was dissolved in 23.3 mL water and split into 384 wells (58 nmol/well). To each well was then added 23.3 μL of 10X ligation buffer, 2.3 μL T4 DNA ligase, and 32.6 μL water. DNA tags were then added (116.2 μL of 1 mM stocks in water) and the ligations were allowed to proceed at 16 degrees C for 16 h. The DNA was precipitated as above, and pellets dissolved in 58 μL of 150 mM borate pH 9.5 per well. The plates were cooled to 4 degrees C. To each well were added 10 equivalents of cyanuric chloride (3 μL of a 200 mM stock in acetonitrile). After 1 h, each well then received 50 equivalents of an amino acid or phenolic acid (15μL of a 200 mM stock in dimethylacetamide) and was removed from cold. The substitution was allowed to

Nature Chemical Biology: doi:10.1038/nchembio.211

proceed for 6 h at 45 degrees C. The library was then pooled and precipitated, and taken to the next step without purification. A total of 32 synthons and 384 DNA tags were used, so that each synthon was encoded by 12 different DNA tags.

Cycle 3. The Cycle 2 product (23.3 μmol) was dissolved in 23.3 mL water and split into 360 wells (64.7 nmol/well). To each well was then added 26 μL of 10X ligation buffer, 2.6 μL T4 DNA ligase, and 36.2 μL water. DNA tags were then added (162 μL of 1 mM stocks in water, 2.5 eq's of tags) and the ligations were allowed to proceed at 16 degrees C for 16 h. The DNA was precipitated as above and dissolved in 60 μL of 150 mM borate pH 9.5 per well. To each well were added 46 equivalents of amine (15 μL of a 200 mM stock in acetonitrile or dimethylacetamide). A total of 340 amines were used. The substitution was allowed to proceed for 6 h at 80 degrees C. The library was then pooled, precipitated, and purified by reverse-phase HPLC to give 14.8 μmol of product. Only a portion of this material was brought forward into Cycle 4.

Cycle 4. The Cycle 3 product (8.0 μmol) was dissolved in 8.0 mL water and arrayed into 384 wells (21 nmol/well). To each well was then added 8.5 μL of 10X ligation buffer, 0.84 μL T4 DNA ligase, and 13 μL water. DNA tags were then added (41 μL of 1 mM stocks in water) and the ligations were allowed to proceed at 16 degrees C for 16 h. After ligation, 5 M NaCl (10% by volume) and 2.5 volumes of cold ethanol were added to each vessel to precipitate the DNA. The DNA pellets recovered after centrifugation were each dissolved in 21 μL 250 mM sodium phosphate pH 5.5. To each well was added 50 equiv. of amine building block (200 mM in 1:1 MeCN/H_2O), 100 equiv. of HCl (200 mM) and 100 equiv. DMT-MM (200 mM in H_2O). A total of 383 different amines were used, leaving one well as a control null. After 48 h, another 50 equiv. DMT-MM was added, followed by an additional 50 equiv. at 60 hours. After 72 h total reaction time, wells were pooled and precipitated, and the crude library purified by reverse-phase HPLC. Final yield was 4.6 μmol, sufficient material for 920 selection experiments at 5 nmol per selection.

Positive-Control Synthesis

The positive control **1** was synthesized according to the literature, (J.-D. Charrier, F. Mazzei, D. Kay, A. Miller, WO 2004/00083) except for the last step:

Cyclopropane carboxylic acid {4-[4-chloro-6-(5-methyl-2H-pyrazol-3-ylamino)-pyrimidin-2-ylsulphanyl]-phenyl} amide (20mg) and t-butyl 2-(piperazin-1-yl)acetate (100mg) were mixed with DMF (1ml). To the mixture was added DIEA (0.4 ml). The reaction mixture was heated at 110^{o}C for 1hr and LCMS showed the reaction was completed. Evaporation of the solvent gave the crude product which was further purified with RP-HPLC to give the desired product.

Nature Chemical Biology: doi:10.1038/nchembio.211

The above product was treated with 50% TFA in CH_2Cl_2 (2 ml) in the presence of TIPS (14µl) at rt for 4hrs. The solvent was evaporated and the product was purified by RP-HPLC to give the pure product **1**. H^1 NMR (DMSO-d^6), δ 10.43 (1H, s), 9.47 (1H, s), 7.72 (2H, d, J=8.4 Hz), 7.48 (2H, d, J=8.4 Hz), 6.02 (1H, br s), 5.44 (1H, s), 4.14 (2H, s), 3.68 (4H, br s), 3.32 (4H, br s), 2.02 (3H, s), 1.82 (1H, p, J=6.4 Hz), 0.81 (4H, d, J=6.4 Hz). C^{13} NMR (DMSO-d^6), δ 171.95, 169.97, 167.28, 161.43, 160.02, 148.05, 140.28, 138.47, 136.55, 122.85, 119.13, 95.19, 80.92, 55.20, 51.03, 40.52, 14.62, 10.34, 7.26. ESIMS 509.33 (M+H, 100%).

To AOP-headpiece solution in pH 9.4 phosphate buffer (1mM, 100µl, 100 nmol) was added **1** (2 mg, 4 µmol, 40eq) in DMF (20ul). The solution was cooled to 4°C and DMT-MM (1.2 mg, 4 µmol, 40eq) in water (20ul) was added. The reaction was left at 4°C overnight. Ethanol precipitation gave the crude product which was further purified by RP-HPLC to give the desired product **1-DNA**. MS: (M-3)/3=1889.91 (cal. 1890.3), (M-4)/4=1417.45 (cal. 1417.5), (M-5)/5=1133.93 (cal. 1133.8), (M-6)/6= 944.87 (cal. 944.7).

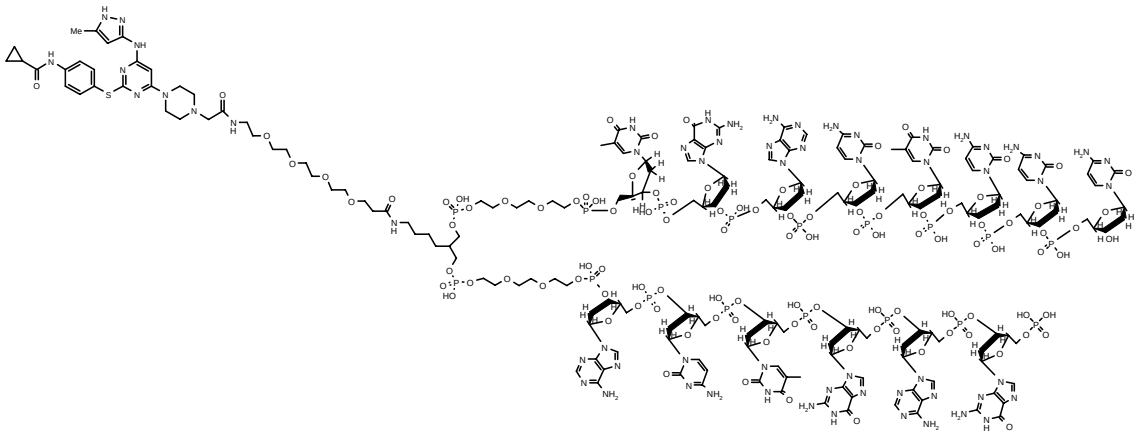


Figure 3. Structure of compound **1-DNA**.

Affinity Selection

Closing Primer Ligation. Prior to selection, 50 nmol aliquots of library were ligated with closing primer using T4 DNA ligase. The primer duplex was then filled by Klenow polymerization. The generic sequence of the closing primer was:

5′ XXXXXXXXXX 3′
3′ AAXXXXXXXXXXNNNNNGTCTGTTCGAAGTGGACG 5′

The variable region X was unique to each closing primer and served as an identifier for the particular library and selection experiment. The variable region N was a randomized sequence used to identify PCR duplication in the sequence data.

Nature Chemical Biology: doi:10.1038/nchembio.211

Aurora A Kinase selection experiments (Method A). DEL A (5 nmol) was incubated with His-tagged Aurora A protein (Upstate) at a concentration of 1μM in 60μL selection buffer [50mM Tris-HCl, pH7.5, 150mM NaCl, 0.1% tween-20, 1mg/mL Sheared Salmon Sperm DNA (Ambion), 1mg/mL Bovine Serum Albumin (Ambion) and 10mM βME] for one hour at room temperature. The solution was then incubated for 30 minutes with 20μL Dynabeads® TALON™ beads (Dynal Biotech). After this incubation, the beads were captured and washed 8 times in 200μL selection buffer. In order to elute the protein/bound library molecules off of the beads, the beads were resuspended in 30uL selection buffer and heated at 72ºC for 5 minutes. The elution was separated from the beads and added to 20μL fresh TALON™ beads and incubated for 15 minutes at room temperature to remove denatured protein. This was repeated a second time with fresh beads. The elution from this round was then incubated with 60uL 500nM His-tagged Aurora A protein (Upstate) in selection buffer for the second round of selection, followed by the Dynabeads® TALON™ bead (10μL of beads) incubation, wash and elution steps described. For the third round, the second round elution was then incubated with 500nM His-tagged Aurora A protein (Upstate) in selection buffer, and again, followed by the Dynabeads® TALON™ bead (10μL of beads) incubation, wash and elution steps described. The final elutions were used as templates for PCR amplification of the DNA codes of the selected molecules. For the spike-in experiment with the Aurora A positive control, Compound **1** was added to the initial incubation at a concentration of 11.7 pM.

Aurora kinase selection experiments (Method B). His-tagged Aurora kinase A (900pmoles, Upstate) was immobilized on 5μL IMAC resin (Phynexus). DEL library (5nmol) in 60 μL selection buffer (50mM Tris-HCl, pH 7.5, 150mM NaCl, 0.1% tween-20, 1mg/mL sheared salmon sperm DNA (Ambion) 1mg/mL BSA (Ambion), and 10mM βME) was incubated with the immobilized aurora kinase for 1 h at room temperature, then washed 10 times with 100μL selection buffer. To elute the protein/bound library molecules, the resin was incubated with 60μL imidazole elution buffer (50mM Tris-HCl, pH 7.5, 150mM NaCl, 0.1% tween-20, 1mg/mL sheared salmon sperm DNA (Ambion) 1mg/mL BSA (Ambion), 10mM βME, and 100mM imidazole) for 5 min. The elution was heated for 5 min at 72°C to denature the aurora kinase protein. To allow for recapture of the denatured protein on IMAC resin, the elution was then diluted 10-fold in selection buffer to lower the imidazole concentration to 10mM. The diluted elution was then incubated with IMAC resin (Phynexus) for 15 min to remove denatured protein and library molecules that bind to IMAC resin. Subsequent rounds of selection were performed by incubating the elution from the previous round with immobilized aurora kinase (Upstate), followed by the wash and elution steps described.

DEL A p38 Selection. DEL library (5nmol) was incubated with 500nM His-tagged p38α protein (R&D Systems) in 60μL selection buffer (50mM Tris-HCl, pH 7.5, 150mM NaCl, 0.1% tween-20, 1mg/mL sheared salmon sperm DNA (Ambion) 1mg/mL BSA (Ambion),

Nature Chemical Biology: doi:10.1038/nchembio.211

and 1mM βME) for 1 h at room temperature. The solution was incubated with 5μL IMAC resin (Phynexus) for 5 min, then washed 10 times with 100μL selection buffer. To elute the protein-bound library molecules, the resin was incubated with 60μL imidazole buffer (50mM Tris-HCl, pH 7.5, 150mM NaCl, 0.1% tween-20, 1mg/mL sheared salmon sperm DNA (Ambion) 1mg/mL BSA (Ambion), 1mM βME, and 100mM imidazole) for 5 min. The eluant was heated for 5 min at 72°C to denature the p38α protein. To allow for recapture of the denatured protein on IMAC resin, the elution was then diluted 10-fold in selection buffer to lower the imidazole concentration to 10mM. The diluted elution was then incubated with IMAC resin (Phynexus) for 15 min to remove denatured protein and library molecules that bind to IMAC resin. Subsequent rounds of selection were performed by incubating the eluant from the previous round with 200nM His-tagged p38α protein (Roche) in selection buffer for the second round of selection, and 200nM His-tagged p38α protein (Roche) in selection buffer for the third round of selection, followed by the wash and elution steps described. Eluted molecules were used as templates for PCR amplification.

DEL B p38 Selection. A 238pmol aliquot of p38α (R&D Systems) was immobilized on 5μL IMAC resin (Phynexus). DEL B library (5nmol) in 60 uL selection buffer (50mM Tris-HCl, pH 7.5, 150mM NaCl, 0.1% tween-20, 1mg/mL sheared salmon sperm DNA (Ambion) 1mg/mL BSA (Ambion), and 1mM βME) was incubated with the immobilized p38α for 1 h at room temperature, then washed 10 times with 100μL selection buffer. To elute the bound library molecules, the resin was incubated with 60μL elution buffer at 72°C for 10min. The elution was then incubated twice with IMAC resin (Phynexus) for 15 min to remove denatured protein and library molecules that bind to IMAC resin. Subsequent rounds of selection were performed by incubating the elution from the previous round with fresh immobilized p38α (R&D Systems), followed by the wash and elution steps described.

Decoding. After selection, library molecules were amplified by PCR (5 min at 95°C, then 20 cycles of 30s at 92°C; 15s at 55°C; 15s at 72°C, followed by 10 min at 72°C) using primers 5'O (5'-GCCTTGCCAGCCCGCTCAGTGACTCCCAAATCGATGTG-3'; 400nM; IDT) and 3'O (5'GCCTCCCTCGCGCCATCAGGCAGGTGAAGCTTGTCTG-3'; 400nM; IDT). The PCR products were purified to remove primers and nucleotides (PCR Clean-Up Kit, Qiagen) and sequenced (454 Life Sciences).

Compound Synthesis

General Procedure for synthesis of Compounds **2 - 12**:

1. ***First Displacement:*** 80 mM stock solutions of cyanuric chloride and amine #1 were freshly prepared in 1:1 acetonitrile:aqueous buffer (250 mM borate, pH 9.4). Equal

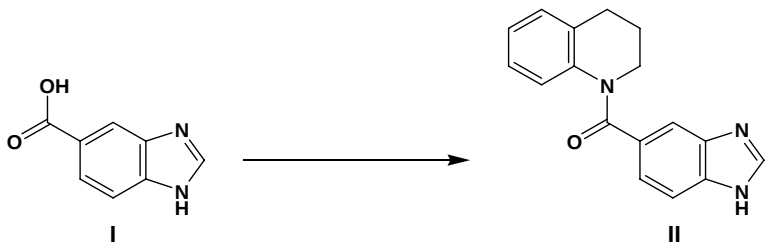
Nature Chemical Biology: doi:10.1038/nchembio.211

volumes of the two stock solutions were combined after chilling over ice. The mono-adduct immediately precipitated upon mixing. The resulting colloidal solution was allowed to warm to RT.

2. ***Second Displacement:*** 1.25 mL of the above mono-adduct solution (100 µmole), 1 equiv. of amine #2, 5 mL acetonitrile, and 50 mg K_2CO_3 (excess) were combined in a 10 mL vessel. The mixture was agitated for 1 hr.

3. ***Third Displacement:*** To the crude di-adduct reaction mixture was added 500 µmole amine #3 (5 equiv.). The reaction mixture was allowed to sit at RT overnight. The solution was then concentrated, reconstituted in 5 mL 10% acetonitrile (aq), and purified by reverse-phase HPLC to yield product (30-50%).

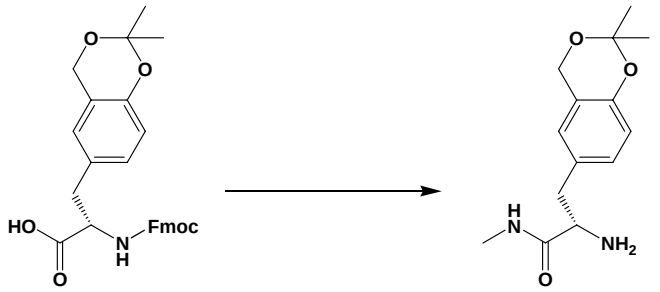
Procedures for Synthesis of Compounds **13** – **15**:



I II

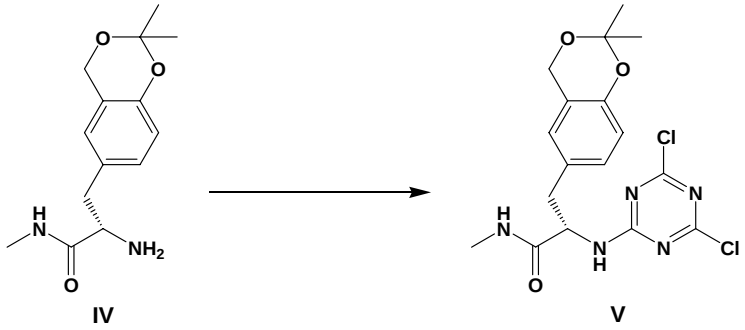
A 100-mL, one-necked, round-bottomed flask equipped with a rubber septum and nitrogen inlet needle was charged with 1*H*-benzimidazole-5-carboxylic acid (**I**) (0.858 g, 5.29 mmol, 1 equiv) and 30 mL of anhydrous benzene. The resulting suspension was cooled to 0 °C, and $SOCl_2$ (5.0 mL, 8.2 g, 69 mmol, 13 equiv) was added via syringe over ca. 1 min. The reaction mixture was stirred 5 min at 0 °C and then allowed to warm to RT. The flask was then equipped with a reflux condenser and the reaction mixture was heated at reflux for 6 h. The reaction mixture was then concentrated to afford a brown solid which was suspended in 15 mL of benzene and concentrated again; this was repeated with two additional 15-mL portions of benzene to ensure complete removal of excess $SOCl_2$. The resulting brown solid was suspended in 40 mL of THF and a solution of tetrahydroquinoline (2.00 mL, 2.11 g, 15.9 mmol, 3 equiv) in 20 mL of THF was added. The resulting reaction mixture was stirred at RT for 16 h, after which point it was concentrated, and the residue was dissolved in CH_2Cl_2 and concentrated onto 8 g of silica gel. Purification on a 120 g silica gel column eluting with 10-30% MeOH-CH_2Cl_2 afforded 0.466 g (32%) of amide **II** as a brown solid (ca. 95% pure by LC/MS): LRMS (ESI) Calcd for $C_{17}H_{15}N_3O$ $(M+H)^+$: 278.12. Found: 278.26.

Nature Chemical Biology: doi:10.1038/nchembio.211



III IV

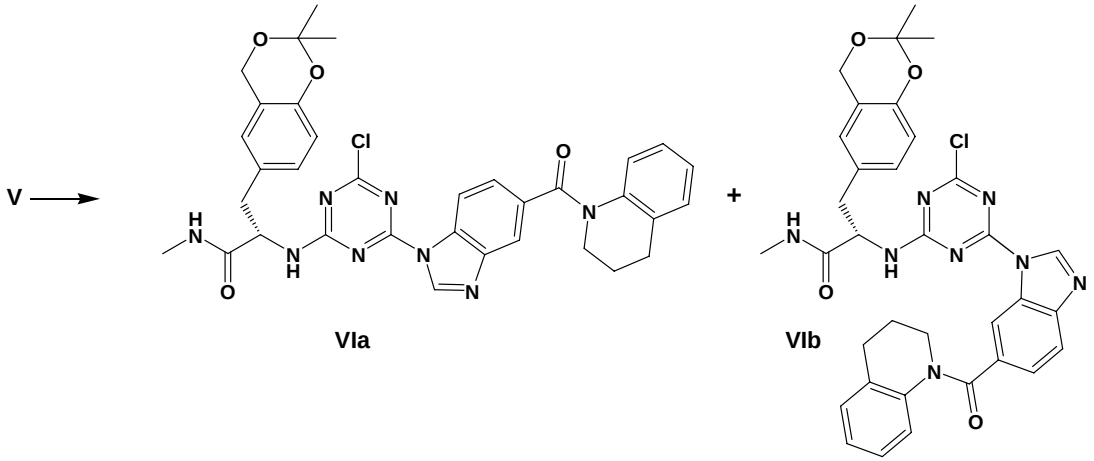
A 100-mL, one-necked, round-bottomed flask equipped with a rubber septum and nitrogen inlet needle was charged with Fmoc-protected amino acid **III** (Bachem, 1.014 g, 2.141 mmol, 1.0 equiv), HATU (1.221 g, 3.212 mmol, 1.5 equiv), and 8 mL of anhydrous THF. *i*-Pr_2EtN (1.1 mL, 0.83 g, 6.4 mmol, 3.0 equiv) was added dropwise via syringe over ca. 1 min. The resulting bright yellow solution was stirred 5 minutes at RT, and then $MeNH_2$·HCl (0.188 g, 2.78 mmol, 1.3 equiv) was added. After stirring 3 h at RT, the reaction mixture was diluted with 30 mL of EtOAc and washed with 20 mL of half-saturated $NaHCO_3$. The aqueous phase was back-extracted with two 30-mL of portions of EtOAc, and the combined organic phases were dried over Na_2SO_4, filtered, and concentrated to afford 3.47 g of an orange liquid. This crude product was dissolved in 50 mL of 30% Et_2NH in CH_3CN and stirred at RT for 1 h. The reaction mixture was then concentrated and purified on a 40 g silica gel column eluting with 0-20% MeOH-(1% Et_3N in CH_2Cl_2) to afford 0.613 g (108%) of amine **IV** (ca. 90% pure by ^1H NMR and LC/MS analysis): LRMS (ESI) Calcd for $C_{14}H_{20}N_2O_3$ $(M+H)^+$: 265.15. Found: 265.09.



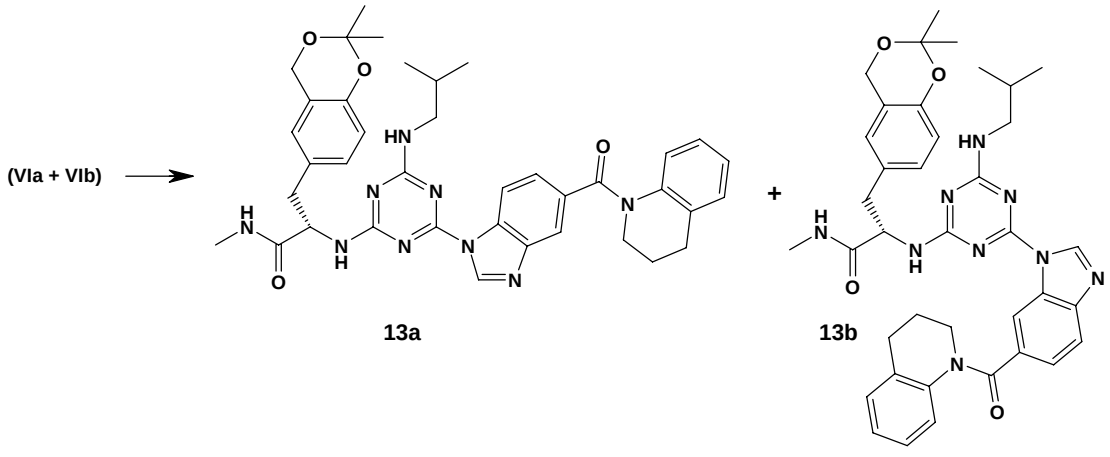
IV V

A vial (ca. 20 mL capacity) was charged with a solution of amine **IV** (0.196 g [92% pure], 0.681 mmol, 1.0 equiv) in 3.4 mL of anhydrous CH_2Cl_2 and cooled to 0 °C. A solution of cyanuric chloride (0.126 g, 0.681 mmol, 1.0 equiv) in 3.4 mL of anhydrous CH_2Cl_2 (pre-cooled to 0 °C) was added to the reaction mixture dropwise via pipette. The resulting reaction mixture was stirred 2 h at 0 °C, 2 h at RT, and then cooled back to 0 °C. *i*-Pr_2EtN (0.36 mL, 0.26 g, 2.0 mmol, 3.0 equiv) was added, the reaction mixture was stirred 10 min at 0 °C, and was then concentrated to afford an orange foam. Purification on a 12 g silica gel column eluting with 10-20% EtOAc-CH_2Cl_2 afforded 0.191 g (68%) of dichlorotriazine **V** as a white solid: LRMS (ESI) Calcd for $C_{17}H_{19}Cl_2N_5O_3$ $(M+H)^+$: 412.09. Found: 412.07.

Nature Chemical Biology: doi:10.1038/nchembio.211

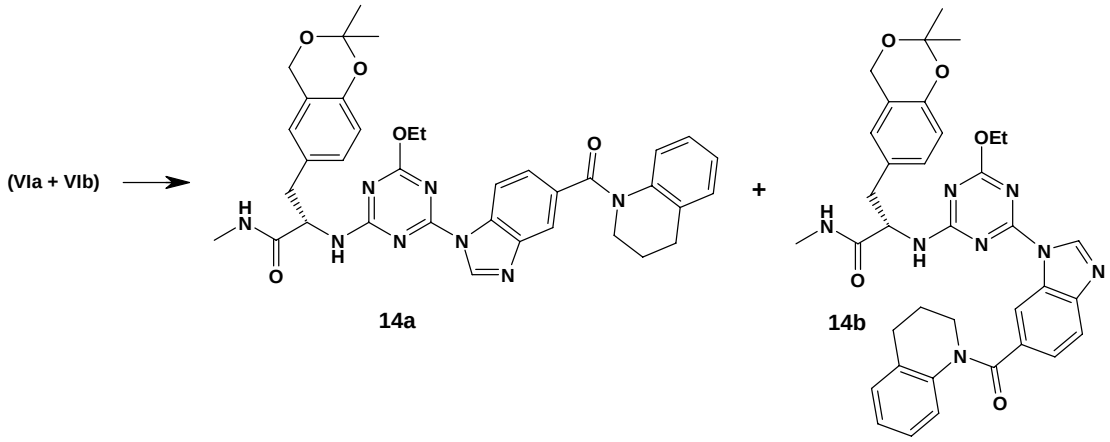


A vial (ca. 20 mL capacity) was charged with NaH (60% dispersion in mineral oil, 0.031 g, 1.30 mmol, 3.0 equiv). The NaH was rinsed with two 1-mL portions of hexanes (the supernatant was carefully removed, leaving solid NaH behind in the reaction vessel), and then a pre-cooled (4 °C) solution of benzimidazole **II** (0.120 g, 0.432 mmol, 1.0 equiv) in 2.2 mL of anhydrous NMP was added. The resulting reaction mixture was stirred at 0 °C for 15 min, and then a solution of **V** (0.178 g, 0.432 mmol, 1.0 equiv) in 2.2 mL of NMP was added. The reaction mixture was stirred 30 min at 0 °C, warmed to RT and stirred an additional 1.5 h. The crude reaction solution was then divided evenly between 3 clean vials and used in the next step without further treatment. LC/MS analysis of the crude reaction mixture confirmed formation of the desired triazine product **VIa/VIb (**mixture of 2 regioisomers in unknown proportion**)**: LRMS (ESI) Calcd for $C_{34}H_{33}ClN_8O_4$ $(M+H)^+$: 653.23. Found: 653.17.



Nature Chemical Biology: doi:10.1038/nchembio.211

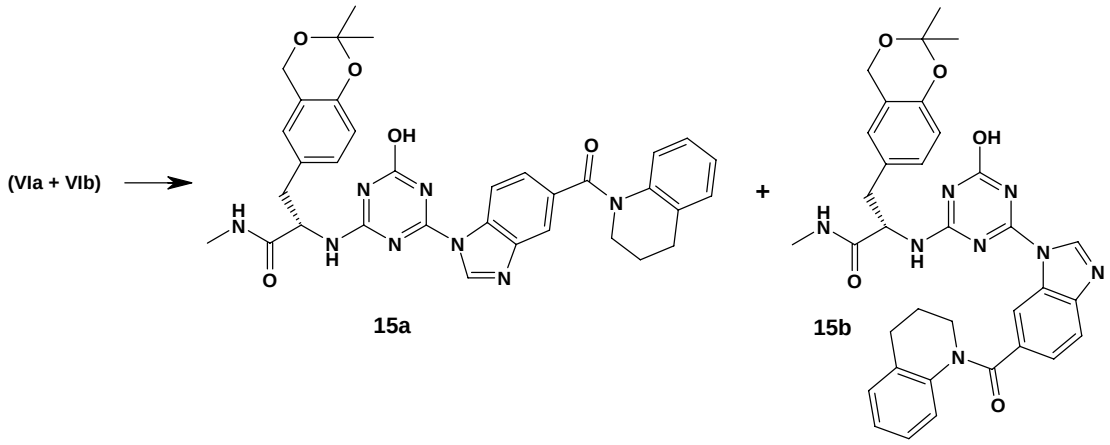
One third of the crude reaction mixture containing triazine **VIa/VIb** (ca. 0.144 mmol in 1.47 mL of NMP) was transferred to a clean vial and isobutylamine (0.215 mL, 0.0.158 g, 2.16 mmol, 15 equiv) was then added. After stirring 22 h at RT the reaction mixture was diluted with 10 mL of half-saturated NaCl, extracted with three 15-mL portions of EtOAc, and the combined organic phases were dried over Na_2SO_4, filtered, and concentrated. Subsequent purification on a 12 g silica gel column eluting with 0-10% MeOH-(0.5% Et_3N in CH_2Cl_2) afforded the desired product in < 90% purity. This material was then dissolved in 8 mL of 5:3 H_2O:CH_3CN and repurified on a reverse phase HPLC column eluting with 25-90% (9:1 CH_3CN:10 mM NH_4OAc in H_2O)-10 mM NH_4OAc (aq, pH 7.6) to deliver a total of 0.013g (13% yield over 2 steps from **V**) of **13a/13b** (mixture of 2 regioisomers in unknown proportion) as a white solid: LRMS (ESI) Calcd for $C_{38}H_{43}N_9O_4$ $(M+H)^+$: 690.34. Found: 690.26.



One third of the crude reaction mixture containing triazine **VIa/VIb** (ca. 0.144 mmol in 1.47 mL of NMP) was transferred to a clean vial and EtOH (0.126 mL, 0.099 g, 2.16 mmol, 15 equiv) was then added. The resulting reaction mixture was stirred at RT for 1.5 h, and then a second portion of EtOH (0.126 mL, 0.099 g, 2.16 mmol, 15 equiv) was added. After stirring an additional 19 h at RT, the reaction mixture was diluted with 10 mL of half-saturated NaCl, extracted with three 15-mL portions of EtOAc, and the combined organic phases were dried over Na_2SO_4, filtered, and concentrated (Note: In some cases it was necessary to add up to 6 equiv. of NaOEt to reach completion). Subsequent purification on a 12 g silica gel column eluting with 0-20% MeOH-(0.5% Et_3N in CH_2Cl_2) afforded the desired product in < 90% purity. This material was then dissolved in 8 mL of 5:3 H_2O:CH_3CN, diluted with 1.5 mL of MeOH, and repurified on a reverse phase HPLC column eluting with 25-90% (9:1 CH_3CN:10 mM NH_4OAc in H_2O)-10 mM NH_4OAc (aq, pH 7.6) to deliver a total of 0.009 g (9% yield over 2 steps from **V**) of **14a/14b** (mixture of 2 regioisomers in unknown proportion) as a white solid: LRMS (ESI) Calcd for $C_{36}H_{38}N_8O_5$ $(M+H)^+$: 663.30. Found: 663.27.

Nature Chemical Biology: doi:10.1038/nchembio.211

Pure samples of each isomer (**14a** and **14b**) for p38 MAP kinase assays and X-ray crystallography were obtained by subsequent purification on a reverse phase HPLC eluting with 62 to 72% B:A (A = 50 mM NH_4OAc; concentrated NH_4OH added to obtain pH = 9; B = 50 mM NH_4OAc in 75:25 CH_3CN:H_2O, pH 9).



One third of the crude reaction mixture containing triazine **VIa/VIb** (ca. 0.144 mmol in 1.47 mL of NMP) was transferred to a clean vial and H_2O (0.039 mL, 0.039 g, 2.16 mmol, 15 equiv) was then added. The resulting reaction mixture was stirred at RT for 1.5 h, and then a second portion of H_2O (0.039 mL, 0.039 g, 2.16 mmol, 15 equiv) was added. After stirring an additional 21 h at RT, the reaction mixture was diluted with 10 mL of half-saturated NaCl, extracted with three 15-mL portions of EtOAc, and the combined organic phases were dried over Na_2SO_4, filtered, and concentrated (Note: In some cases it was necessary to add up to 5 equiv. of NaOH to reach completion). Subsequent purification on a 12 g silica gel column eluting with 10-20% MeOH-(0.5% Et_3N in CH_2Cl_2) afforded the desired product in < 90% purity. This material was then dissolved in 8 mL of 5:3 H_2O:CH_3CN and repurified on a reverse phase HPLC eluting with 20-85% (9:1 CH_3CN:10 mM NH_4OAc in H_2O)-10 mM NH_4OAc (aq, pH 7.6) to deliver a total of 0.011g (12% yield over 2 steps from **V**) of **15a**/**15b** (mixture of 2 regioisomers) as a white solid: LRMS (ESI) Calcd for $C_{34}H_{34}N_8O_5$ $(M+H)^+$: 635.27. Found: 635.27. Partial separation of the 2 regioisomers was achieved during preparative HPLC purification, allowing for a pure sample of **15b**, a sample enriched in **15a** (≥ 80:20 **15a:15b**), as well as an unresolved mixture of regioisomers (ca. 56:44 **15a:15b**) to be assayed.

Pure samples of each isomer (**15a** and **15b**) for p38 MAP kinase assays and X-ray crystallography were obtained by subsequent purification on a reverse phase HPLC eluting with 46 to 55% B:A (A = 50 mM NH_4OAc; concentrated NH_4OH added to obtain pH = 9; B = 50 mM NH_4OAc in 75:25 CH_3CN:H_2O, pH 9).

Nature Chemical Biology: doi:10.1038/nchembio.211

Analysis. All synthesized compounds were characterized by LCMS and showed >95% purity and expected mass data. A few representative compounds were characterized more fully. Analysis of compounds was complicated by rotameric equilibria. In some cases, elevated temperature was required to obtain suitable NMR spectra.

Compound 4 ^{1}H NMR (400 MHz, DMSO-d_6, 75° C): δ 11.45 (bs, 1H), 9.57 (bs, 1H), 8.47 (bs, 1H), 7.85 – 7.55 (m, 3H), 7.36 – 7.34 (m, 1H), 7.2 – 7.0 (m, 4H), 4.41 (bs, 2H), 3.90 (bs, 1H), 3.70 (s, 3H), 2.26 (s, 3H), 2.25 (s, 3H), 1.06 (d, 3H, J = 7 Hz), 0.87 (s, 9H). ^{13}C NMR (100 MHz, DMSO-d_6, 75° C) δ 163.0 (br), 156.3 (b), 137.8 (br), 137.0, 130.1 (br), 127.9, 127.8, 127.3, 125.4 (br), 124.4, 62.7, 54.0, 43.4 (br), 34.0, 25.7, 20.5, 17.3, 14.9. LRMS(ESI) 478.3 (M+H, 100%), 954.5 (2M+H, 10%).

Compound 12.^{1}H NMR (400 MHz, DMSO-d_6, 75 °C): δ 9.06 (br s, 1 H), 8.71 (dd, *J* = 4.4, 1.6 Hz, 1 H), 8.57 (d, *J* = 2.4 Hz, 1 H), 8.07 (app dd, *J* = 9.0, 2.2 Hz, 2 H), 7.89 (d, *J* = 8.8 Hz, 1 H), 7.38 (dd, *J* = 8.4, 4.4 Hz, 1 H), 7.15-7.23 (m, 2 H), 6.95 (d, *J* = 8.0 Hz, 1 H), 6.83 -6.90 (m, 1 H), 6.67 (br s, 1 H), 6.58 (br s, 1 H), 3.78 (s, 3 H), 3.53-3.64 (m, 2 H), 2.84-2.96 (m, 5 H). ^{13}C NMR (100 MHz, DMSO-d_6, 75 °C): δ 166.2, 165.6, 164.0, 157.1, 147.5, 143.8, 138.6, 134.4, 129.6, 128.4, 128.2, 127.6, 127.0, 124.0, 120.9, 120.0, 114.0, 110.7, 55.1, 40.2, 29.6, 26.9. LRMS (ESI) Calcd for $C_{22}H_{23}N_7O$ $(M+H)^+$: 402.20. Found: 402.27.

Compound 14a (mixture of rotamers, confirmed by VT ^{1}H NMR experiments, no convergence at high T): ^{1}H NMR (400 MHz, DMSO-d_6, 25 °C): δ 9.08 (s, 1 H), 8.98 (s, 1 H), 8.64 (d, *J* = 8.0 Hz, 1 H), 8.57 (d, *J* = 8.8 Hz, 1 H), 8.43 (d, *J* = 7.6 Hz, 1 H), 8.30 (d, *J* = 8.4 Hz, 1 H), 8.13-8.19 (m, 1 H), 8.03-8.09 (m, 1 H), 7.74 (s, 1 H), 7.69 (s, 1 H), 7.42 (d, *J* = 8.8 Hz, 1 H), 7.33 (d, *J* = 8.8 Hz, 1 H), 7.22 (d, *J* = 7.2 Hz, 2 H), 7.10 (d, *J* = 8.4 Hz, 2 H), 6.93-7.04 (m, 4 H), 6.79-6.88 (m, 2 H), 6.68-6.76 (m, 2 H), 6.67 (d, *J* = 8.8 Hz, 1 H), 6.61 (d, *J* = 8.0 Hz, 1 H), 4.55-4.78 (m, 6 H), 4.32-4.48 (m, 4 H), 3.73-3.90 (m, 4 H), 2.96-3.06 (m, 2 H), 2.80-2.90 (m, 6 H), 2.60 (d, *J* = 4.8 Hz, 3 H), 2.56 (d, *J* = 4.4 Hz, 3 H), 1.93-2.07 (m, 4 H), 1.20-1.45 (m, 18 H).^{13}C NMR (100 MHz, DMSO-d_6, 25 °C): δ 171.24, 171.21, 170.3, 170.1, 169.2, 166.84, 166.77, 162.1, 162.0, 149.3, 149.2, 143.9, 143.4, 142.9, 139.23, 139.19, 132.26, 132.24, 132.09, 131.6, 131.5, 129.7, 129.6, 128.9, 128.8, 128.5, 128.4, 125.6, 125.5, 125.4, 125.0, 124.8, 124.24, 124.20, 120.2, 119.0, 118.9, 116.1, 116.0, 115.9, 115.4, 99.0, 98.9, 63.3, 63.2, 60.0, 59.9, 56.8, 56.4, 44.3, 36.9, 26.29, 26.25, 25.6, 25.5, 24.7, 24.5, 24.4, 24.0, 23.71, 23.69, 14.2, 14.1. LRMS (ESI) Calcd for $C_{36}H_{38}N_8O_5$ $(M+H)^+$: 663.30. Found: 663.27.

Biochemical Assay

Aurora A kinase assay. Compounds were assayed for inhibition of Aurora A kinase activity using a 96 well plate radiometric kinase assay. Aurora A kinase (Upstate) was pre-diluted in 50mM Tris, pH 7.5, 0.1mM EGTA, 0.1 mM sodium orthovanadate, 0.1% beta-mercaptoethanol, 0.03% Brij-35, and 1 mg/ml BSA and then added at a final concentration of 5 nM to assay buffer (20mM HEPES 7.4, 10mM $MgCl_2$, 25mM beta-glycerophosphate, 1mM DTT) containing 200 uM Kemptide (LRRASLG). Compounds

Nature Chemical Biology: doi:10.1038/nchembio.211

pre-diluted in assay buffer with 10% DMSO for a final concentration of 1% DMSO in the assay were then added to the wells. Compounds were pre-incubated with kinase for 20 min at room temperature before the reaction was initiated by the addition of 10 uM ATP/0.02 uCi/ul [gamma-33P]ATP. The reaction plates were incubated at 30C for 2h. The reactions were stopped by the addition of 200 mM phosphoric acid and transferred to a 96 well Millipore phosphocelluose filter plate. The filter plates were washed repeatedly with 100 mM phosphoric acid to remove excess [gamma-33P] ATP and the filters were then dried and 20 ul of scintillation fluid (Microscint 40) was added to each well. The filter plates were counted on a TopCount-NXT scintillation counter and the data was processed using Prism (GraphPad) curve fitting software.

p38α Kinase Assay (Compounds 2-4). Compounds were assayed for inhibition of p38α kinase activity using a 96 well plate radiometric kinase assay. Assay buffer (20mM HEPES 7.4, 10mM $MgCl_2$, 25mM beta-GP, 1mM DTT, 0.1 mg/ml MBP) containing 2 nM active p38α kinase (R&D Systems) was added to the wells followed by the addition of compounds pre-diluted in assay buffer with 10% DMSO for a final concentration of 1% DMSO in the assay. Compounds were pre-incubated with kinase for 20 min at room temperature before the reaction was initiated by the addition of 10 uM ATP/0.02 uCi/ul [gamma-33P]ATP. The reaction plates were incubated at 30C for 4h. The reactions were stopped by the addition of 200 mM phosphoric acid and transferred to a 96 well Millipore phosphocelluose filter plate. The filter plates were washed and counted in the same way as in the Aurora assay protocol and the data was processed using Prism curve fitting software. Under these conditions, accurate IC_{50} values for tight-binding ligands could not be determined due to ligand depletion.

p38α Binding Assay (Compounds 13- 15). Binding affinities of compounds to the p38 active site were measured using a competitive displacement assay (DiscoveRx HitHunter p38 MAPK Binding Assay). Displacement of a modified p38 inhibitor by the test compound activates beta-galactosidase, resulting in the production of luminescence. Compounds were prediluted in assay buffer supplied with the kit and transferred into 96 well plates. p38α (Upstate) was added to final concentration of 6.3 nM and the plates were incubated at room temperature for 30 minutes. Detection reagents were added according to the assay instructions. The luminescent signal was read using a Safire 2 (Tecan) and the data was processed using Prism curve fitting software. Under these conditions, accurate EC_{50} values for tight-binding ligands could not be determined due to ligand depletion.

Crystallization and structure determination

Crystals of AurA in complex with compound **10** were grown by sitting-drop vapor diffusion from a solution of 16%-20% polyethylene glycol 3350 buffered with 100 mM Na Citrate pH 5.6-6.4. The drops were formed by mixing 2 µl of the above solution with

Nature Chemical Biology: doi:10.1038/nchembio.211

2 µl of 25 mg/ml protein complex. The crystals appeared within 3 days at room temperature. Crystals were allowed to equilibrate in a similar crystallization solution, in the presence of 20% ethylene glycol, prior to freezing in liquid nitrogen.

Crystals of p38 in complex with compound 14b were obtained by using of 18% polyethylene glycol (PEG) 8000/0.2 M $Mg(OAc)_2$/0.1 M Hepes, pH 7.0. The protein concentration was ≈12 mg/ml in a buffer of 50 mM NaCl/1 mM EDTA/10 mM DTT/25 mM Hepes, pH 7.4. Crystals were cryoprotected at a final concentration of 35% PEG 400 added to mother liquor. Crystals were then flash-frozen in liquid nitrogen in a nylon loop.

X-ray diffraction data were collected using the MARCCD165 detector at the Industrial Macromolecular Crystallography Association beam line 17-ID at the Advanced Photon Source in the Argonne National Laboratory (Chicago, IL). Diffraction data were indexed and scaled using the HKL2000 software. The structures were determined by molecular replacement and the structures were refined using program REFMAC. The coordinates have been deposited in the Protein Data Bank (accession code 3HA6 and 3HA8).

Table 1 Data collection and refinement statistics

	Aurora A with **10**	p38 with **14b**
Data collection		
Space group	$P3_121$	$P2_12_12_1$
Cell dimensions		
a, *b*, *c* (Å)	86.4, 86.4, 143.3	46.2, 87.9, 121.5
α, β, γ (°)	90.0, 90.0, 120.0	90.0, 90.0, 90.0
Resolution (Å)	2.36 (2.44-2.36)	2.48 (2.57-2.48)
R_{sym} or R_{merge}	0.079 (0.272)	0.064 (0.580)
$I / \sigma I$	27.0 (2.0)	26.9 (2.0)
Completeness (%)	84.0 (45.1)	97.2 (86.5)
Redundancy	9.2 (5.2)	6.3 (4.0)
Refinement		
Resolution (Å)	25-2.36	25-2.48
No. reflections	19335	15683
R_{work} / R_{free}	0.223/0.252	0.214/0.247
No. atoms		
Protein	2458	2732
Ligand/ion	38	49
Water	29	49
B-factors		
Protein	62.4	49.1
Ligand/ion	63.8	46.4
Water	58.6	49.5
R.m.s. deviations		
Bond lengths (Å)	0.012	0.007
Bond angles (°)	1.173	1.069

Nature Chemical Biology: doi:10.1038/nchembio.211

Characterization of DEL-A

Starting material (primer-ligated AOP headpiece)

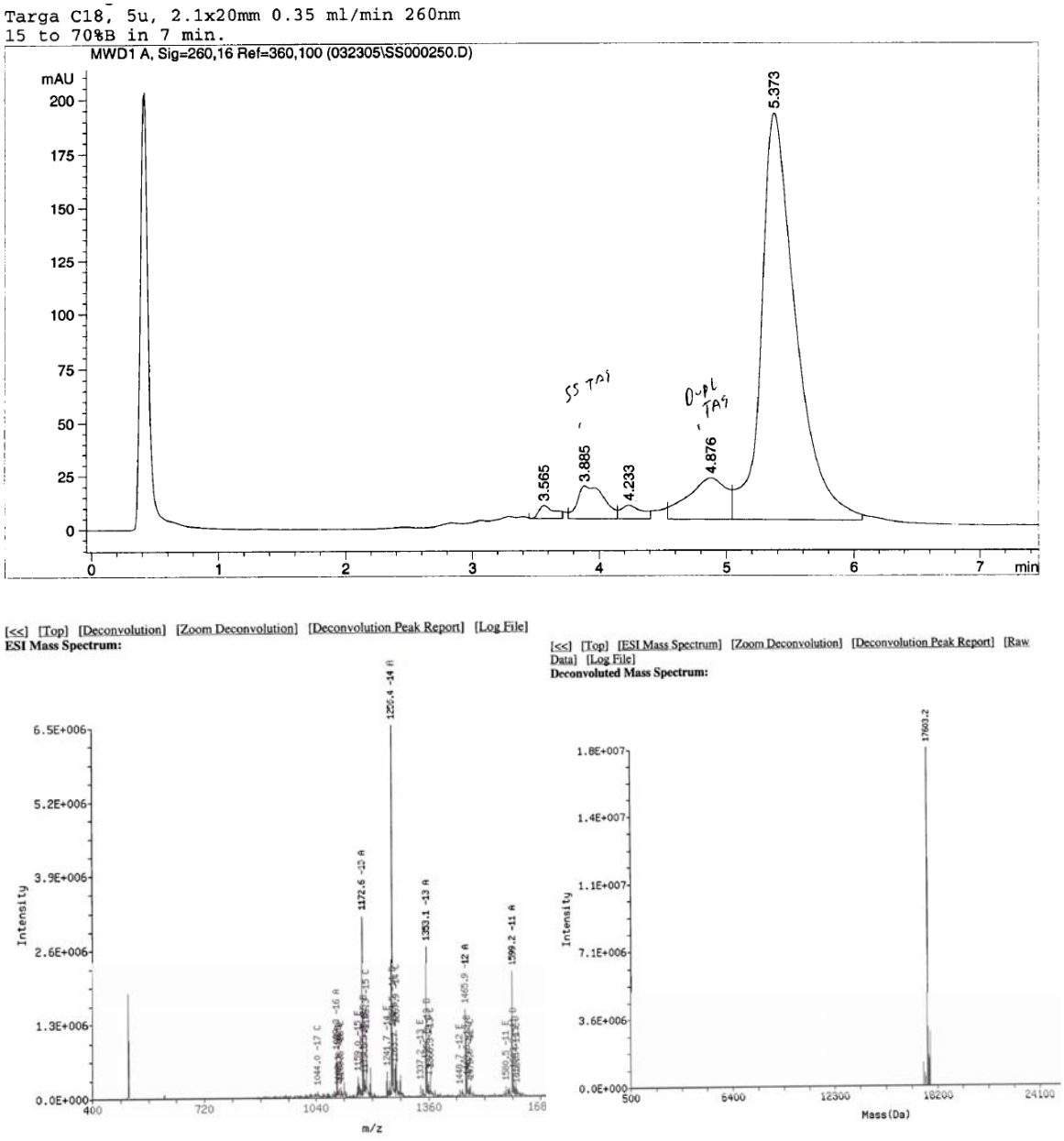
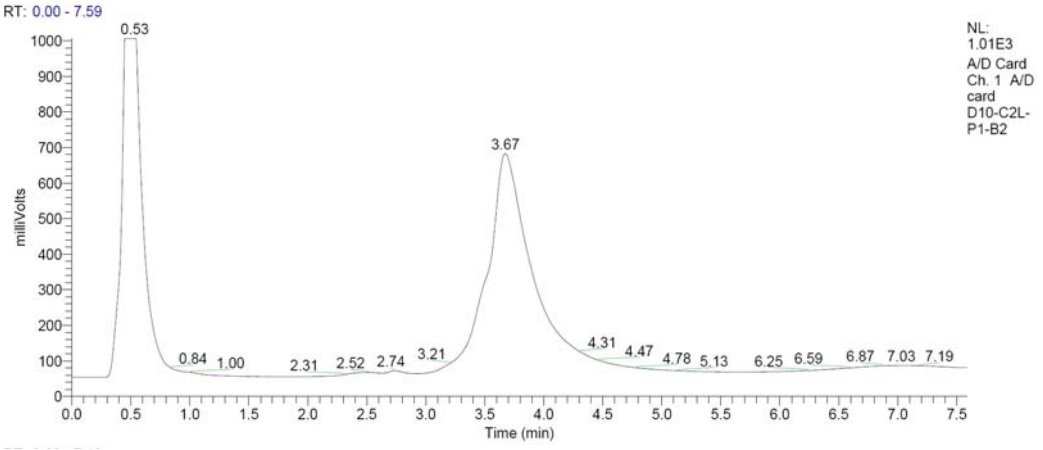


Figure 4. LC chromatogram, raw ESI mass spectrum, and deconvoluted mass spectrum of primer-ligated AOP headpiece. Expected mass = 17606; Observed mass =17603.

Nature Chemical Biology: doi:10.1038/nchembio.211

Cycle 1 Analysis

Plate 1 Well B2 Ligation (1 component)



[<<] [Top] [ESI Mass Spectrum] [Zoom Deconvolution] [Deconvolution Peak Report] [View Data] [Log File]

Deconvoluted Mass Spectrum:

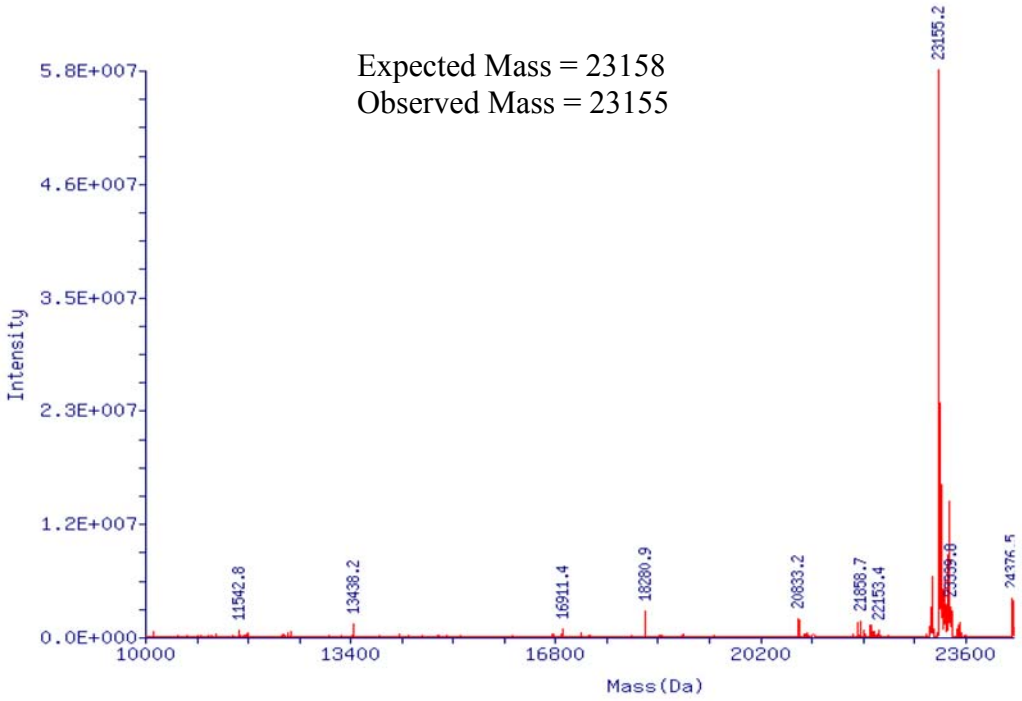
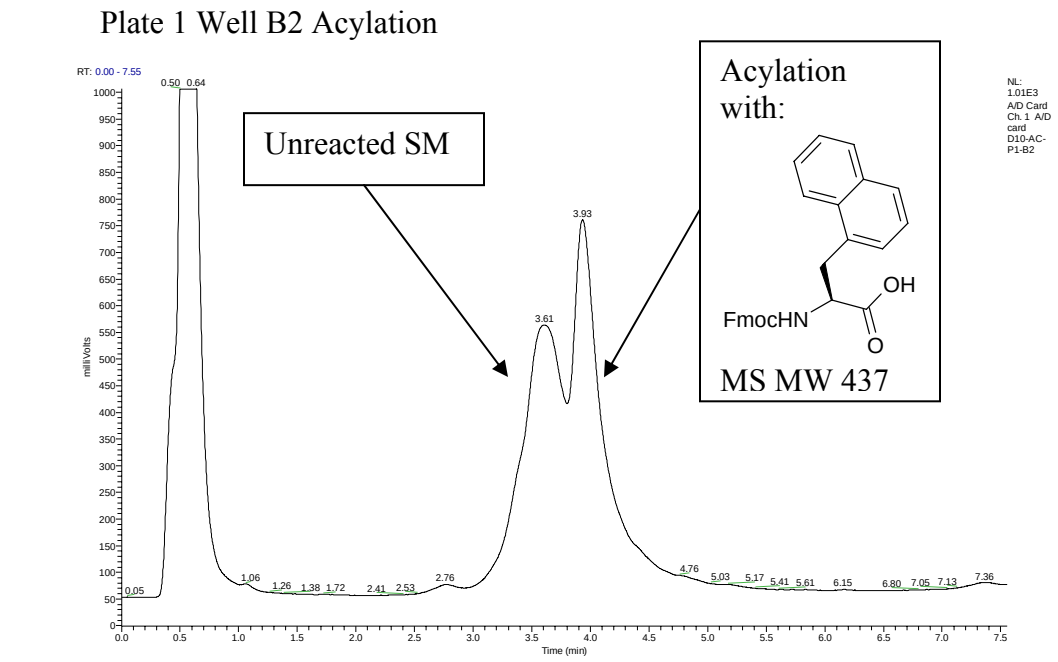


Figure 5. LC chromatogram and deconvoluted mass spectrum of ligation of plate 1 well B2.

Nature Chemical Biology: doi:10.1038/nchembio.211



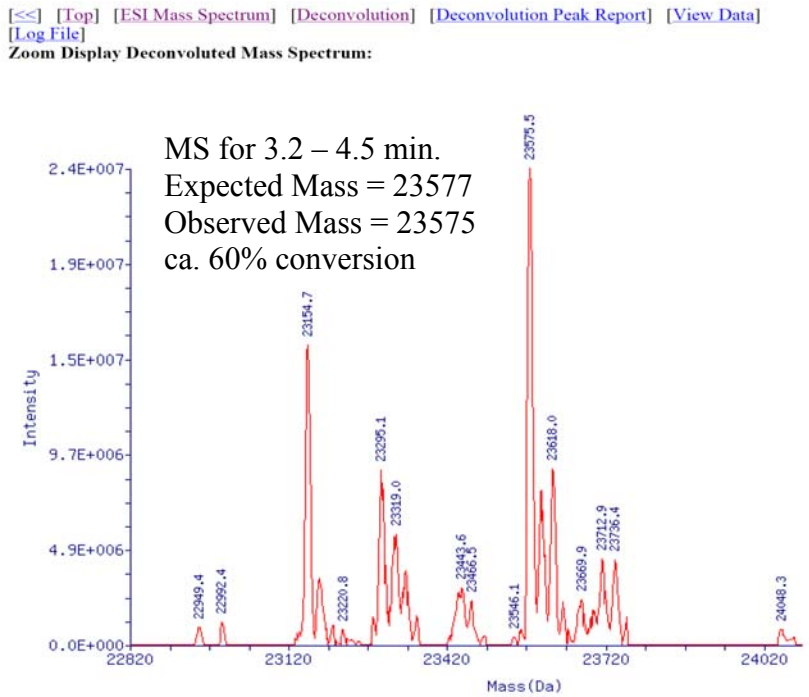
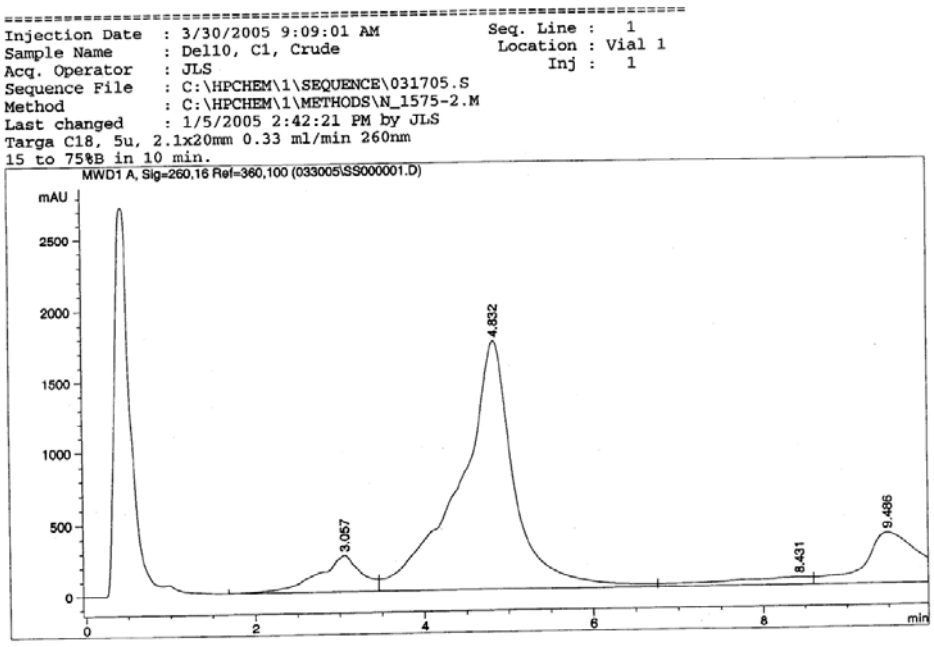


Figure 6. LCMS analysis of acylation step for well B2 from plate 1. Acylation with Fmoc-naphthylalanine proceeded around 60%. Additional adducts in the acylation mass spectrum are a dimethoxytriazinyl adduct from DMT-MM (+139).

Nature Chemical Biology: doi:10.1038/nchembio.211

Crude pooled Cycle 1 product (192 components)



[<<] [Top] [ESI Mass Spectrum] [Deconvolution] [Deconvolution Peak Report] [Raw Data] [Log File]

Zoom Display Deconvoluted Mass Spectrum:

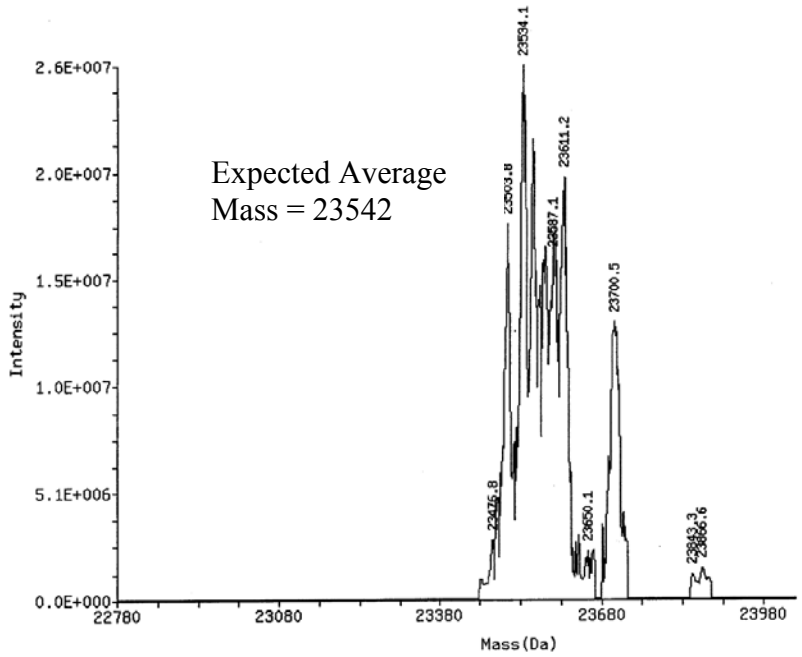
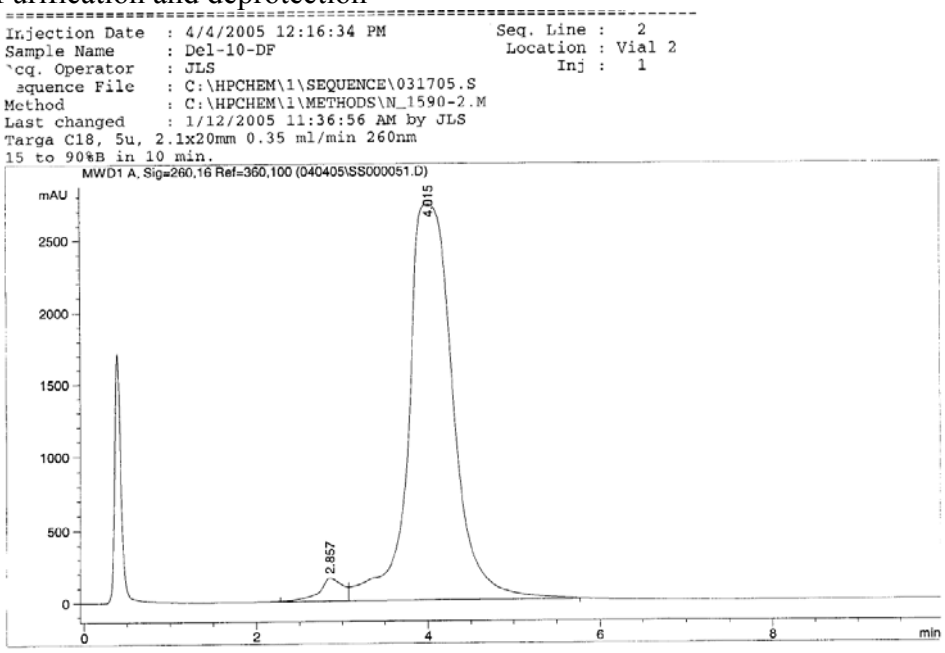


Figure 7. LC chromatogram and deconvoluted mass spectrum of crude Cycle 1 pooled product.

Nature Chemical Biology: doi:10.1038/nchembio.211

Purification and deprotection



[<<] [Top] [ESI Mass Spectrum] [Deconvolution] [Deconvolution Peak Report] [Raw Data] [Log File]

Zoom Display Deconvoluted Mass Spectrum:

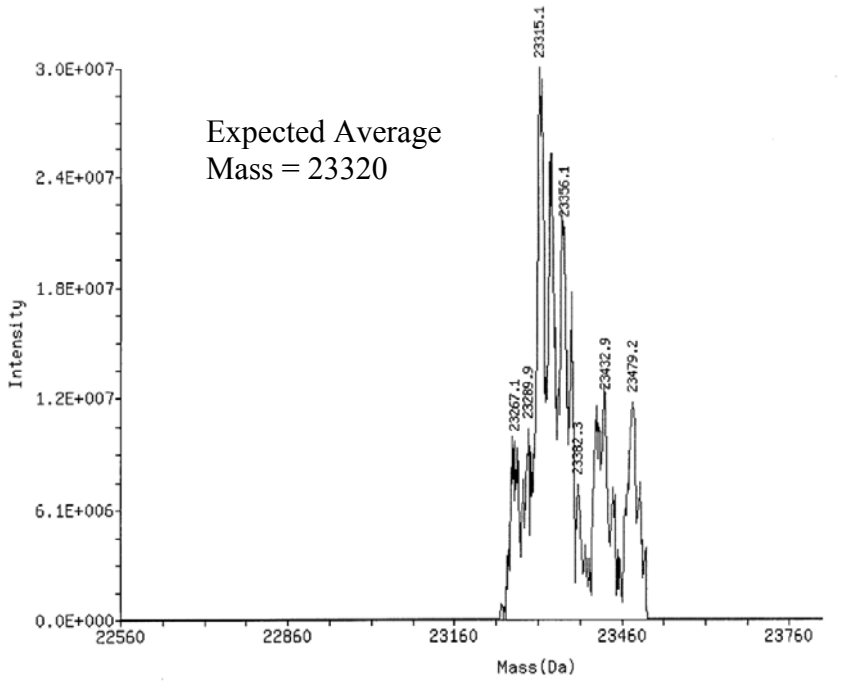


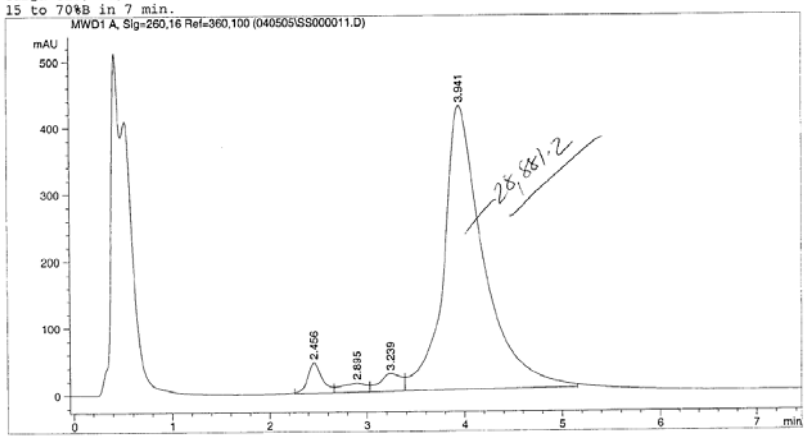
Figure 8. LC chromatogram and deconvoluted mass spectrum of Cycle 1 pooled product after reverse phase LC purification and Fmoc deprotection.

Nature Chemical Biology: doi:10.1038/nchembio.211

Cycle 2 Analysis

Plate 3 Well B2 Ligation (192 components)

```
Injection Date  : 4/5/2005 10:34:04 AM          Seq. Line :  11
Sample Name     : DEL-10-P4-C3-B2               Location : Vial 58
Acq. Operator   : JLS                              Inj :   1
Sequence File   : C:\HPCHEM\1\SEQUENCE\031705.S
Method          : C:\HPCHEM\1\METHODS\N1570-7M.M
Last changed    : 4/4/2005 3:21:54 PM by JLS
Targa C18, 5u, 2.1x20mm 0.35 ml/min 260nm
15 to 70%B in 7 min.
```



[<<] [Top] [ESI Mass Spectrum] [Deconvolution] [Deconvolution Peak Report] [Raw Data] [Log File]

Zoom Display Deconvoluted Mass Spectrum:

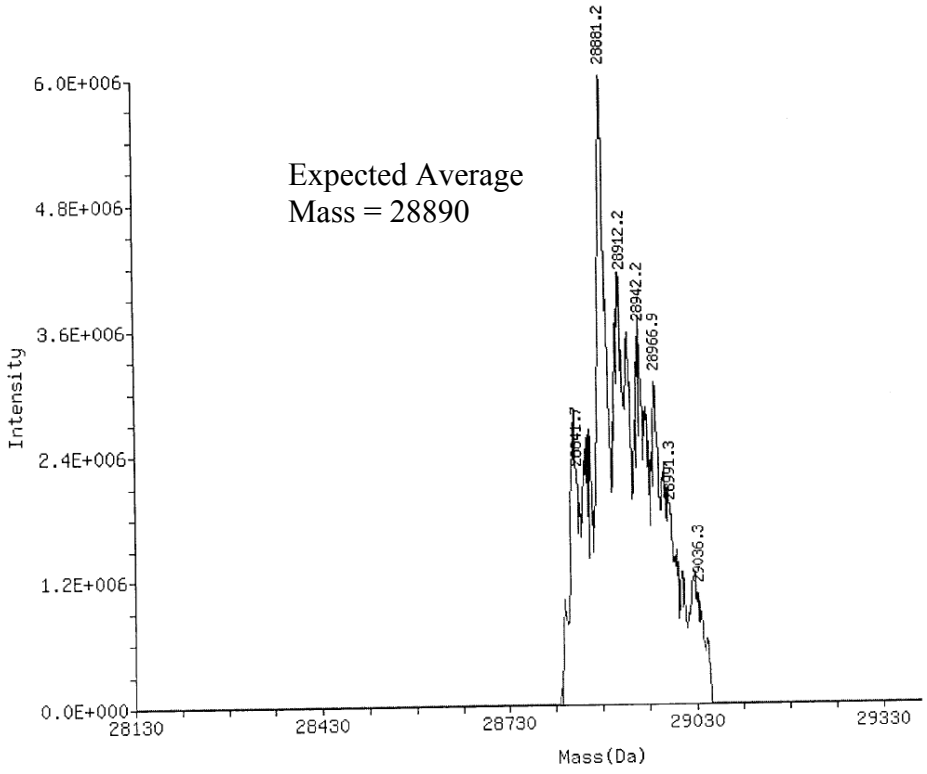
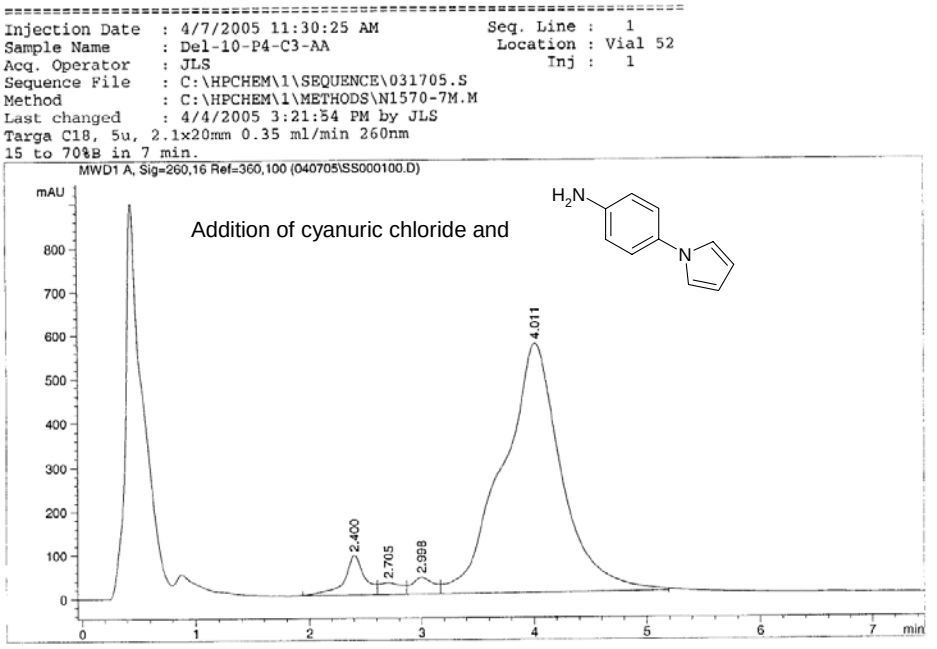


Figure 9. LC chromatogram and deconvoluted mass spectrum of Cycle 2 ligation for well B2 of plate3.

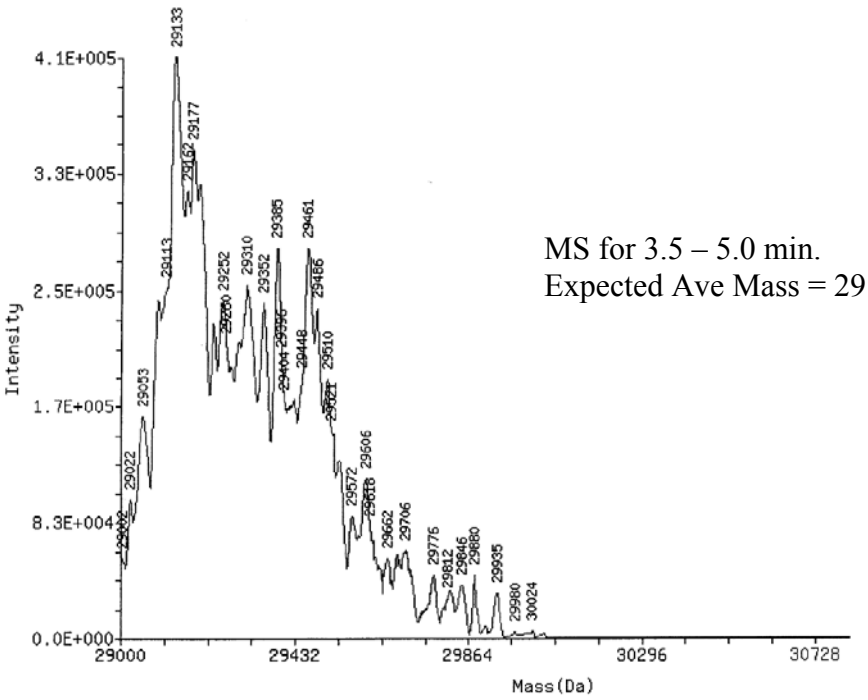
Nature Chemical Biology: doi:10.1038/nchembio.211

Plate 3 Well B2 Triazine Installation and Amine Substitution



[<<] [Top] [ESI Mass Spectrum] [Deconvolution] [Deconvolution Peak Report] [Raw Data] [Log File]

Zoom Display Deconvoluted Mass Spectrum:

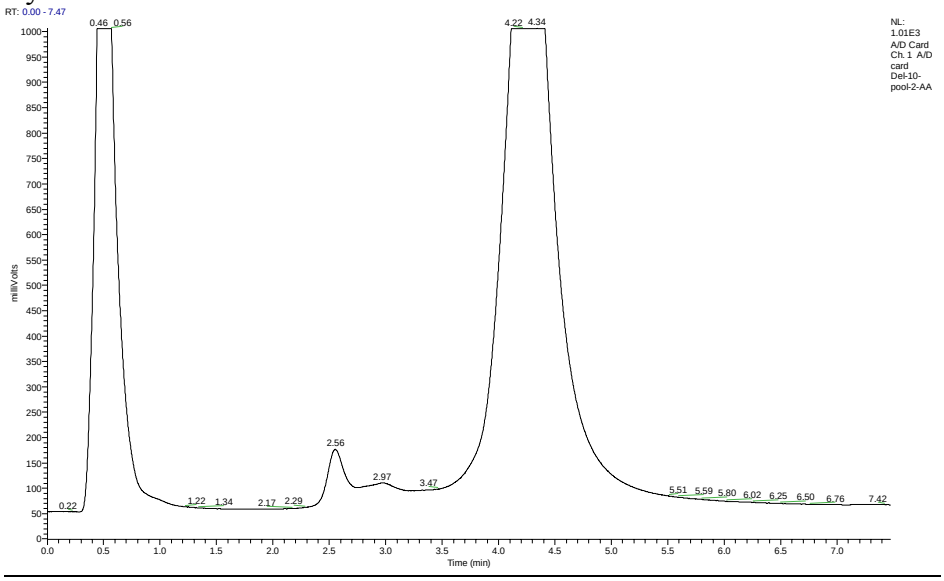


MS for 3.5 – 5.0 min.
Expected Ave Mass = 29159

Figure 10. LC chromatogram and deconvoluted mass spectrum of Cycle 2 triazine installation and amine substitution from well B2 of plate3.

Nature Chemical Biology: doi:10.1038/nchembio.211

Cycle 2 Pooled Product



[<<] [Top] [ESI Mass Spectrum] [Deconvolution] [Deconvolution Peak Report] [View Data] [Log File]

Zoom Display Deconvoluted Mass Spectrum:

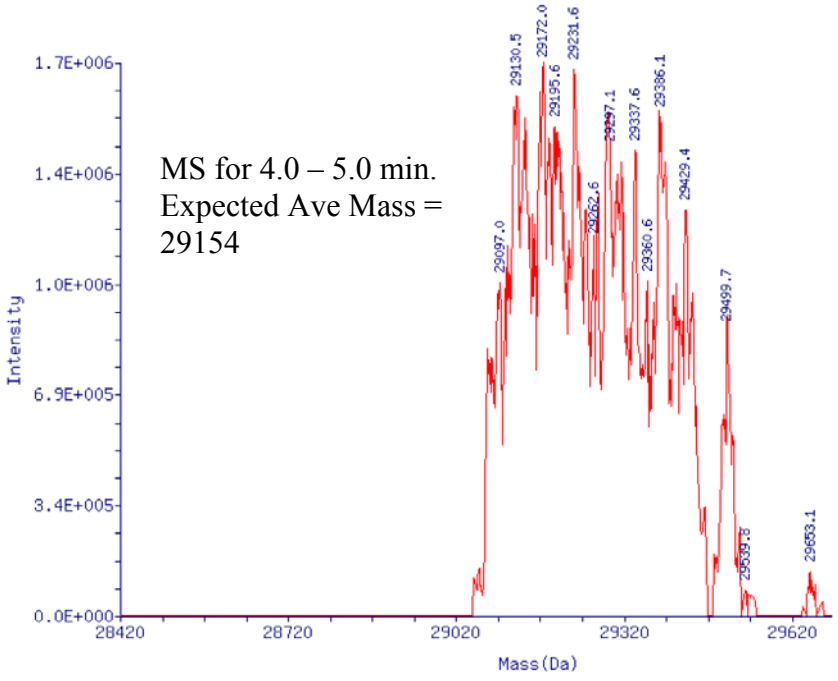
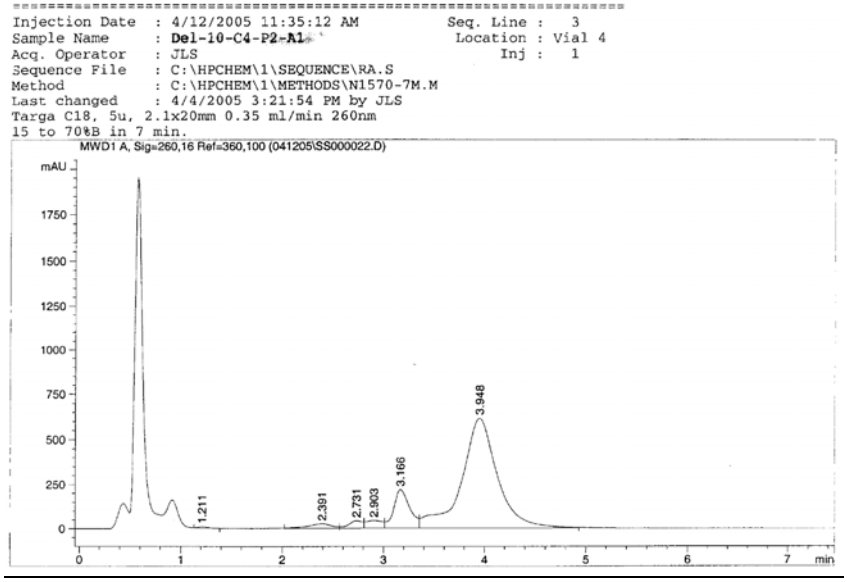


Figure 11. LC chromatogram and deconvoluted mass spectrum of pooled Cycle 2 product. No purification was performed before entry into Cycle 3.

Nature Chemical Biology: doi:10.1038/nchembio.211

Cycle 3 Analysis

Plate 2 Well A1 Ligation (36,864 components)



[<<] [Top] [ESI Mass Spectrum] [Deconvolution] [Deconvolution Peak Report] [Raw Data] [Log File]

Zoom Display Deconvoluted Mass Spectrum:

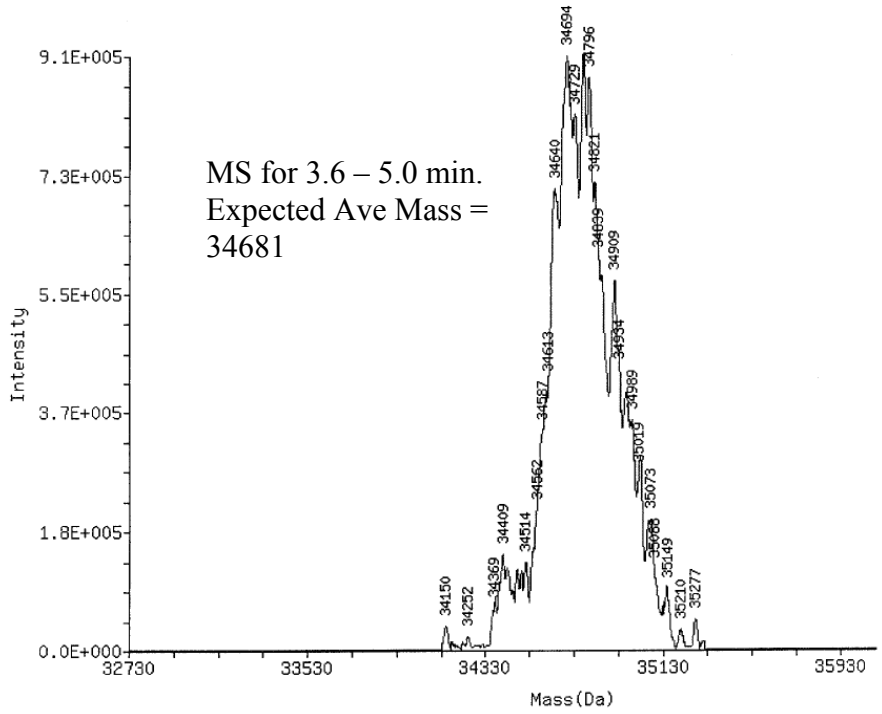
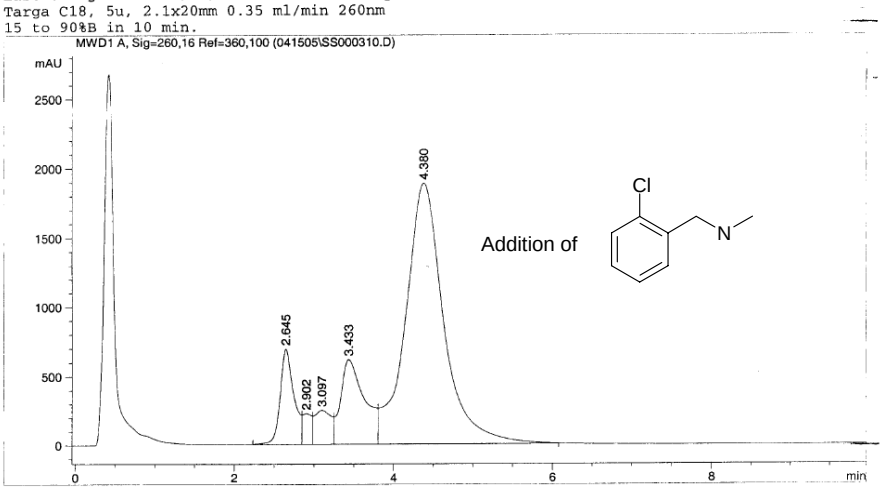


Figure 12. LC chromatogram and deconvoluted mass spectrum of Cycle 3 ligation from well A1 of plate 2.

Nature Chemical Biology: doi:10.1038/nchembio.211

Plate 2 Well A1 Amine Substitution

```
==================================================================
Injection Date  : 4/15/2005 5:35:54 PM          Seq. Line :  11
Sample Name     : Del-10-P2-A1-2AA               Location : Vial 31
Acq. Operator   : JLS                               Inj :   1
Sequence File   : C:\HPCHEM\1\SEQUENCE\RA.S
Method          : C:\HPCHEM\1\METHODS\N_1590-2.M
Last changed    : 1/12/2005 11:36:56 AM by JLS
Targa C18, 5u, 2.1x20mm 0.35 ml/min 260nm
15 to 90%B in 10 min.
```



[<<] [Top] [ESI Mass Spectrum] [Deconvolution] [Deconvolution Peak Report] [Raw Data] [Log File]

Zoom Display Deconvoluted Mass Spectrum:

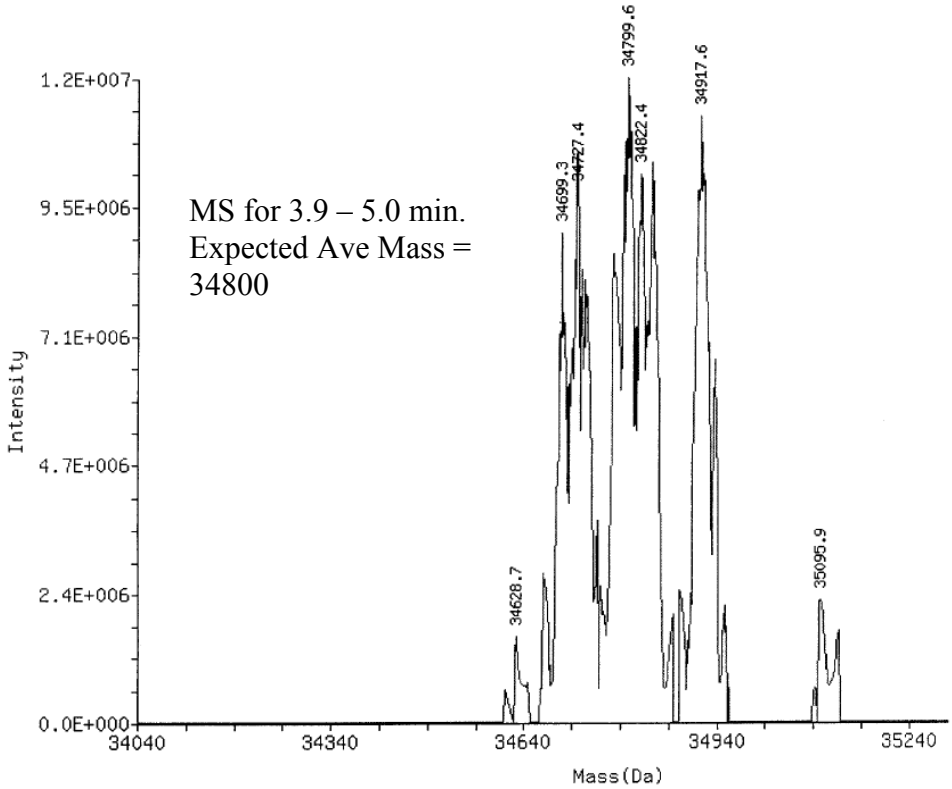


Figure 13. LC chromatogram and deconvoluted mass spectrum of Cycle 3amine subsitution from well A1 of plate 2.

Nature Chemical Biology: doi:10.1038/nchembio.211

Final library, pooled and purified

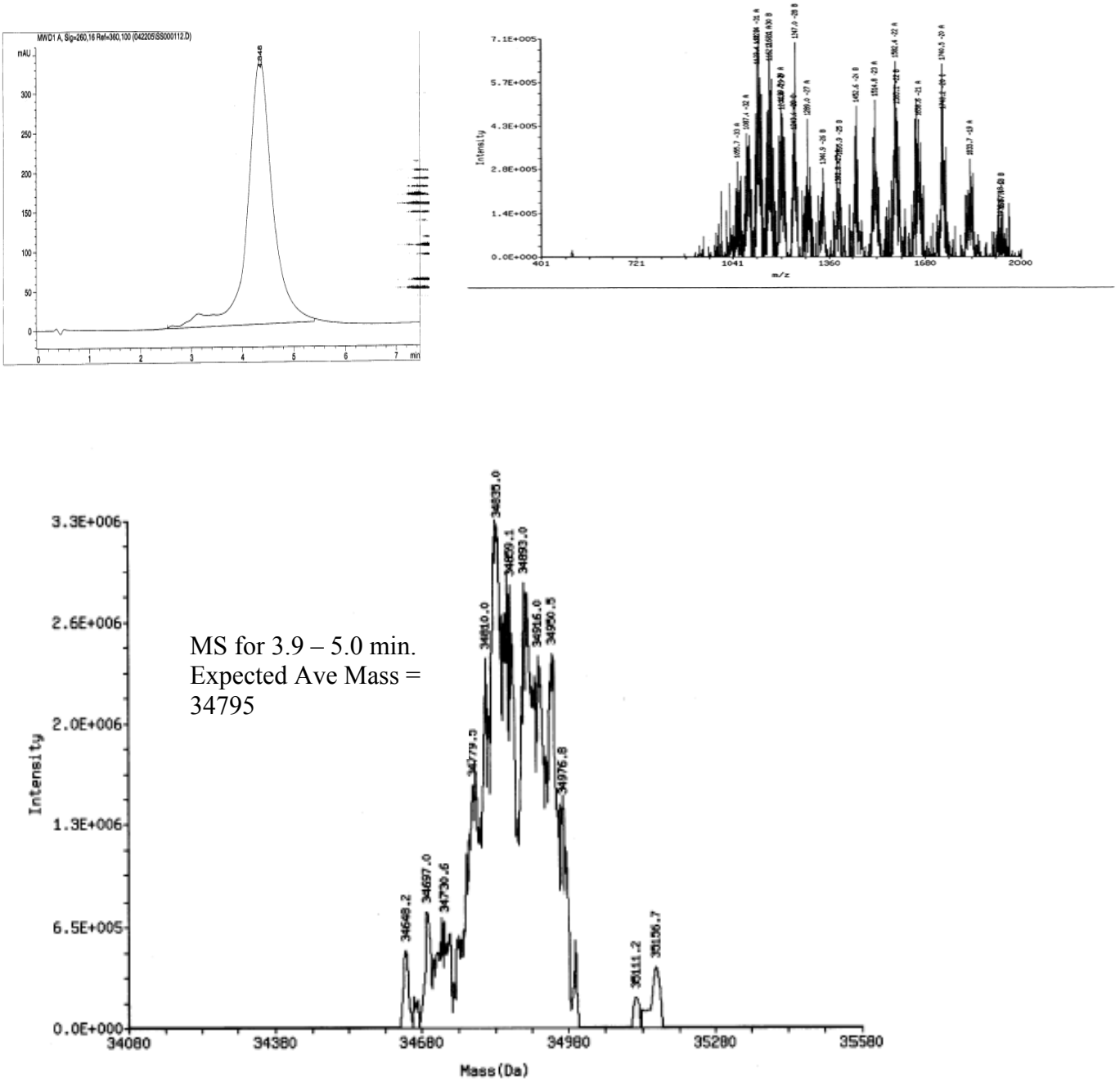


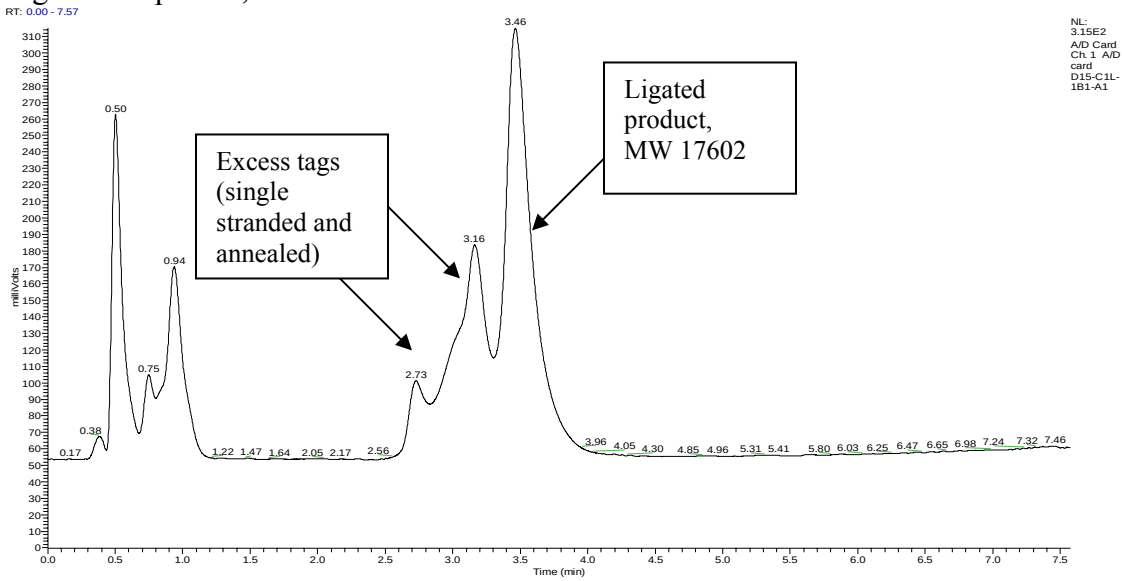
Figure 14. LC chromatogram, raw ESI mass spectrum, and deconvoluted mass spectrum of DEL A final product (ca. 7 million components)

Nature Chemical Biology: doi:10.1038/nchembio.211

Characterization of DEL B

Cycle 1 Analysis

Ligation of plate 2, well A1



Acylation of plate 2, well A1

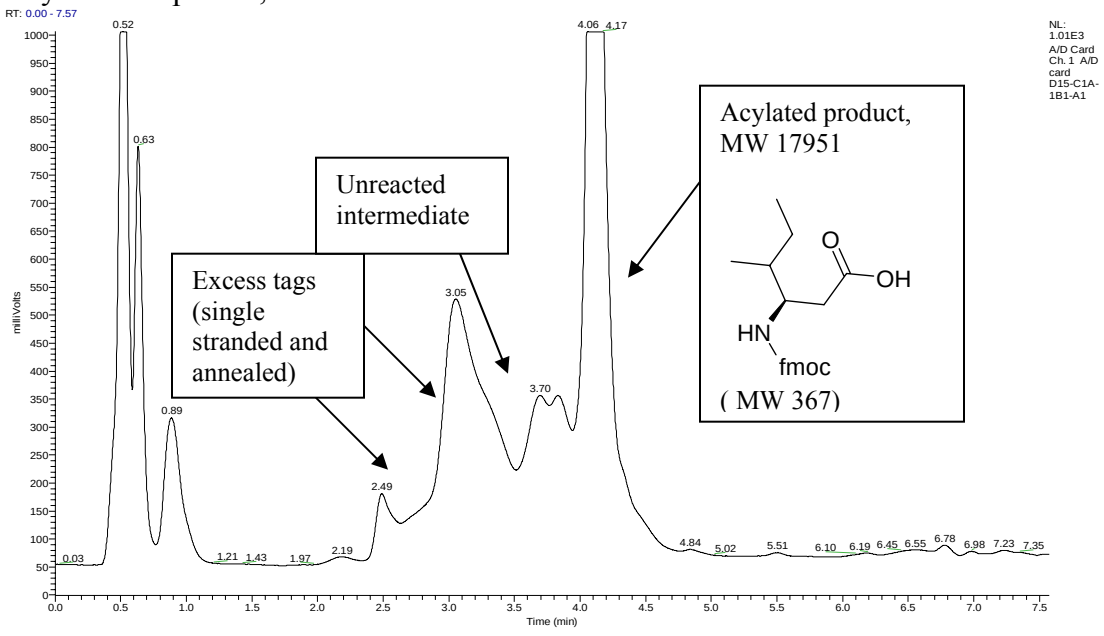
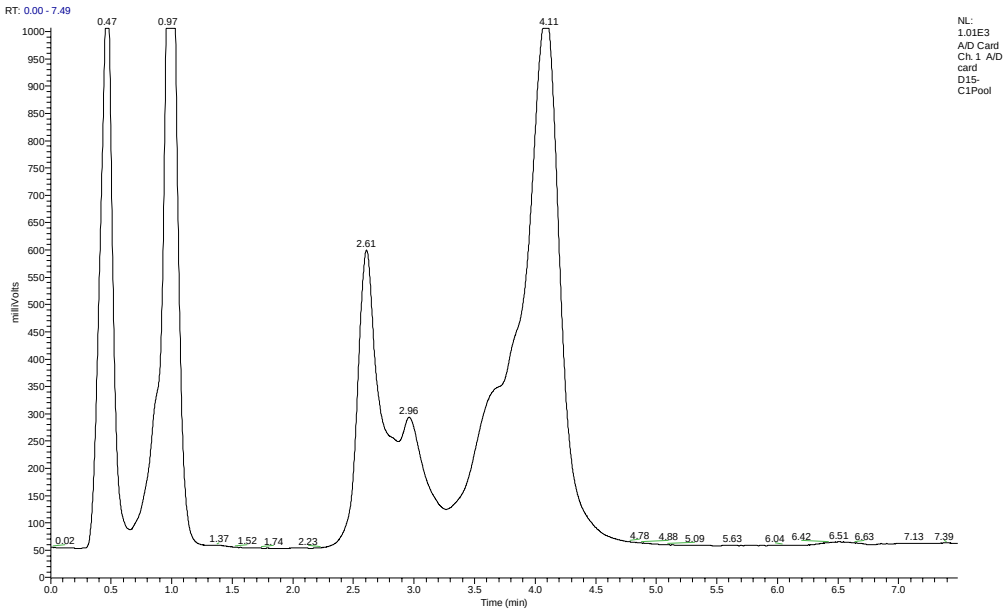


Figure 15. LC chromatogram of ligation and acylation reactions from plate 2 well A1.

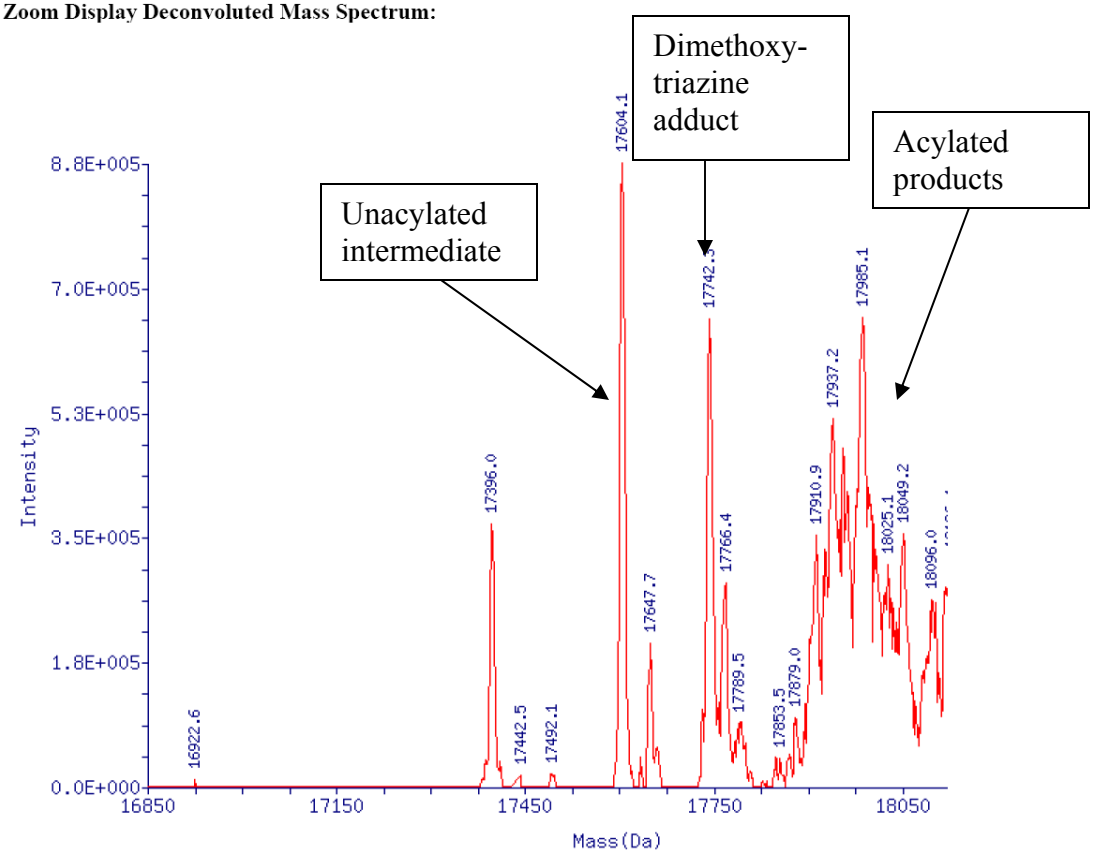
Nature Chemical Biology: doi:10.1038/nchembio.211

Crude, pooled Cycle 1 product



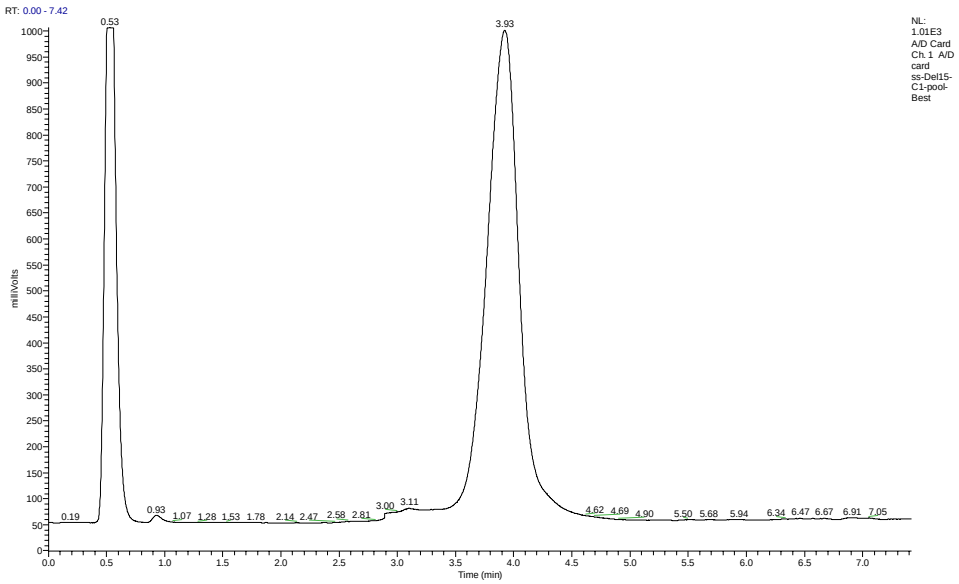
[<<] [Top] [ESI Mass Spectrum] [Deconvolution] [Deconvolution Peak Report] [View Data] [Log File]

Zoom Display Deconvoluted Mass Spectrum:



Nature Chemical Biology: doi:10.1038/nchembio.211

Figure 16. LC chromatogram and deconvoluted mass spectrum of crude pooled Cycle 1 product.
Purified Cycle 1 product (before deprotection)



[<<] [Top] [ESI Mass Spectrum] [Deconvolution] [Deconvolution Peak Report] [View Data] [Log File]

Zoom Display Deconvoluted Mass Spectrum:

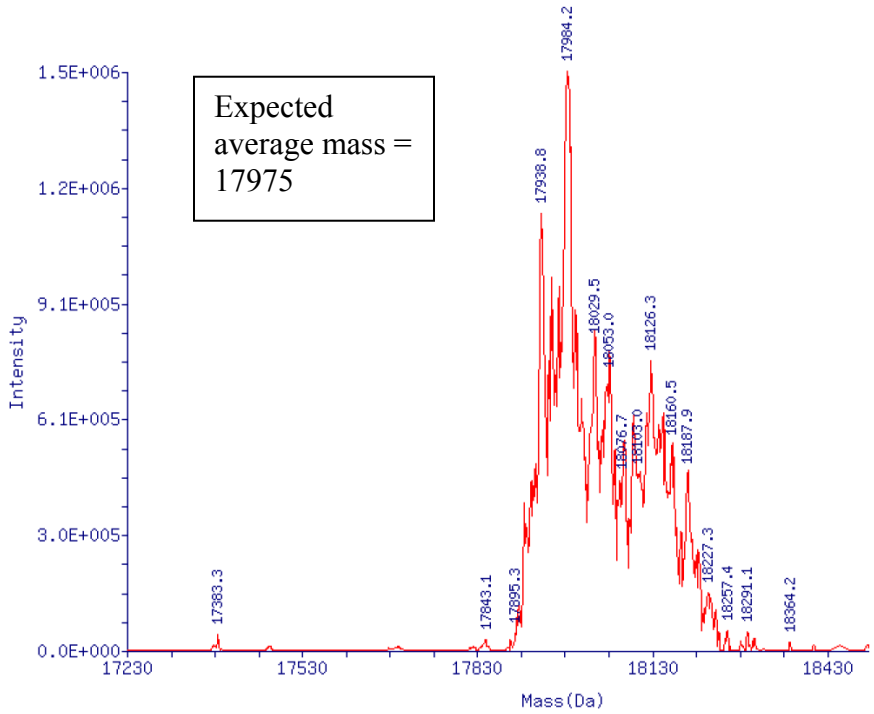
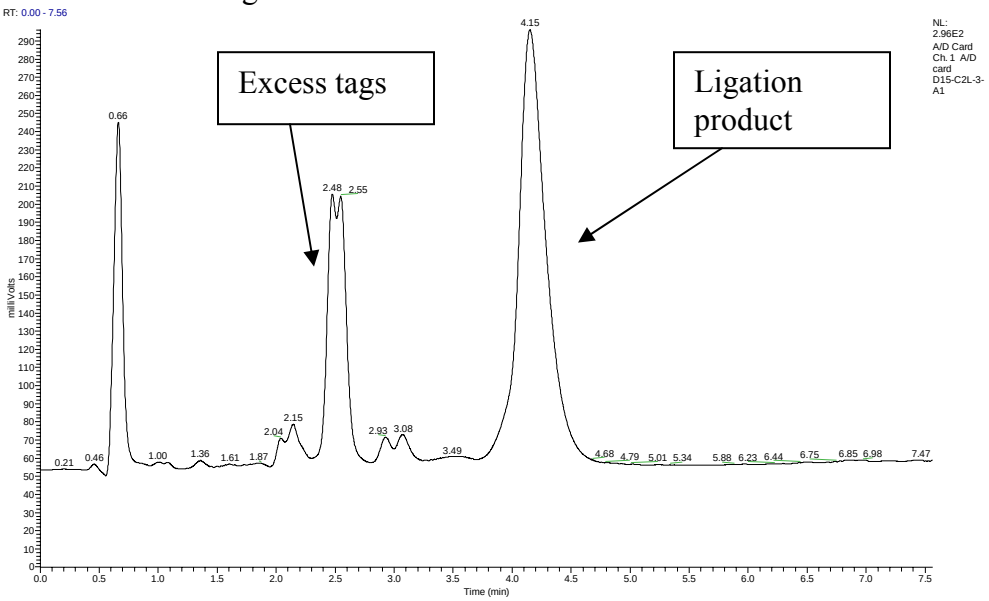


Figure 17. LC chromatogram and deconvoluted mass spectrum of purified Cycle 1 product.

Nature Chemical Biology: doi:10.1038/nchembio.211

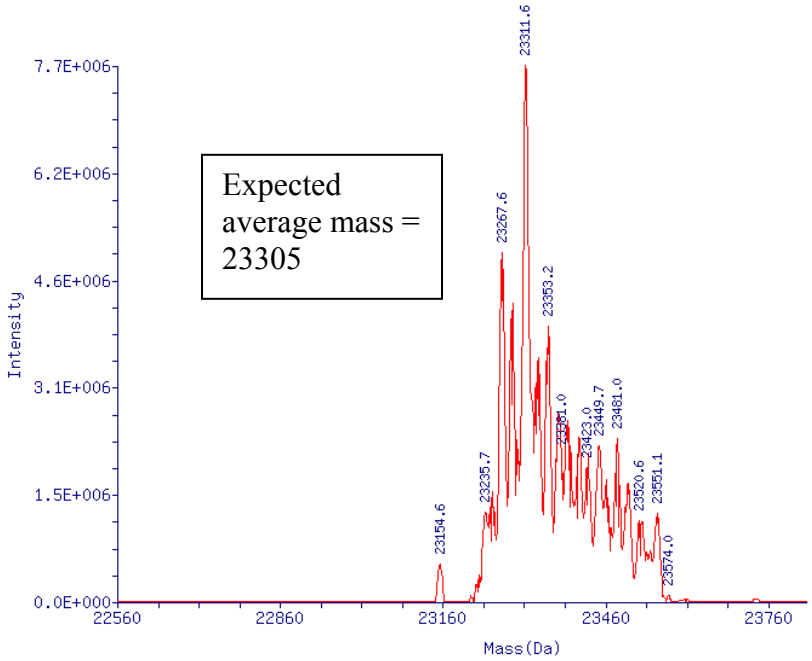
Cycle 2 Analysis

Plate 3 Well B2 Ligation



[<<] [Top] [ESI Mass Spectrum] [Deconvolution] [Deconvolution Peak Report] [View Data] [Log File]

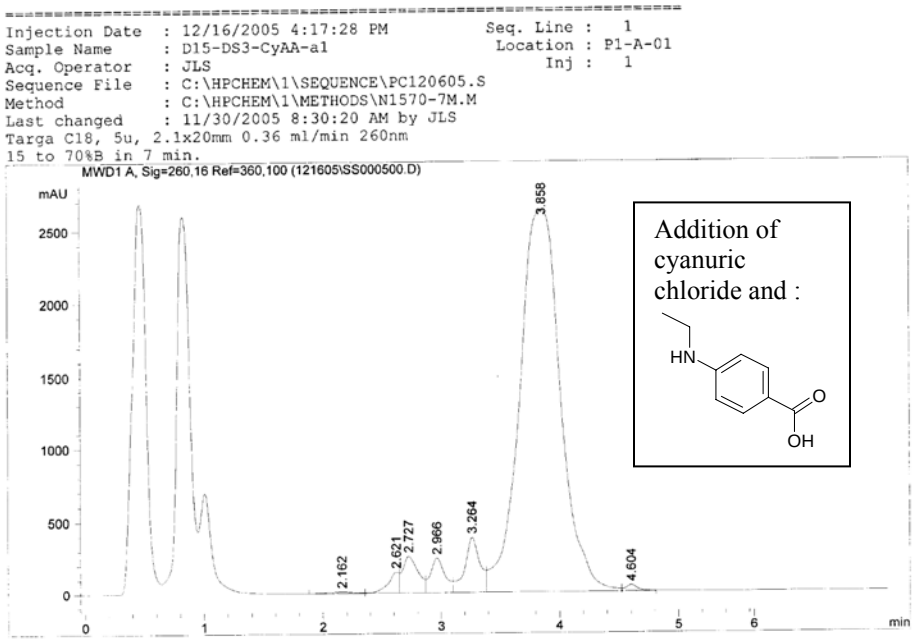
Zoom Display Deconvoluted Mass Spectrum:



Nature Chemical Biology: doi:10.1038/nchembio.211

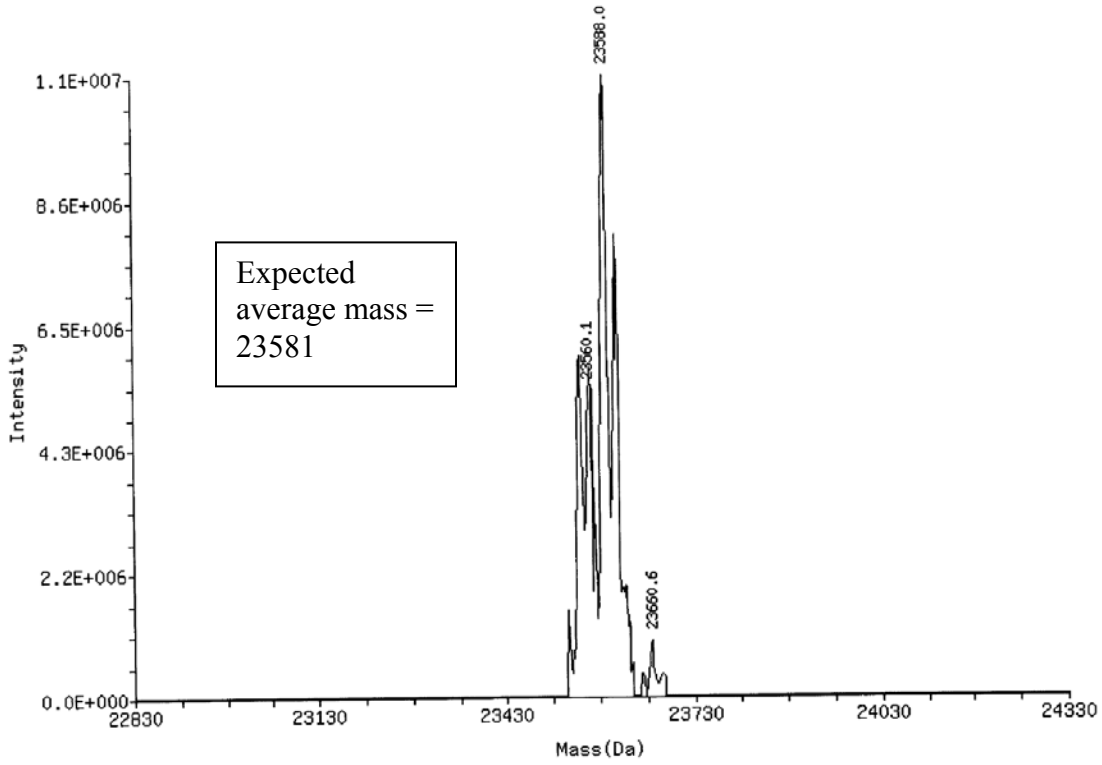
Figure 17. LC chromatogram and deconvoluted mass spectrum of ligation reaction from Cycle 2 plate 3 well B2.

Plate 3 Well B2 Triazine Installation and Amine Substitution



[<<] [Top] [ESI Mass Spectrum] [Deconvolution] [Deconvolution Peak Report] [Raw Data] [Log File]

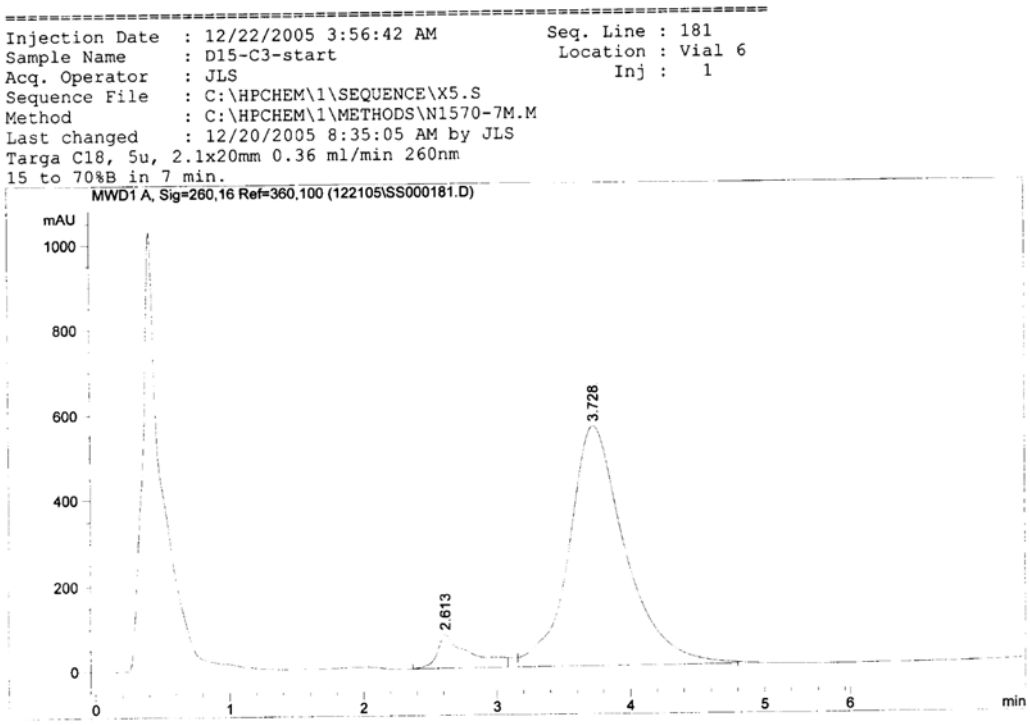
Zoom Display Deconvoluted Mass Spectrum:



Nature Chemical Biology: doi:10.1038/nchembio.211

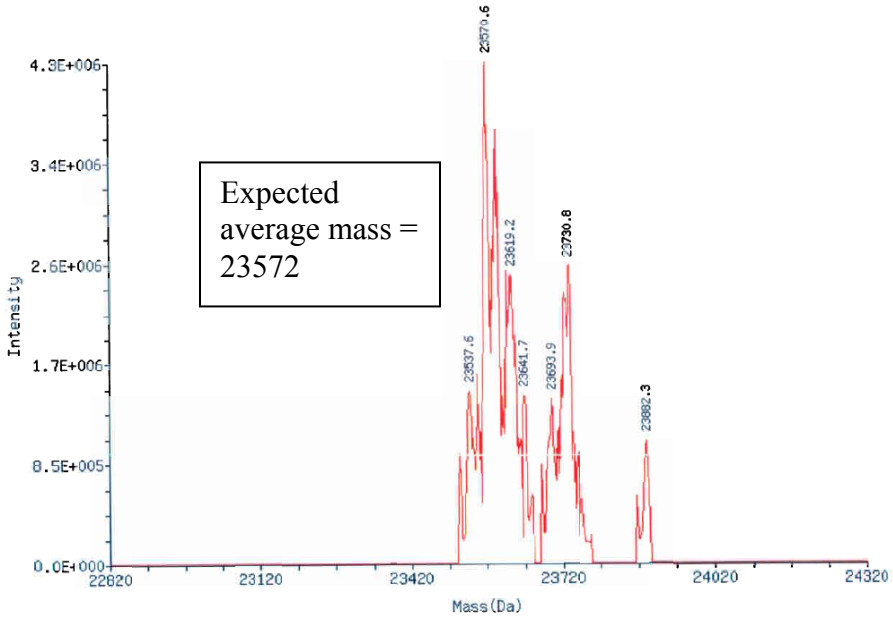
Figure 18. LC chromatogram and deconvoluted mass spectrum of triazine installation and amine substitution reactions from Cycle 2 plate 3 well B2.

Cycle 2 Product, pooled and precipitated.



[<<] [Top] [ESI Mass Spectrum] [Deconvolution] [Deconvolution Peak Report] [Raw Data] [Log File]

Zoom Display Deconvoluted Mass Spectrum:

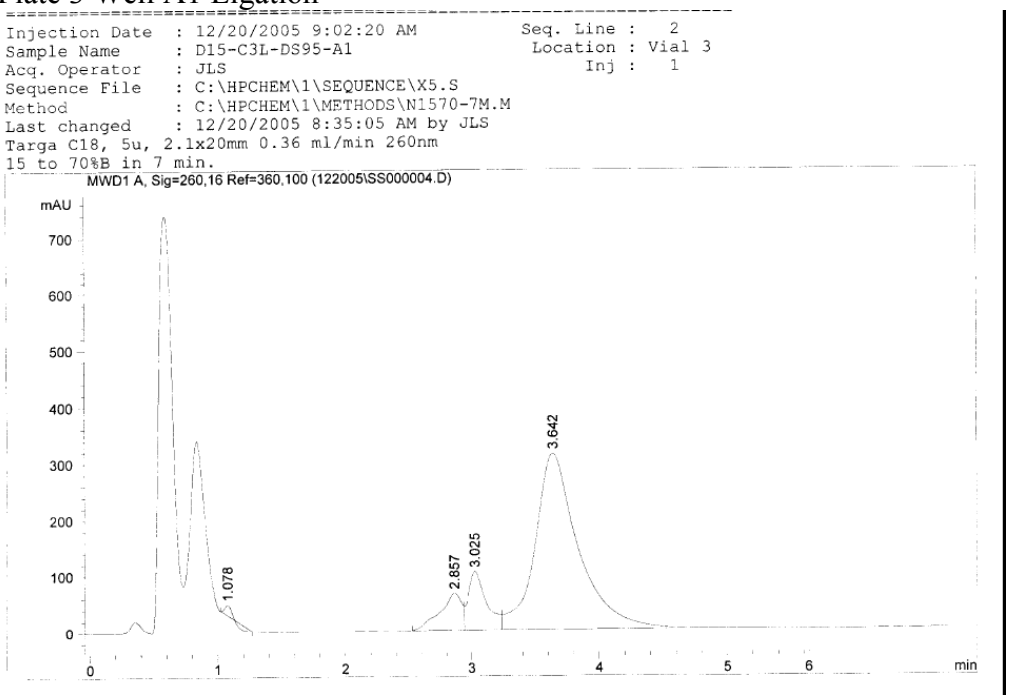


Nature Chemical Biology: doi:10.1038/nchembio.211

Figure 19. LC chromatogram and deconvoluted mass spectrum of pooled Cycle 2 product. (6,144 components)

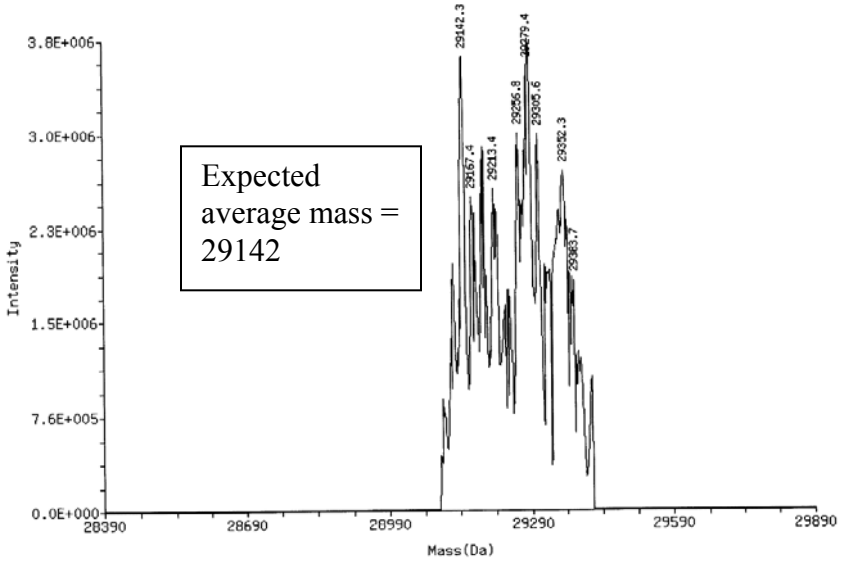
Cycle 3 Analysis

Plate 3 Well A1 Ligation



[<<] [Top] [ESI Mass Spectrum] [Deconvolution] [Deconvolution Peak Report] [Raw Data] [Log File]

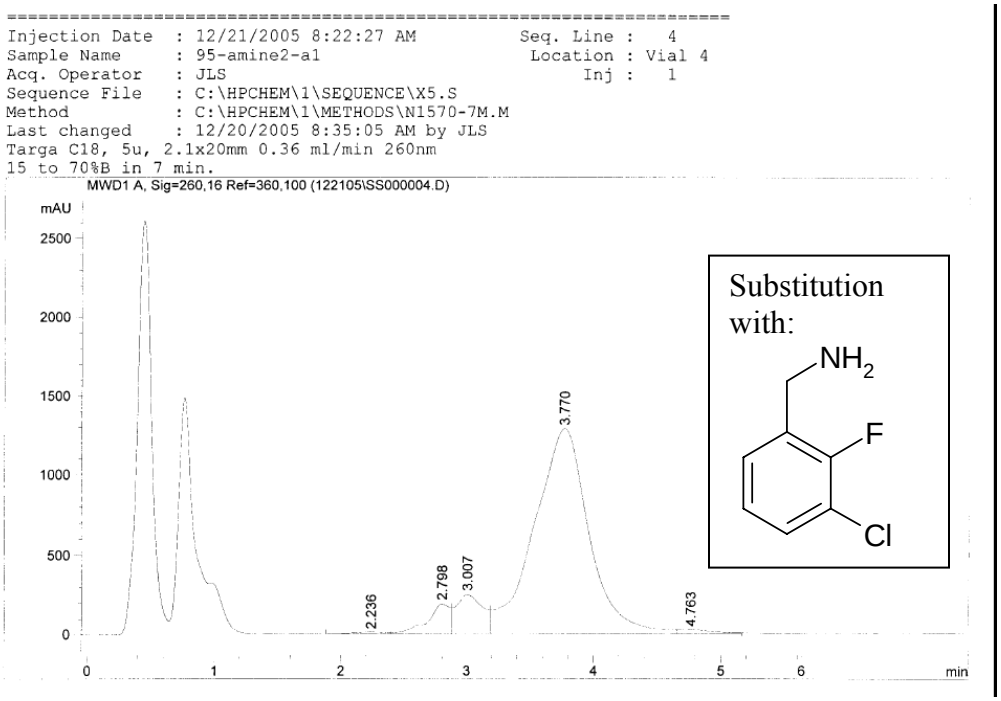
Zoom Display Deconvoluted Mass Spectrum:



Nature Chemical Biology: doi:10.1038/nchembio.211

Figure 20. LC chromatogram and deconvoluted mass spectrum of ligation reaction from Cycle 3 plate 3 well A1.

Plate 3 Well A1 Amine Substitution



Nature Chemical Biology: doi:10.1038/nchembio.211

[<<] [Top] [ESI Mass Spectrum] [Deconvolution] [Deconvolution Peak Report] [Raw Data] [Log File]
Zoom Display Deconvoluted Mass Spectrum:

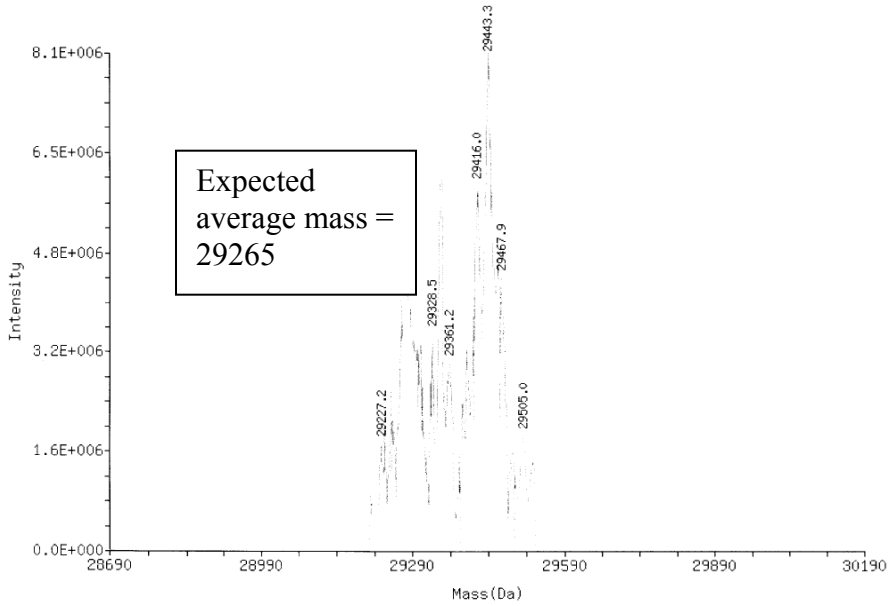
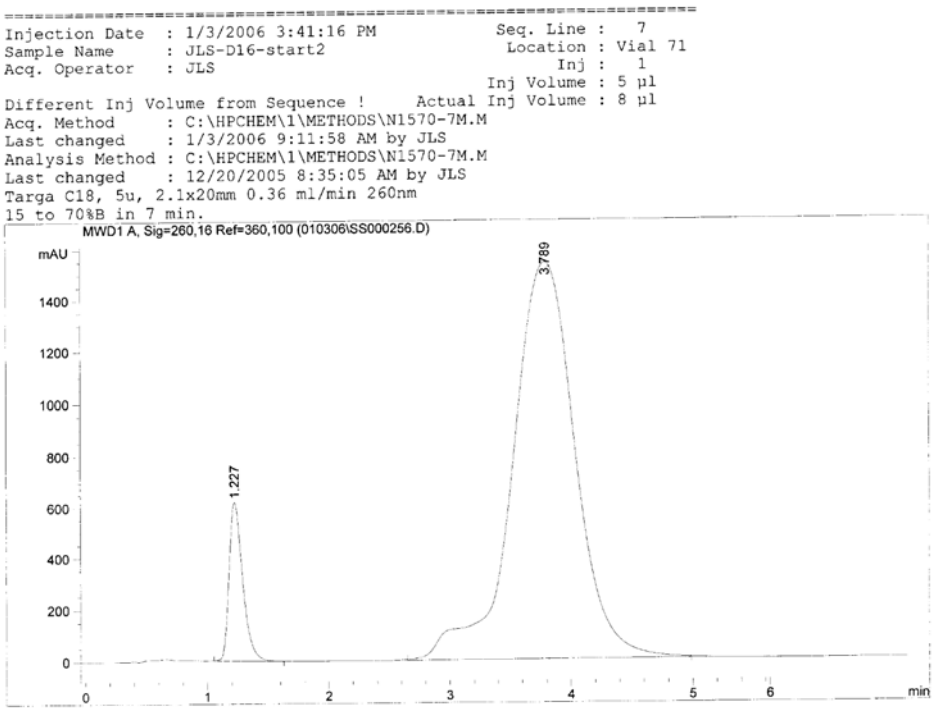


Figure 21. LC chromatogram and deconvoluted mass spectrum of substitution reaction from Cycle 3 plate 3 well A1.

Pooled and Purified Cycle 3 Product



Nature Chemical Biology: doi:10.1038/nchembio.211

[<<] [Top] [ESI Mass Spectrum] [Deconvolution] [Deconvolution Peak Report] [Raw Data] [Log File]
Zoom Display Deconvoluted Mass Spectrum:

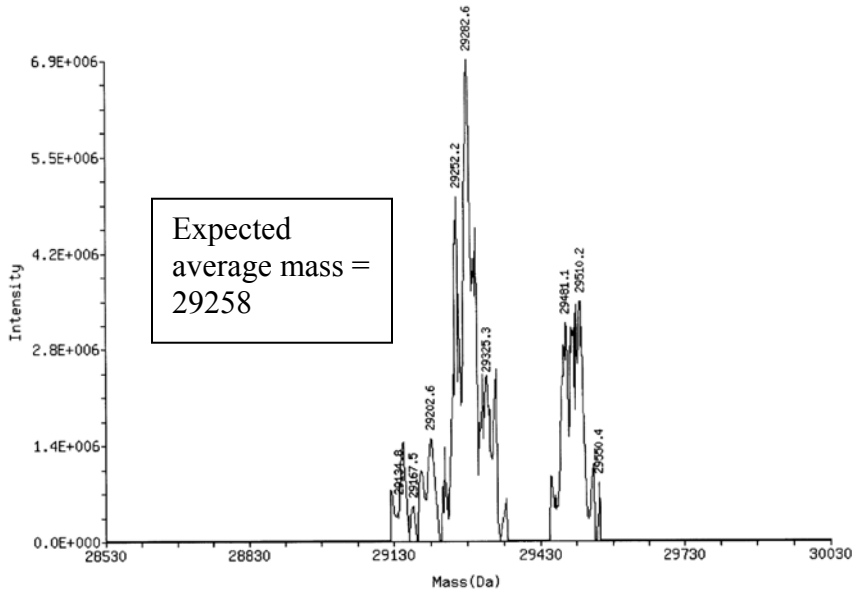
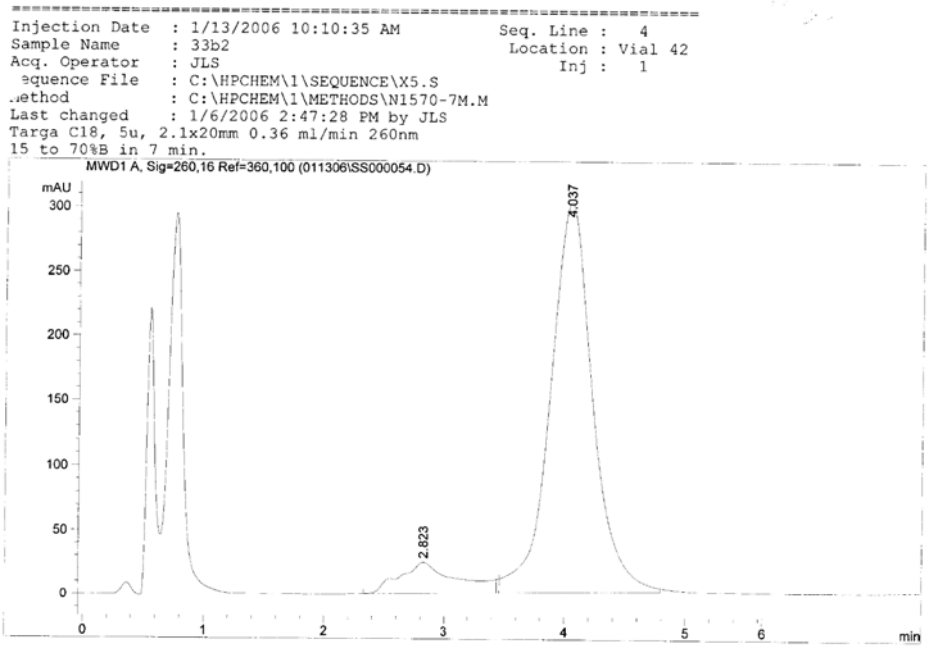


Figure 22. LC chromatogram and deconvoluted mass spectrum of purified cycle 3 product (ca. 2 million components).

Cycle 4 Analysis

Plate 1 Well B2 Ligation



Nature Chemical Biology: doi:10.1038/nchembio.211

[<<] [Top] [ESI Mass Spectrum] [Deconvolution] [Deconvolution Peak Report] [Raw Data] [Log File]
Zoom Display Deconvoluted Mass Spectrum:

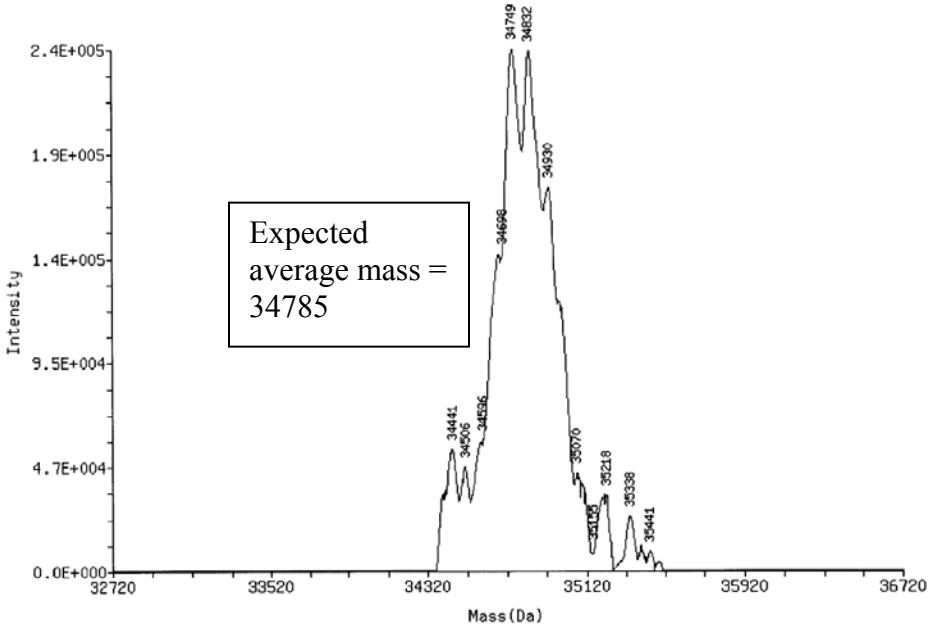
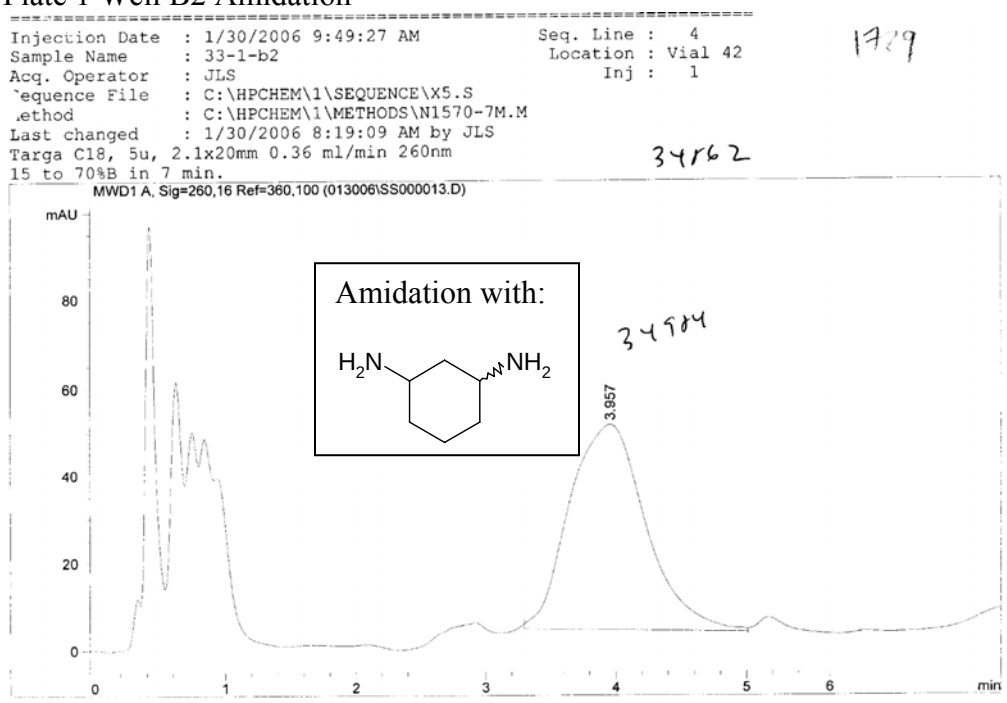


Figure 23. LC chromatogram and deconvoluted mass spectrum of ligation reaction from Cycle 4 plate 1 well B2.

Plate 1 Well B2 Amidation



Nature Chemical Biology: doi:10.1038/nchembio.211

[<<] [Top] [ESI Mass Spectrum] [Deconvolution] [Deconvolution Peak Report] [Raw Data] [Log File]
Zoom Display Deconvoluted Mass Spectrum:

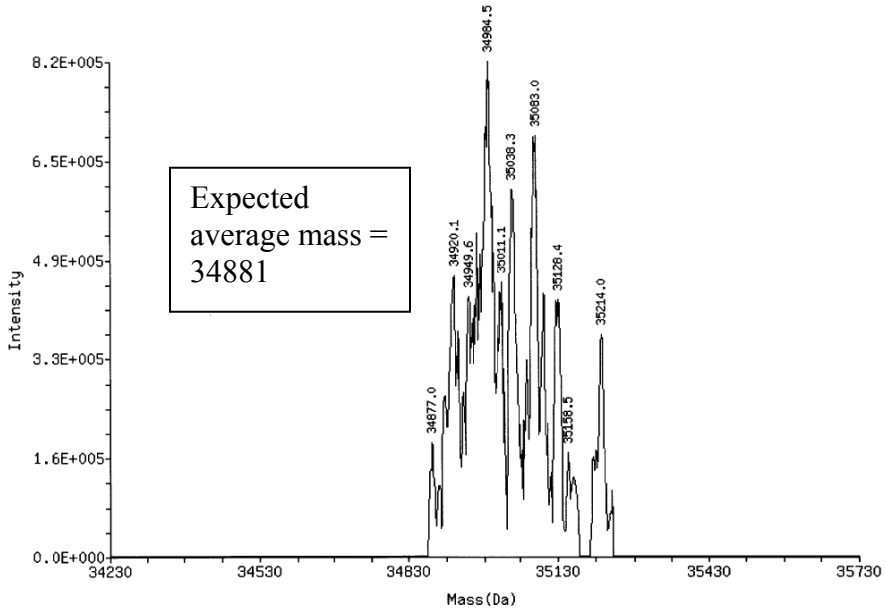
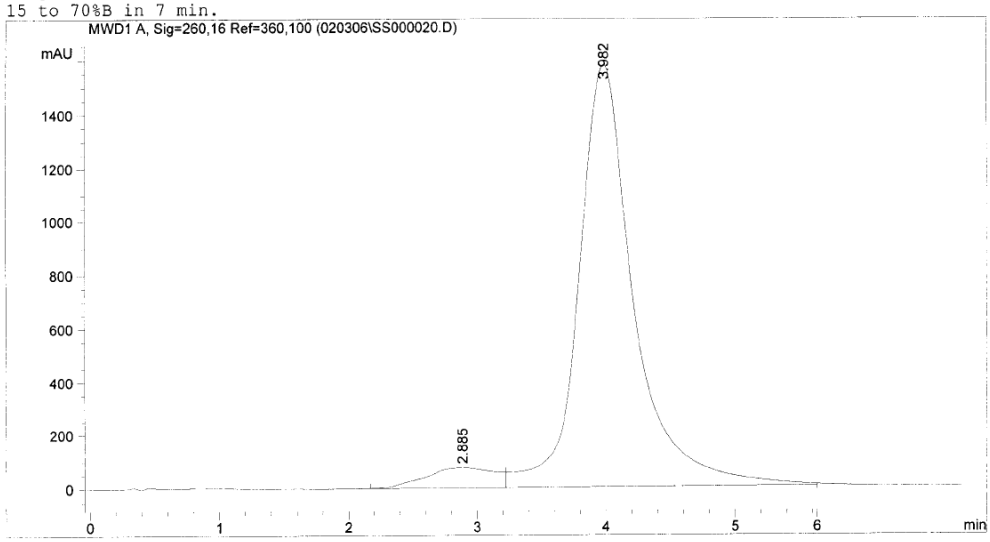


Figure 24. LC chromatogram and deconvoluted mass spectrum of amidation reaction from Cycle 4 plate 1 well B2.

DEL B Final Product, Pooled and Purified

```
Injection Date  : 2/3/2006 1:39:20 PM            Seq. Line :  20
Sample Name     : DEL-16-Final                   Location : Vial 41
Acq. Operator   : JLS                                Inj :   1
Sequence File   : C:\HPCHEM\1\SEQUENCE\X5.S
Method          : C:\HPCHEM\1\METHODS\N1570-7M.M
Last changed    : 1/30/2006 8:19:09 AM by JLS
Targa C18, 5u, 2.1x20mm 0.36 ml/min 260nm
15 to 70%B in 7 min.
```



Nature Chemical Biology: doi:10.1038/nchembio.211

[<<] [Top] [ESI Mass Spectrum] [Deconvolution] [Deconvolution Peak Report] [Raw Data] [Log File]
Zoom Display Deconvoluted Mass Spectrum:

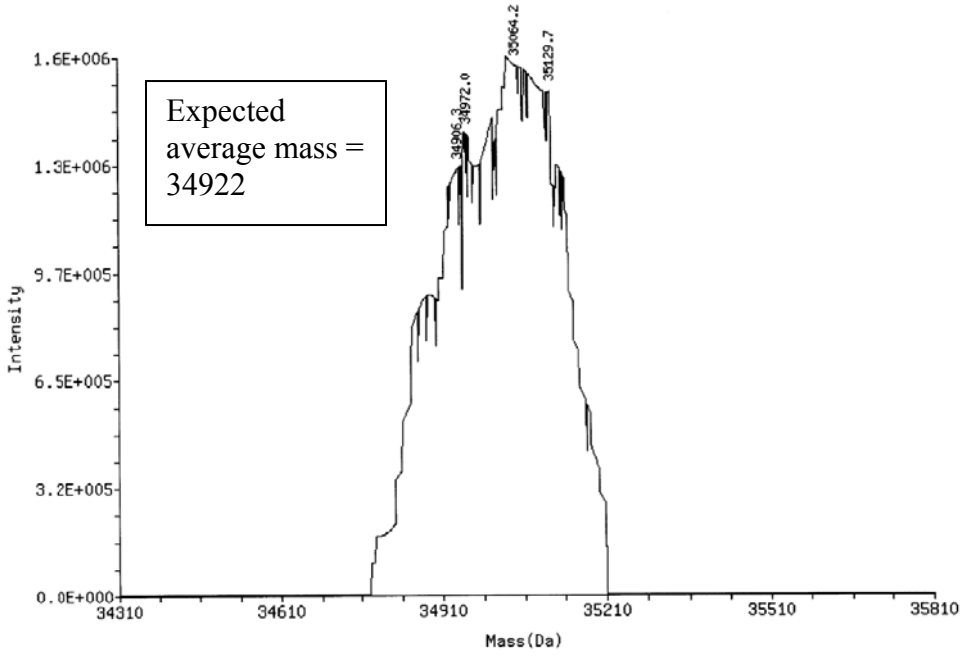


Figure 25. LC chromatogram and deconvoluted mass spectrum of purified DEL-B final product (ca. 800 million components).

Summary of Additional Files

Synthon Lists

Supplementary Dataset 1 (DEL-A Synthon List)
Supplementary Dataset 2 (DEL-B Synthon List)

Synthon lists for DEL-A and DEL-B are included as Excel spreadsheets. They include cycle number, synthon structure, and the tag sequences. Many synthons occur multiple times since they were encoded by multiple tags within a given cycle. The tag sequence is written in the 5′→3′ direction and corresponds to the bottom strand of the tags as represented in the tag descriptions above.
Thus the DEL-B Cycle 2 tag: ATGCAGGCT would correspond to the following full tag structure:

5′ CCTGCATGT 3′
3′ TCGGACGTA 5′

Nature Chemical Biology: doi:10.1038/nchembio.211

Selection Output

The filtered output from all of the selections described in the manuscript are included. These Excel files contain the structures of the synthons (as SMILES strings) for a selected library member, as well as the number of times that library member was observed.

Supplementary Dataset 3 (DEL-A Selection - p38 MAPK):

Only species that occurred 10 or more times are shown, replicating data shown in the cube in Figure 3c. A background-binding disynthon, corresponding to the combination of 4-azidophenylalanine in Cycle 1 and 2-aminopyridine N-oxide in Cycle 2, is visible in addition to the selected AMMB-containing structures. As stated in the text, species that were selected in the No Target Controls were designated background binders and omitted from analysis of target-based selections.

Supplementary Dataset 4 (DEL-A Selection - Aurora A, Method A)
Supplementary Dataset 5 (DEL-A Selection - Aurora A, Method B):

Two files are included, corresponding to selection Methods A and B. Both show species that occurred 4 or more times. The background binder mentioned above is visible in the Method B selection.

Supplementary Dataset 6 (DEL-B Selection - p38 MAPK):

Only species contained in the cube shown in detail in Figure 5 are included. These are species with occurences greater than 1, and containing benzimidazole-5-carboxylate at Cycle 2.

Nature Chemical Biology: doi:10.1038/nchembio.211