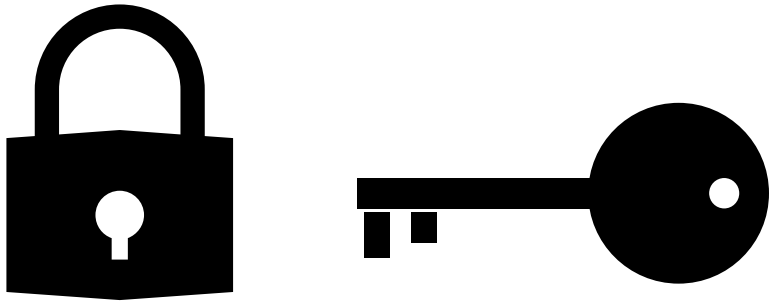
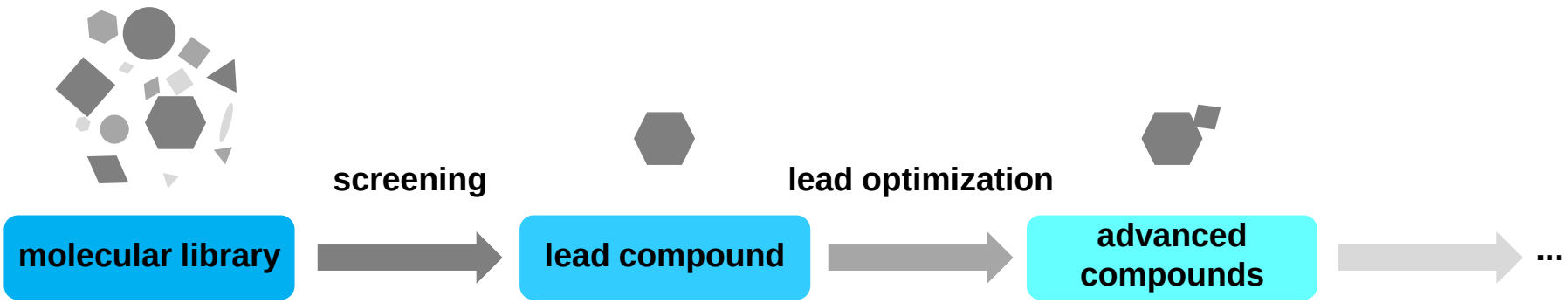
Encoding and Decoding Molecular Library
A Recent Example



Hiroaki Itoh

Sep. 9, 2023 | Literature Seminar

Library for Screening Bioactive Compounds



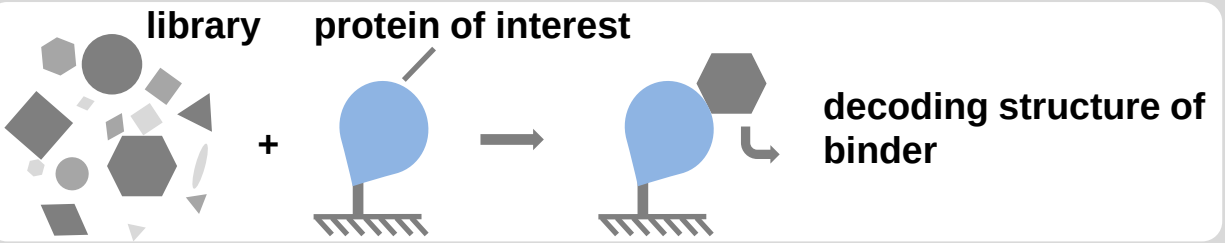
physical
(compound collection
/synthetic
/genetically encoded)
virtual

Molecular library and its screening system are important for efficient discovery of useful bioactive compounds as drug/agrochemical candidates

Library Strategy for Affinity Selection

high-throughput screening by functional assays

affinity selection[1)]



phage display library[2)]
peptides comprising proteinogenic amino acids, 10^6–10^{11}
decoding by DNA sequencing

mRNA display library
peptides/cyclic peptides proteinogenic (and nonproteinogenic) amino acids, 10^{12}–10^{13}
decoding by DNA sequencing

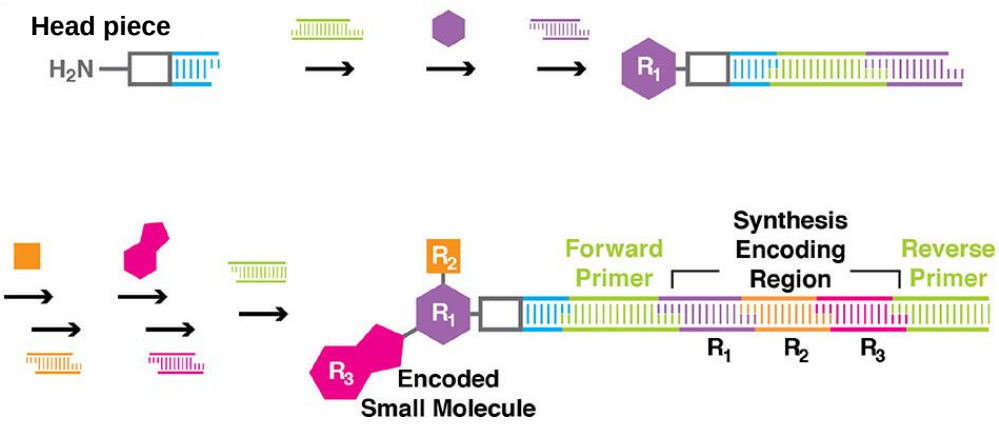
DNA-encoded library (DEL)[3)]
synthetic small molecules, 10^5–10^9
decoding by DNA sequencing

“self-encoded” library
current applications are limited to synthetic peptides/peptidomimetics,[4,5)] 10^8
decoding by MS^n

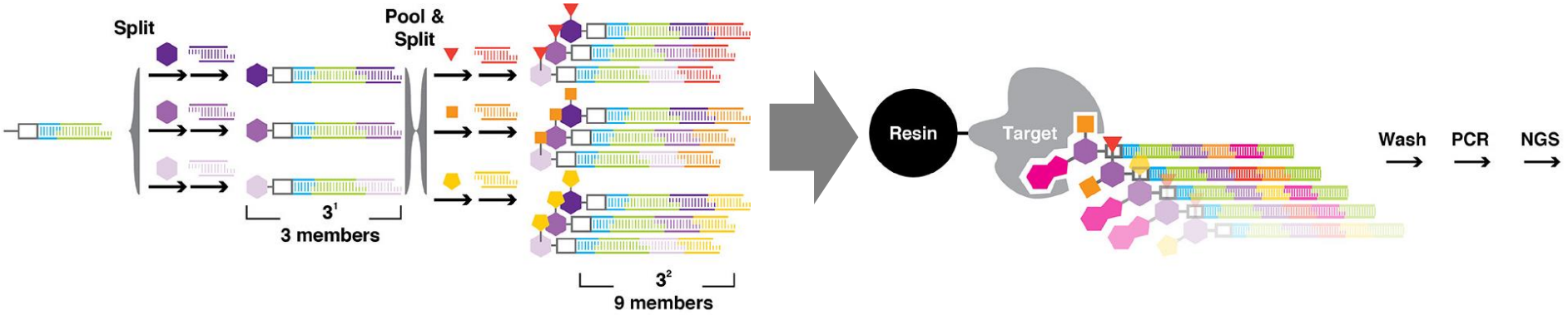
1) Mata, J. M.; van der Nol, E.; Pomplun, S. J. *J. Am. Chem. Soc.* **2023**, doi: 10.1021/jacs.3c04899. 2) For the previous literature seminar for phage display, see: 180616_LS_Koichi_Kamiya_Library_Strategy_for_Cyclic_Peptides 3) For the previous literature seminar for DEL, see: 171125_LS_Yuri_Takada_DNA-encoded_library
4) For SpeedScreen, see: Muckenschnabel, I.; Falchetto, R.; Mayr, L.; Filipuzzi, I. *Anal. Biochem.* **2004**, *324*, 241.
5) For ALIS, see: Annis, D. A.; Athanasopoulos, J.; Curran, P. J.; Felsch, J. S.; Kalghatgi, K.; Lee, W. H.; Nash, H. M.; Orminati, J. P. A.; Rosner, K. E.; Shipps, G. W.; Thaddupathy, G. R. A.; Tyler, A. N.; Vilenchik, L.; Wagner, C. R.; Wintner, E. A. *Int. J. Mass Spectrom.* **2004**, *238*, 77.

DNA-Encoded Library

- encoded synthesis by enzymatic ligation and building block coupling



- split-and-mix diversification/affinity selection/PCR amplification and structure decoding by DNA sequencing

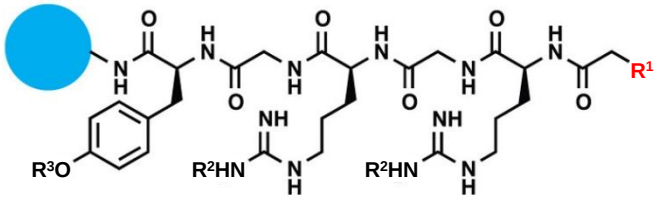


DNA-encoded library (DEL) is widely used for screening synthetic small molecules

1) Fitzgerald, P. R.; Paegel, B. M. *Chem. Rev.* **2021**, *121*, 7155.

What is the “DNA-Compatible” Reaction?

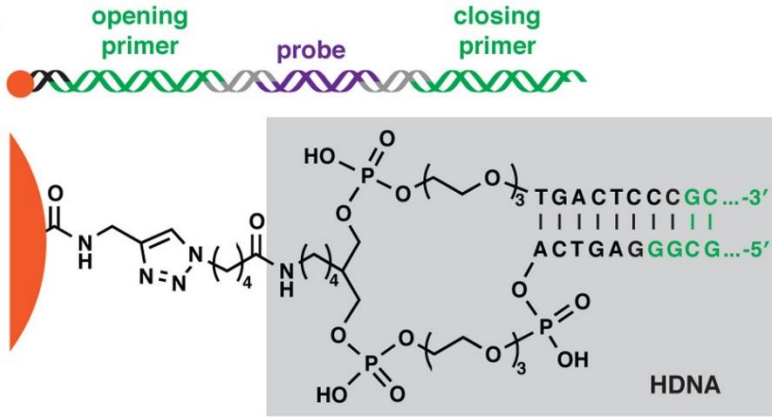
■ **synthesis resin or modified oligonucleotides**

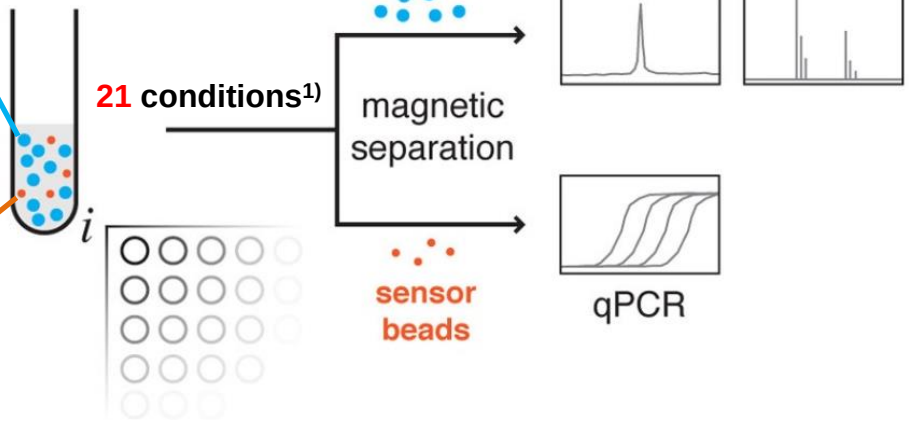


3'-TTACGGTCA-5'

$(\ )_6NHR^4$

■ **magnetic sensor bead**



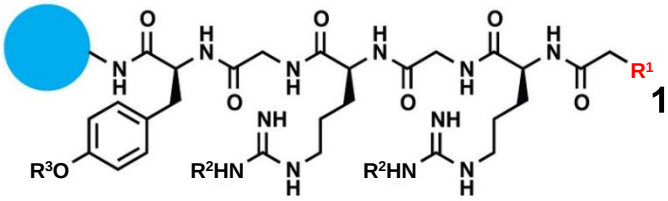


Yield of each reaction and amplifiable DNA were separately evaluated

1) Malone, M. L.; Paegel, B. M. *ACS Comb. Sci.* **2016**, *18*, 182.

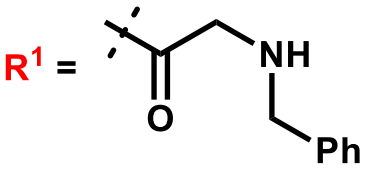
Difference in Yield and DNA Remaining

- selected example of solid-phase reaction[1)]

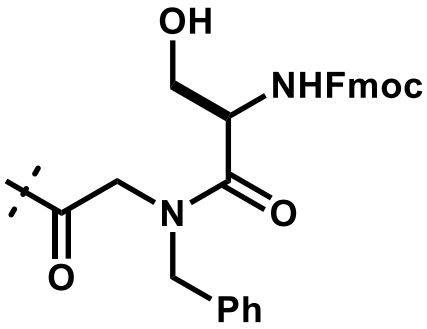


1. Fmoc-L-Ser(*t*-Bu)-OH
   i-PrN=C=N*i*-Pr
   Oxyma Pure
   2,4,6-collidine, DMF
   3 h

2. separation of sensor beads;
   CF_3CO_2H/CH_2Cl_2 /*i*-Pr_3SiH (90/5/5)

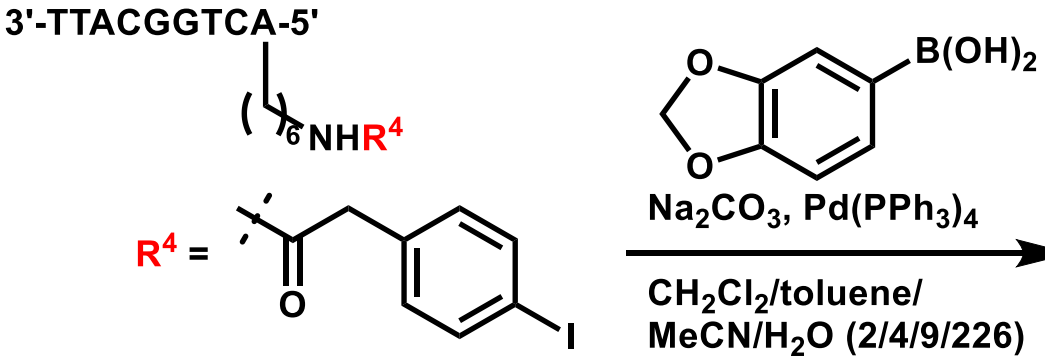


R^2 = Pbf, R^3 = *t*-Bu



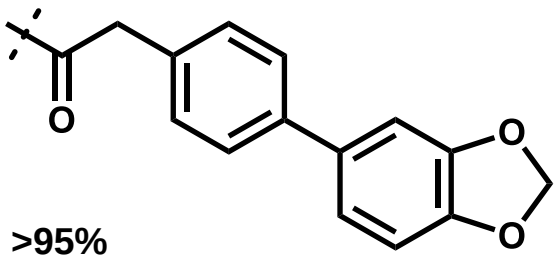
$R^2 = R^3$ = H
yield = >95%
% amplifiable DNA remaining = 15%

- selected example of solution-phase reaction[1)]



Na_2CO_3, $Pd(PPh_3)_4$

CH_2Cl_2/toluene/
MeCN/H_2O (2/4/9/226)



yield = >95%
% amplifiable DNA remaining = 30%

Robust tagging system is important for high-fidelity decoding of the library

1) Malone, M. L.; Paegel, B. M. *ACS Comb. Sci.* **2016**, *18*, 182. 2) For PNA-encoded library, see: Zambaldo, C.; Barluenga, S.; Winssinger, N. *Curr. Opin. Chem. Biol.* **2015**, *26*, 8.

Prof. Bradley L. Pentelute

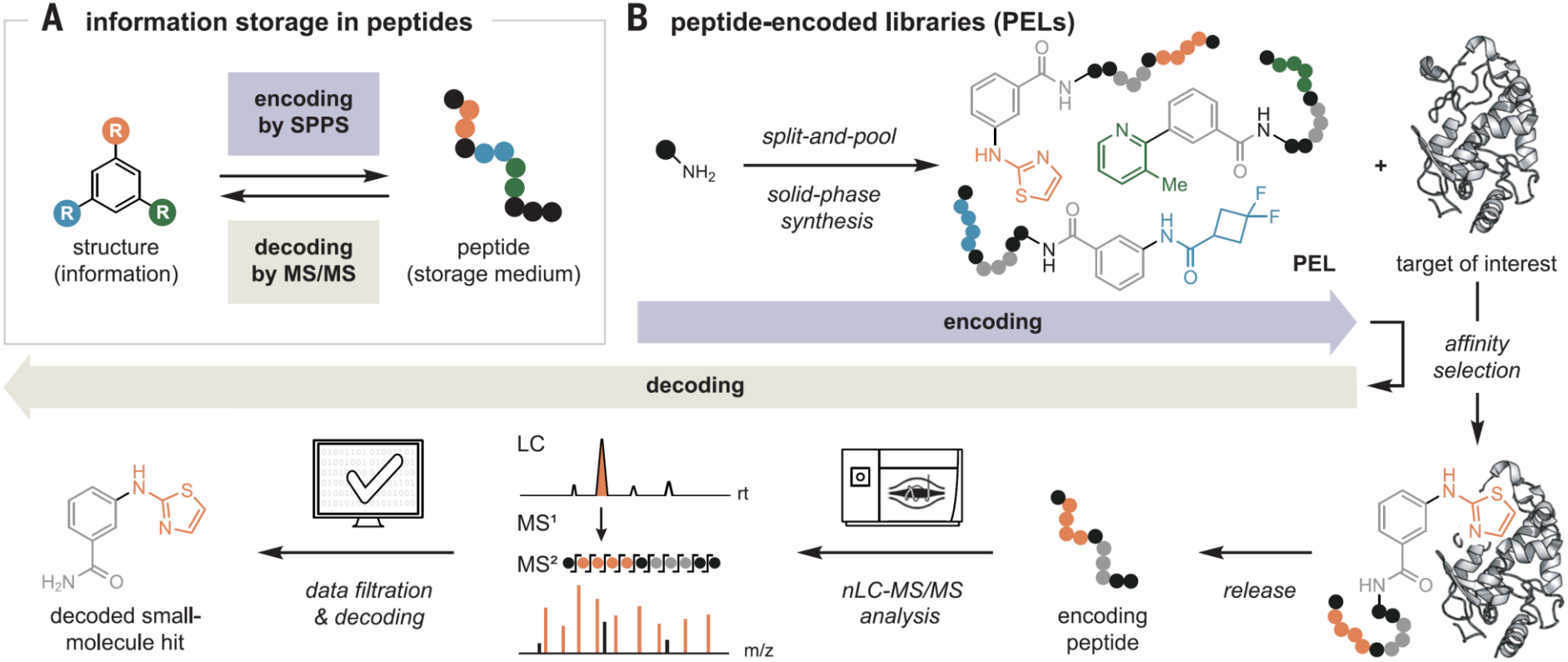
2021–present	Professor, Department of Chemistry, MIT
2016–2021	Associate Professor, Department of Chemistry, MIT
2011–2016	Assistant Professor, Department of Chemistry, MIT
2008–2011	Postdoctoral Fellow, Department of Microbiology and Molecular Genetics, Harvard Medical School, (Prof. R. John Collier)
1/2008–9/2008	Senior Scientist, Ethos Pharmaceuticals
2004–2008	Graduate Student, Department of Chemistry University of Chicago (Prof. Stephen B. H. Kent)
2003	B.S. in Chemistry and B.A. in Psychology from University of Southern California

Research focus:
Development of chemistry for protein modification to enhance the therapeutic properties

- **Arylation of unprotected peptides and proteins[1)]**
- **Automated flow biopolymer synthesis**
- **Intracellular delivery of biomolecules**
- **Peptidomimetic binders to proteins**

1) For the previous literature seminar for π-clamp, see:
160604_LS_Kotaro_Tokumoto_protein_bioconjugation_using_pi-clamp

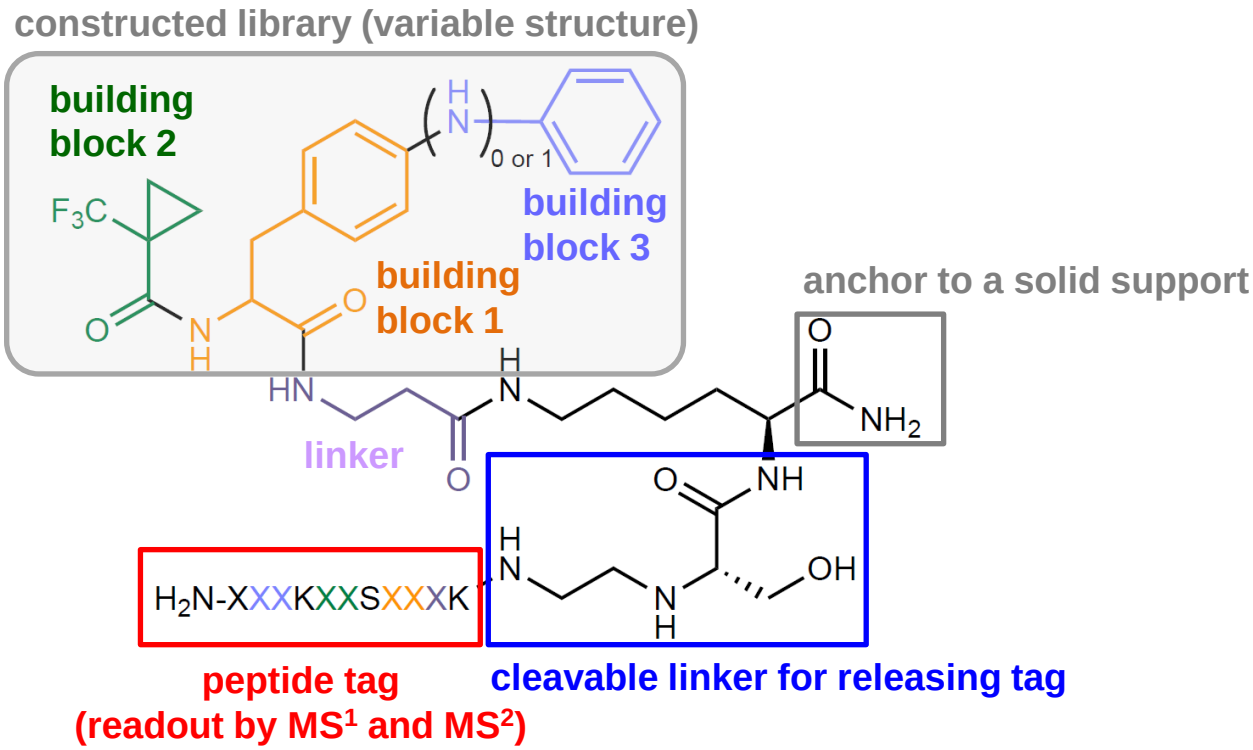
Peptide-Encoded Library



Structures are encoded by solid-phase peptide synthesis and decoded by MS [1)]

1) Rössler, S. L.; Grob, N. M.; Buchwald, S. L.; Pentelute, B. L. *Science* **2023**, *379*, 939.

Peptide-Encoded Library



Structures are encoded by solid-phase peptide synthesis and decoded by MS [1)]

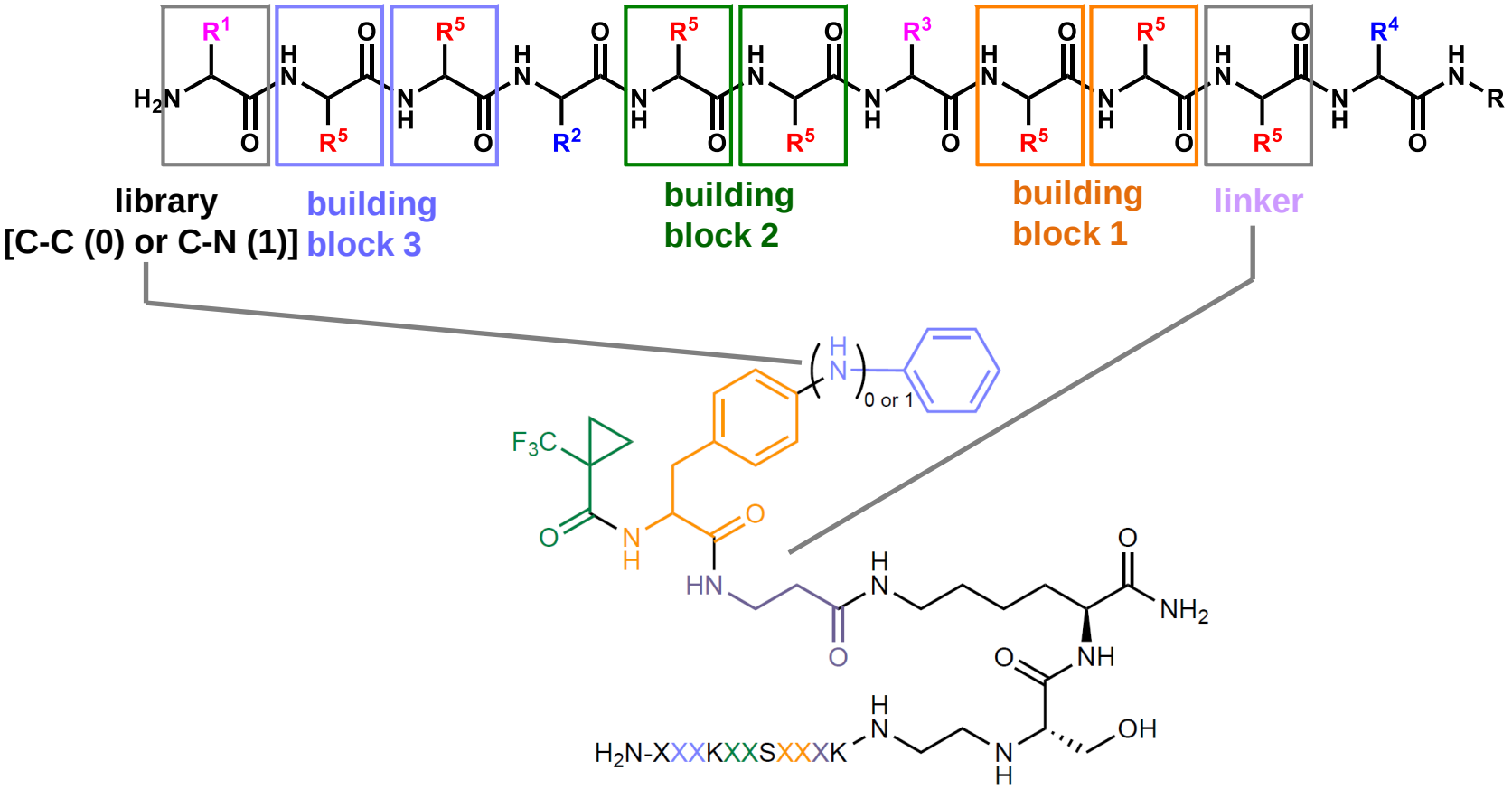
1) Rössler, S. L.; Grob, N. M.; Buchwald, S. L.; Pentelute, B. L. *Science* **2023**, *379*, 939.

Peptide Tag for Decoding

R^1, R^3: apolar/polar (non-charged) spacer monomer for enhancing sequencing confidence and solubility

R^2, R^4: basic group for increasing sequencing confidence and polarity tuning

R^5: coding region for each building block

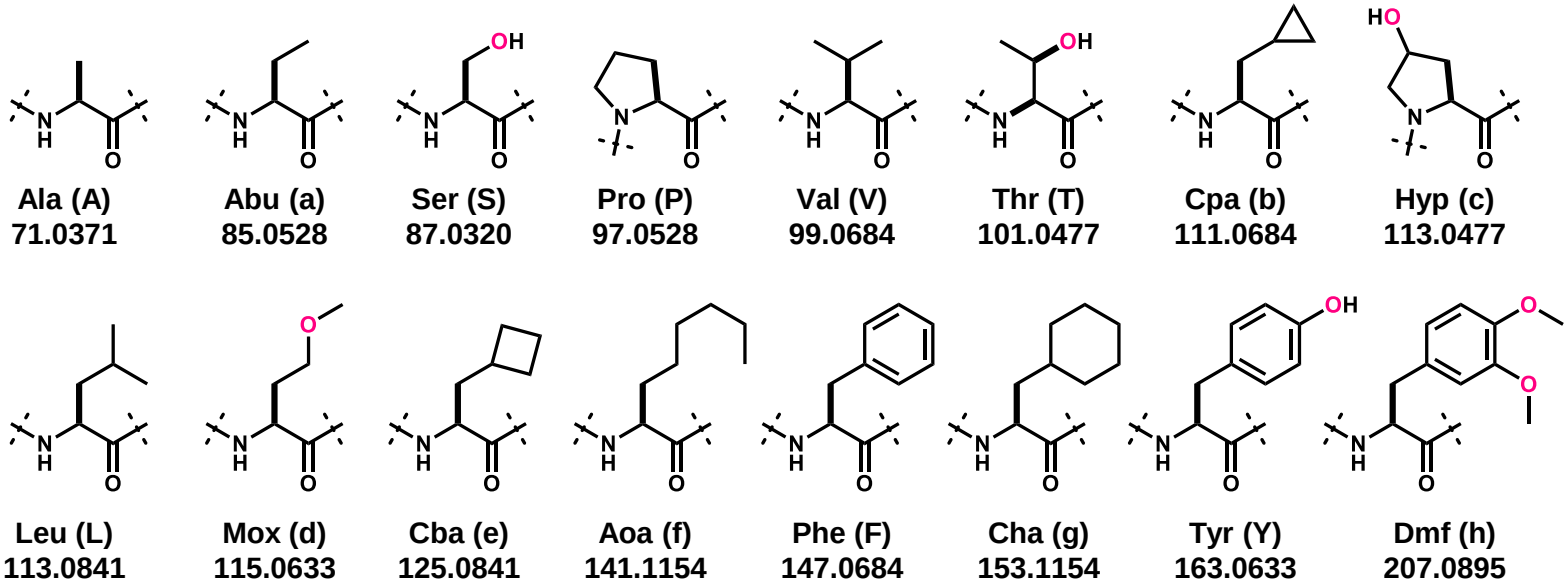


Amino Acids for Encoding

8 residues are coding region (8-digit memory cell)
16 amino acids are used for coding

theoretically maximal memory: $16^8 = 2^{32} =$ **4 GiB**
c.f. DNA (quaternary: ATGC): $4^8 = 2^{16} = 64$ KiB

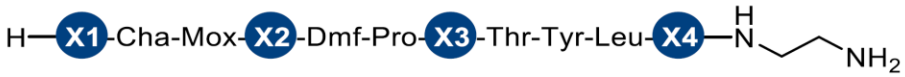
16 amino acid components for encoding



16 amino acids were adopted for encoding

Optimization of Four Residues

- **Optimization of residues X1–X4 for high-fidelity sequencing**



investigation using three basic amino acids (K, R, H) and two Val

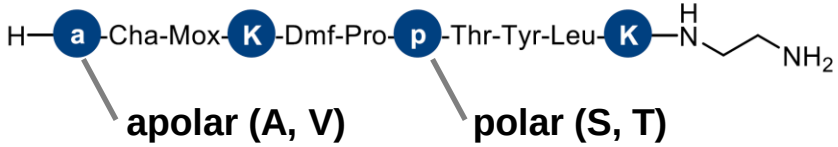
optimal position: X1 and X3 = K, X2 and X4 = V (based on ALC and the number of found sequences)

investigation using two of A, V, L, T and two Lys

KK*xx*
*x*K*x*K
K*xx*K
K*x*K*x*
*x*KK*x*
*xx*KK

***x* = A, V, L, or T**

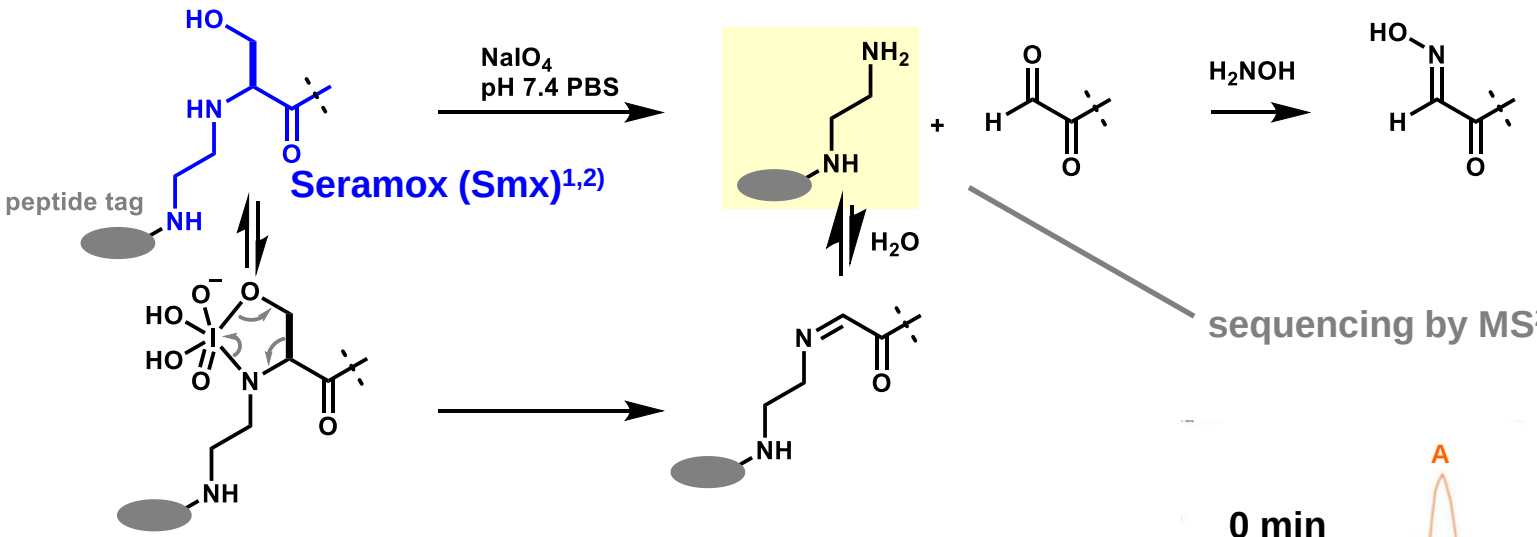
	Overall		Position X1		Position X2		Position X3		Position X4	
	avg. ALC (%)	found	avg. ALC (%)	found	avg. ALC (%)	found	avg. ALC (%)	found	avg. ALC (%)	found
A	95.5	32/48	96.2	9/12	97.2	6/12	93.6	8/12	94.9	9/12
V	96.1	25/48	95.0	8/12	96.9	7/12	94.5	4/12	98.2	6/12
L	87.3	12/48	90.7	6/12	82.0	2/12	93.0	2/12	83.5	2/12
T	98.0	24/48	n/a	0/12	97.6	8/12	97.4	9/12	98.9	7/12

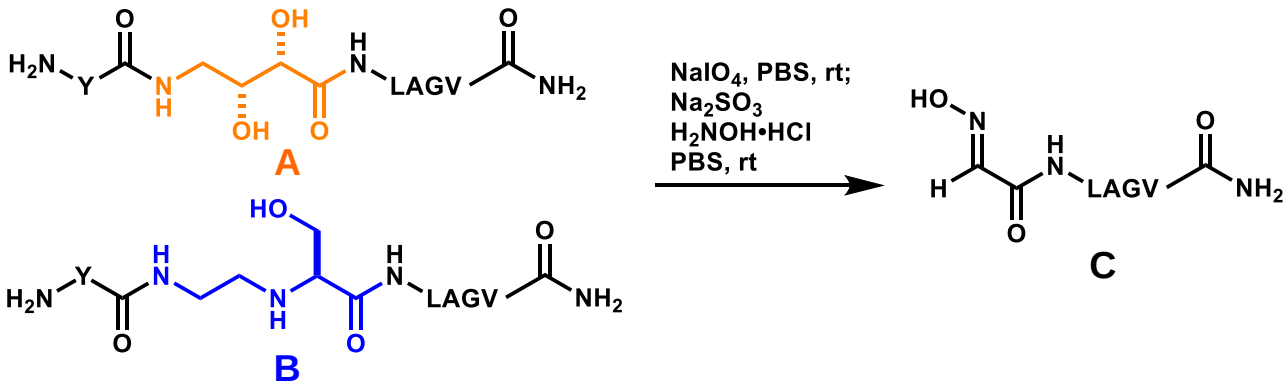


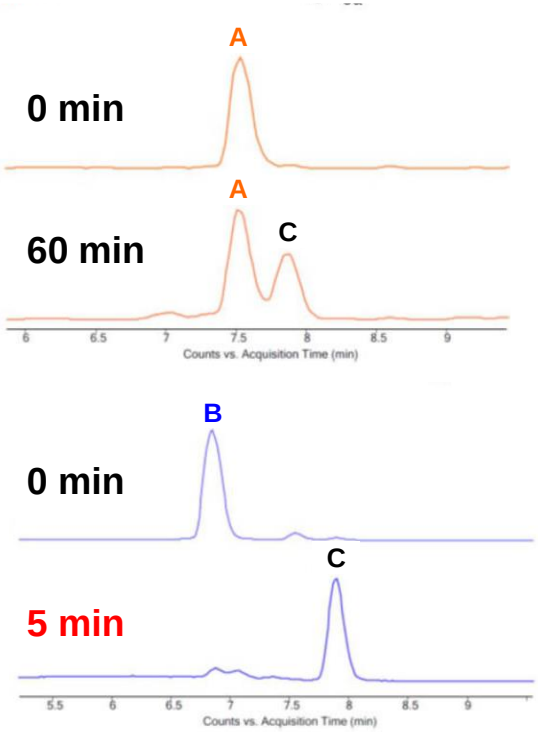
e.g. local confidence score in PEAKS

average local confidence (ALC): 91%

Releasing Tag by Oxidative Cleavage





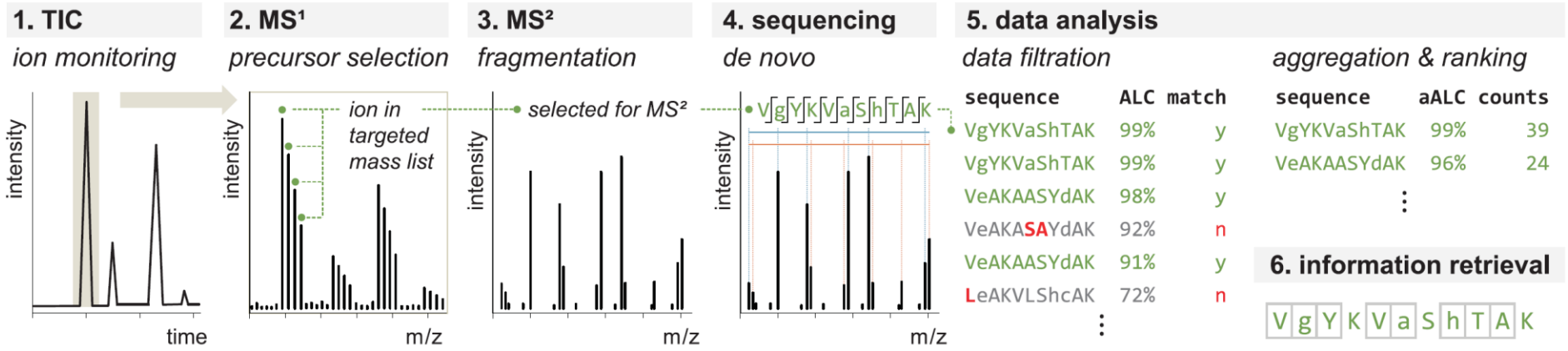


Seramox linker realized rapid and selective cleavage

1) Pomplun, S.; Shugrue, C. R.; Schmitt, A. M.; Schissel, C. K.; Farquhar, C. E.; Pentelute, B. L. *Angew. Chem. Int. Ed.* **2020**, *59*, 11566. 2) Methylene amine moiety of Seramox is recognized as a non-hydrolyzable bioisostere of amide bond. See also: Patani, G. A.; LaVoie, E. J. *Chem. Rev.* **1996**, *96*, 3147.

MS-Based Decoding

- **decoding**



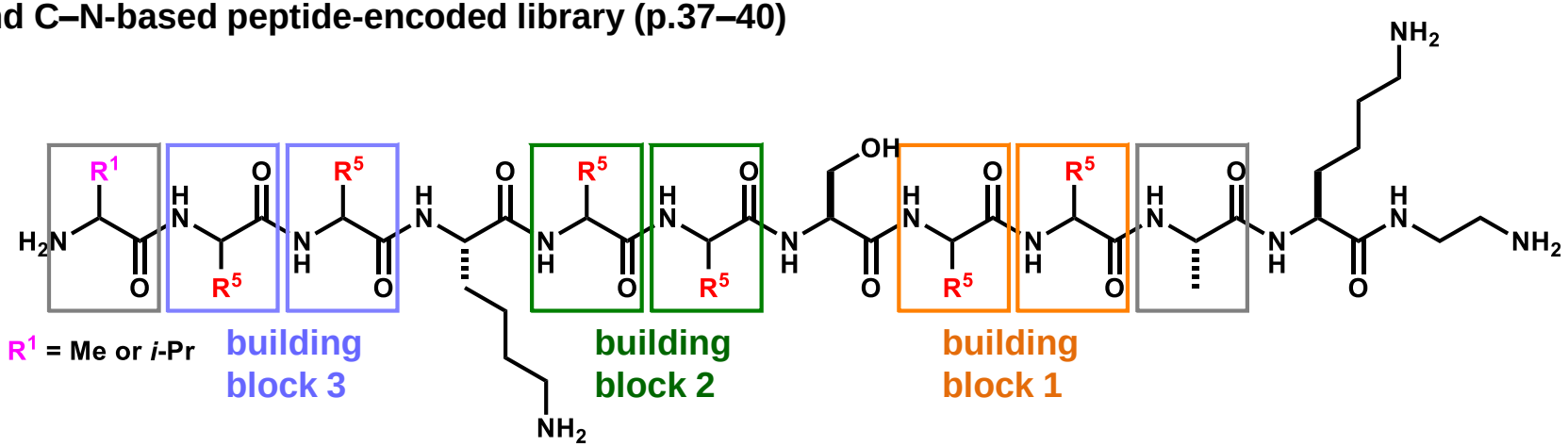
nLC-MS/MS: EASY-nLC 1200 + Orbitrap Fusion Eclipse Tribrid (ETD, EThcD)
de novo sequencing: PEAKS 8.5[1,2)]
detection limit: ~10 fmol/member

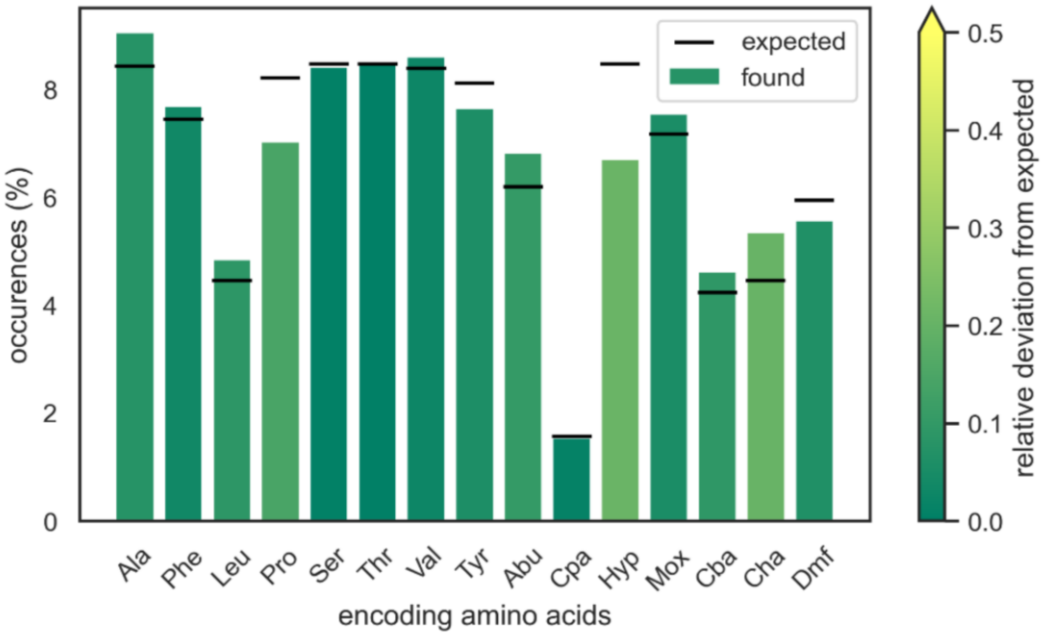
The structure of each library component could be decoded by MS-based sequencing at two-digit fmol scale

1) Vinogradov, A. A.; Gates, Z. P.; Zhang, C.; Quartararo, A. J.; Halloran, K. H.; Pentelute, B. L. *ACS Comb. Sci.* **2017**, *19*, 694. 2) Ma, B.; Zhang, K.; Hendrie, C.; Liang, C.; Li, M.; Doherty-Kirby, A.; Lajoie, G. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 2337.

Sequencing Bias Analysis

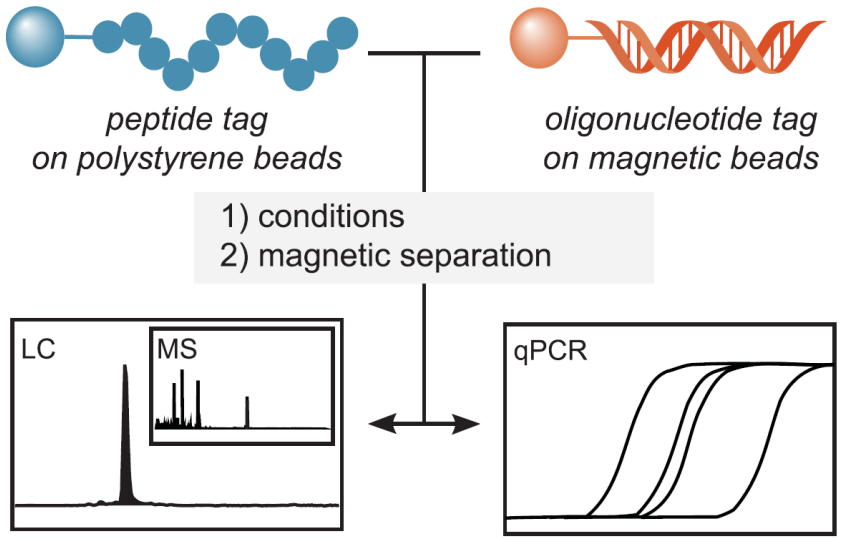
- analysis of peptides corresponding to the encoding tags of the 39k- and 41k-membered C–C- and C–N-based peptide-encoded library (p.37–40)





Sequencing bias was negligible

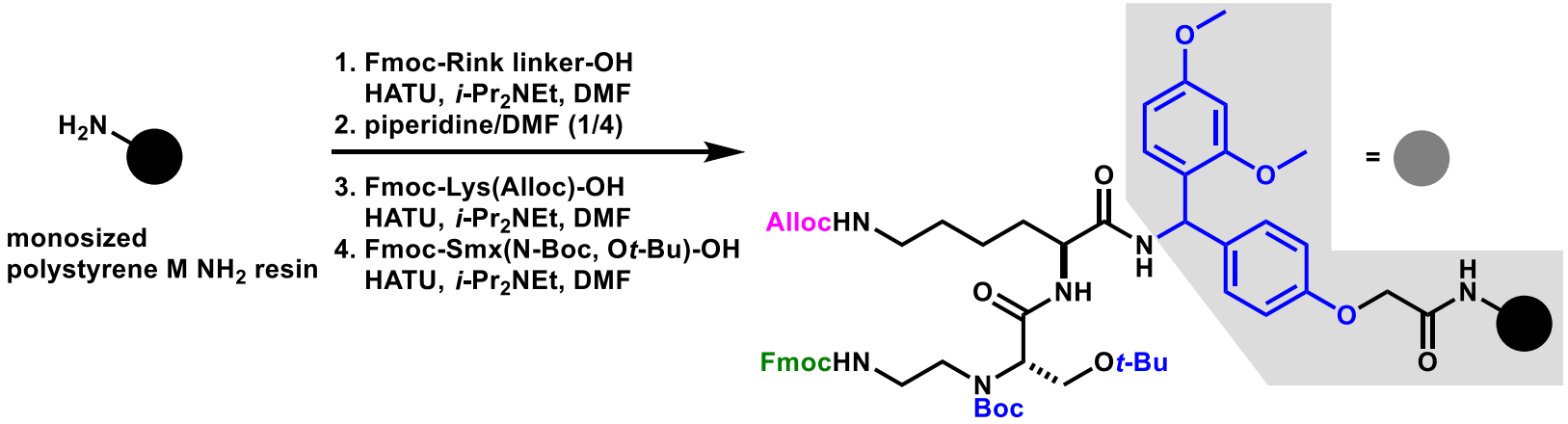
Stability of Peptide Tags

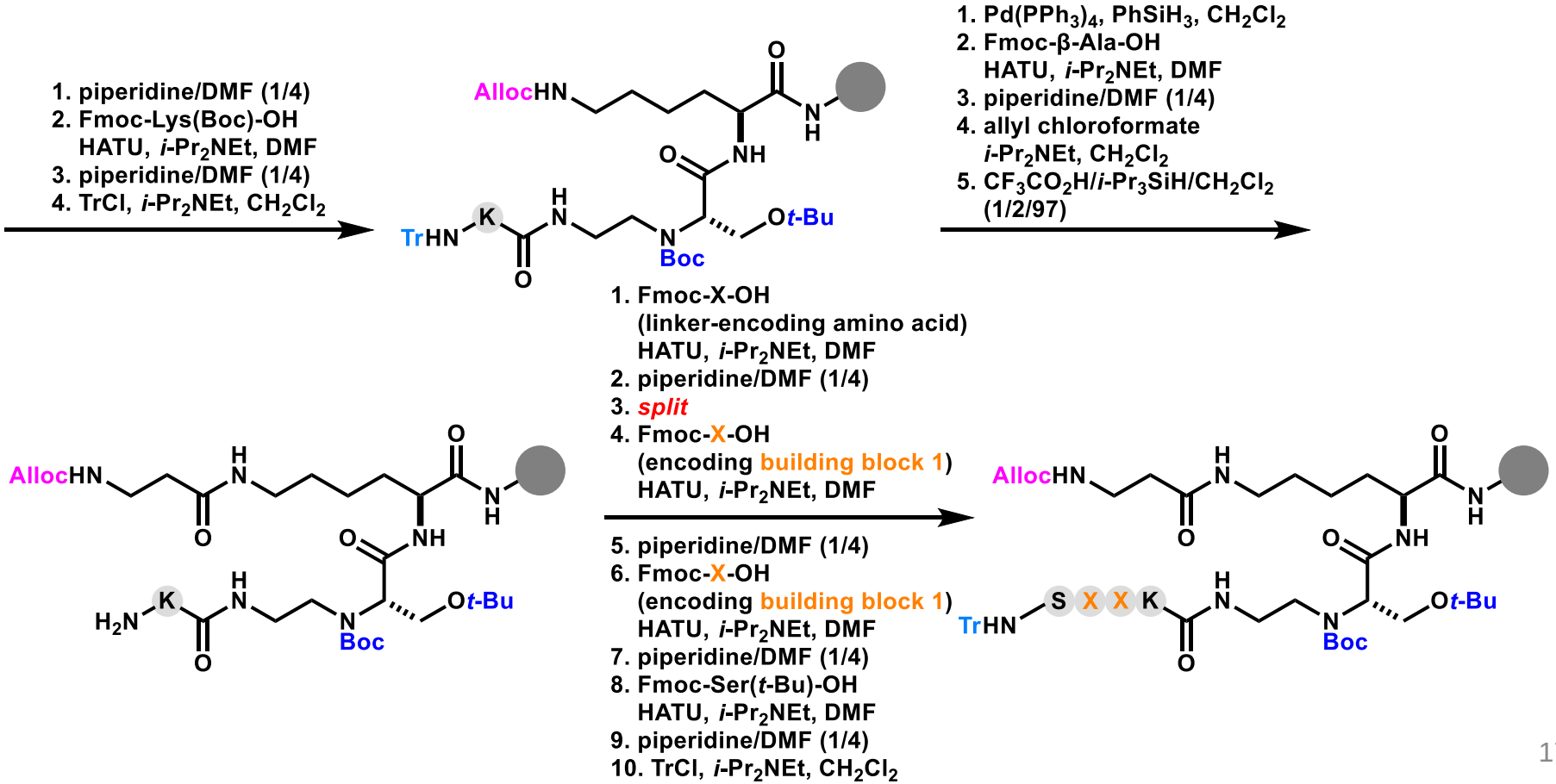


conditions	peptide	DNA
5% $Pd(OAc)_2$ in THF	>95%	<1%
5% CF_3CO_2H in H_2O	>95%	22%
5% H_2O_2 in H_2O	>95%	>95%
0.1 M NaOH in H_2O	>95%	22%
5% $PhI(OAc)_2$ in MeCN	>95%	74%
5% $[Cu(MeCN)_4]PF_6$ in MeCN	>95%	88%
1% Grubbs G2 in CH_2Cl_2	>95%	23%
20% piperidine in DMF	>95%	17%
5% Et_3N in DMF	>95%	23%
1 M NaO*t*-Bu in THF	<1%	16%
UV (254 nm, 4 h) in H_2O	>95%	83%

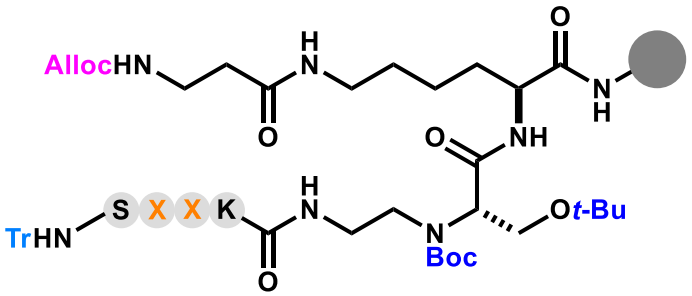
Compared to DNA tags, wider range of reaction conditions and solvents are acceptable to peptide tags

Synthesis of PEL (1)



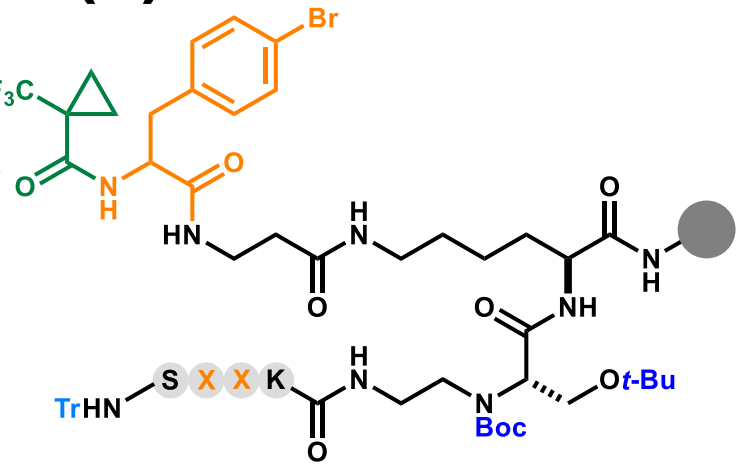


Synthesis of PEL (2)



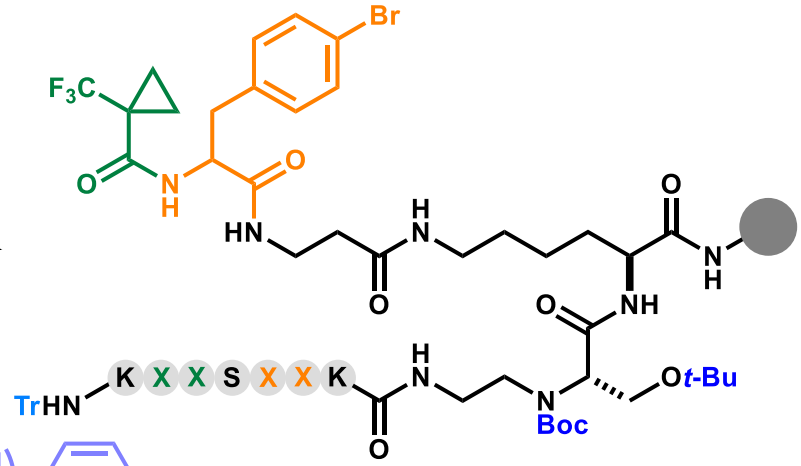
1. $Pd(PPh_3)_4$, $PhSiH_3$, CH_2Cl_2
2. bulding block 1
   HATU, *i*-Pr_2NEt, DMF
3. *pool*

4. piperidine/DMF (1/4)
5. *split*
6. building block 2
   HATU, *i*-Pr_2NEt, DMF
7. CF_3CO_2H/*i*-Pr_3SiH/CH_2Cl_2
   (1/2/97)



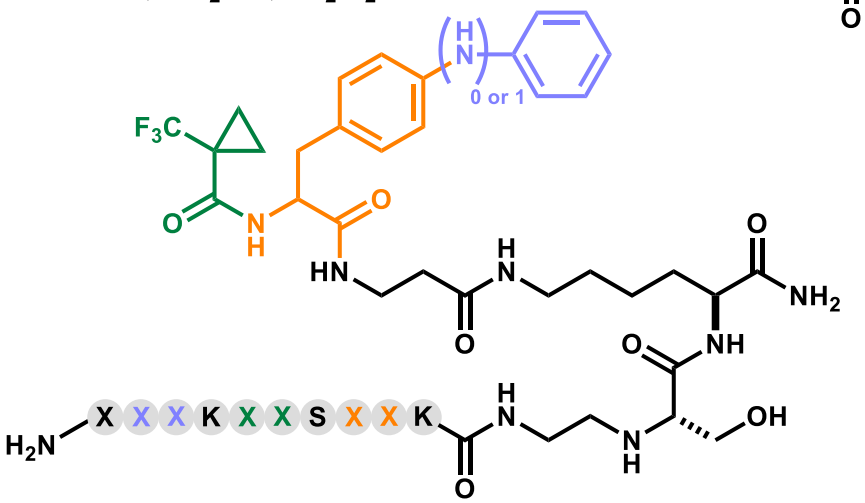
1. Fmoc-X-OH
   (encoding building block 2)
   HATU, *i*-Pr_2NEt, DMF
2. piperidine/DMF (1/4)
3. Fmoc-X-OH
   (encoding building block 2)
   HATU, *i*-Pr_2NEt, DMF
4. *pool*

5. piperidine/DMF (1/4)
6. Fmoc-Lys(Boc)-OH
   HATU, *i*-Pr_2NEt, DMF
7. piperidine/DMF (1/4)
8. TrCl, *i*-Pr_2NEt, CH_2Cl_2



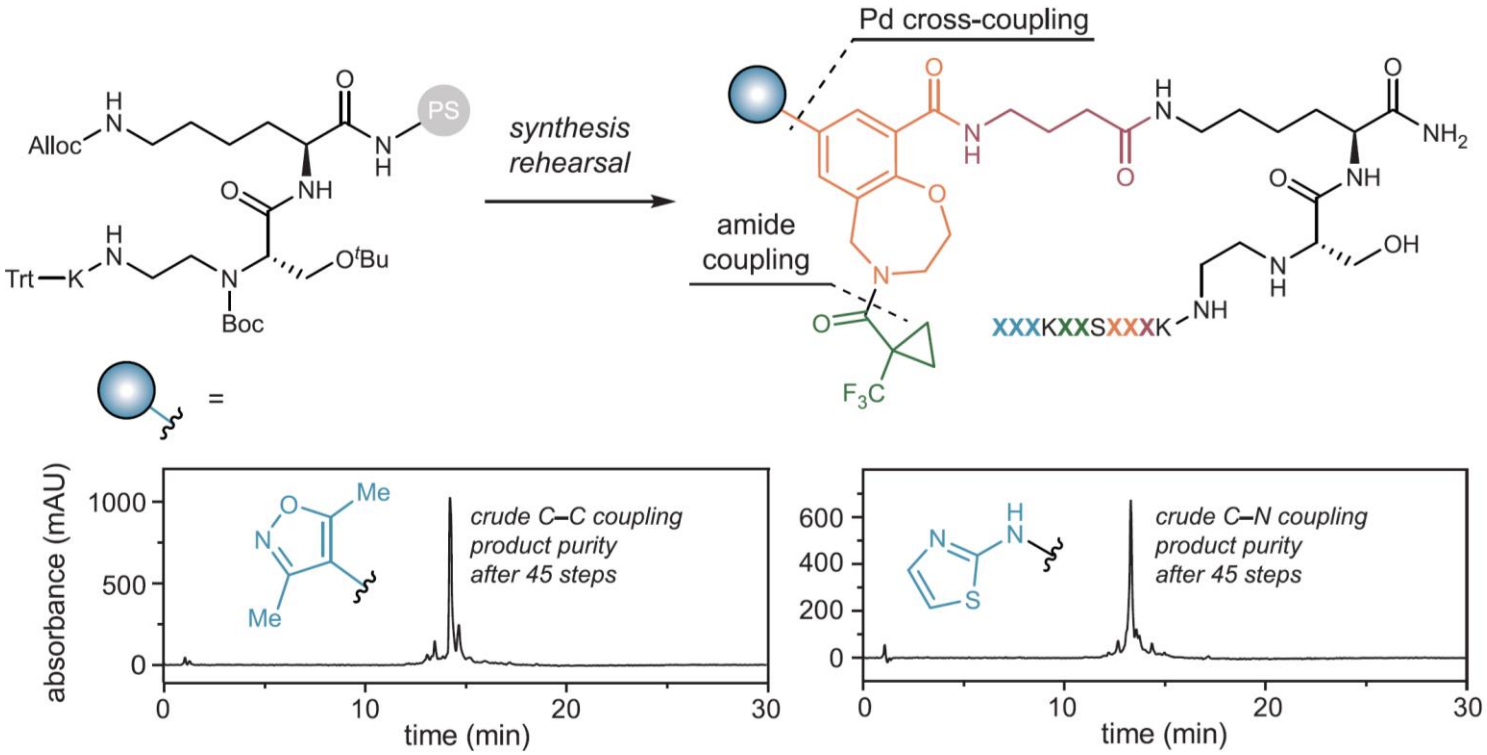
1. *split*
2. Pd-mediated C-C or C-N cross coupling
   (For conditions, see p.19)
3. CF_3CO_2H/*i*-Pr_3SiH/CH_2Cl_2 (1/2/97)
4. Fmoc-X-OH (encoding building block 3)
   HATU, *i*-Pr_2NEt, DMF

5. piperidine/DMF (1/4)
6. Fmoc-X-OH (encoding building block 3)
   HATU, *i*-Pr_2NEt, DMF
7. *pool*
8. piperidine/DMF (1/4)
9. Fmoc-X-OH (encoding library)
   HATU, *i*-Pr_2NEt, DMF
10. piperidine/DMF (1/4)
11. CF_3CO_2H/$HS(CH_2)_2SH$/H_2O/*i*-Pr_3SiH
    (94/2.5/2.5/1)



The peptide and small molecule were sequentially functionalized using orthogonal protecting groups (Fmoc, Tr and Alloc)

Evaluation of Coupling Efficiency



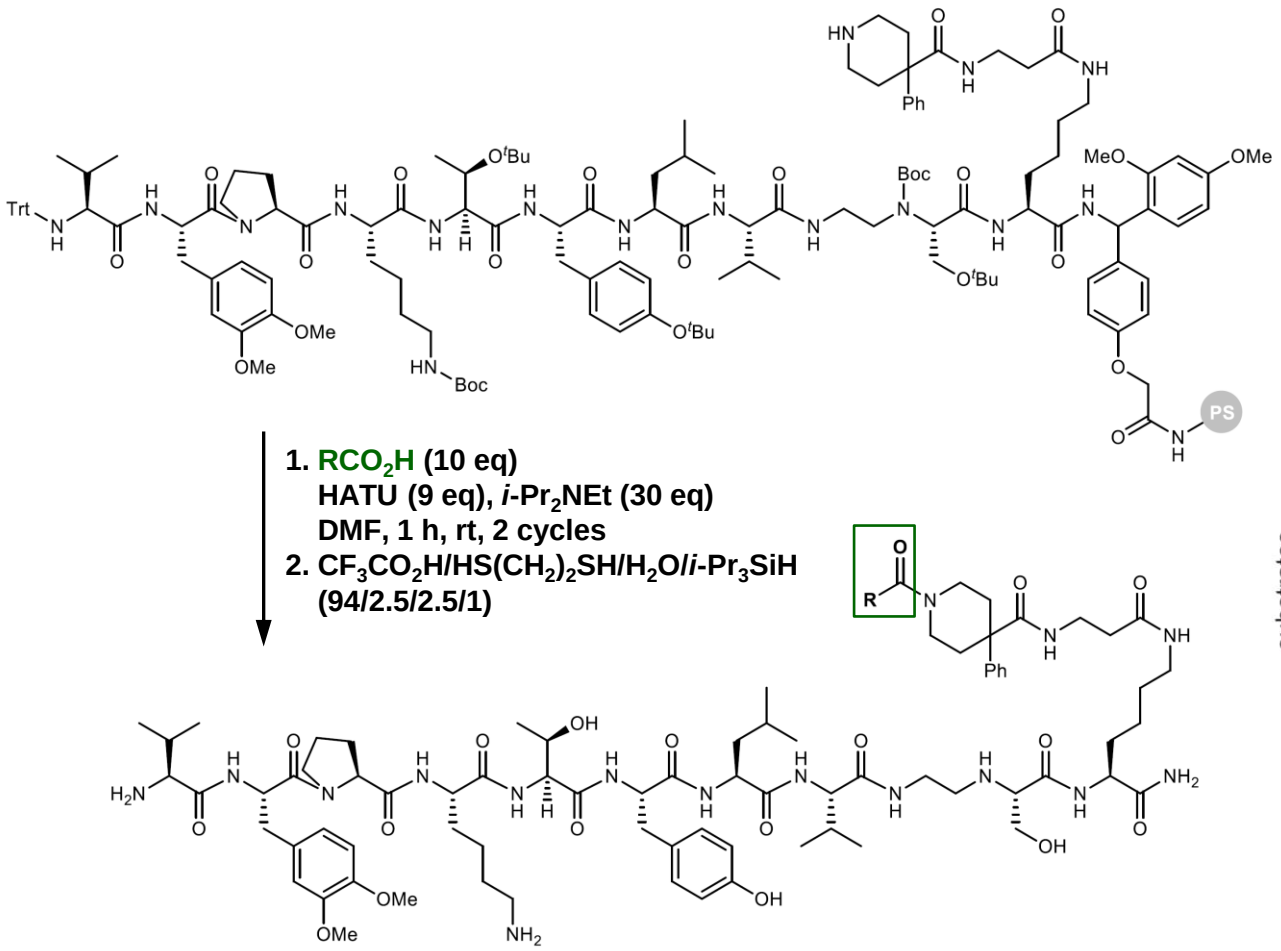
amide coupling conditions
RCO_2H (10 eq)
HATU (9 eq), *i*-Pr_2NEt (30 eq)
DMF, 1 h, rt, 2 cycles

C-C coupling conditions
$Ar\text{-}B(OH)_2$ (5 eq)
G4 Pd XPhos (1.2 eq), XPhos (1.2 eq)
K_3PO_4 (5 eq), THF/H_2O, rt, 24 h

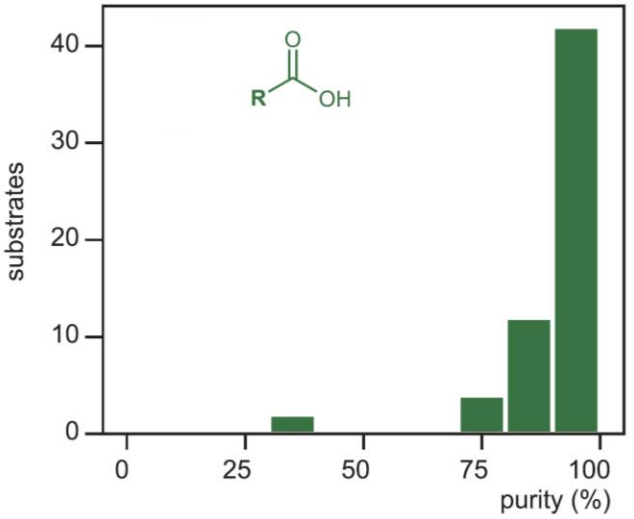
C-N coupling conditions
$Ar\text{-}NH_2$ (3 eq)
[(AlPhos)Pd]$_2$(cod) (0.6 eq)
DBU (5 eq), 2-MeTHF, 50 °C, N_2, 2 h

In the synthesis rehearsal, high purities of the compounds were confirmed by HPLC analysis

Efficiency of Amide Coupling

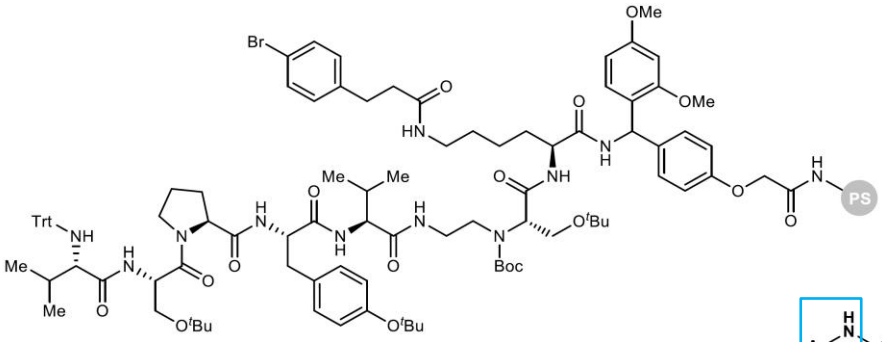


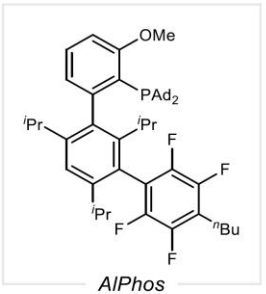
1. RCO_2H (10 eq)
   HATU (9 eq), *i*-Pr_2NEt (30 eq)
   DMF, 1 h, rt, 2 cycles
2. $CF_3CO_2H/HS(CH_2)_2SH/H_2O/i\text{-}Pr_3SiH$
   (94/2.5/2.5/1)



LC-MS spectra were automatically integrated to afford peak area percent as purity (For full list of RCO_2H, see p.28 and 29 in Appendix)

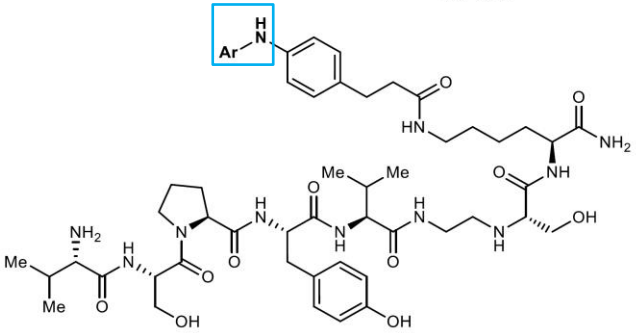
Efficiency of C-N Cross Coupling

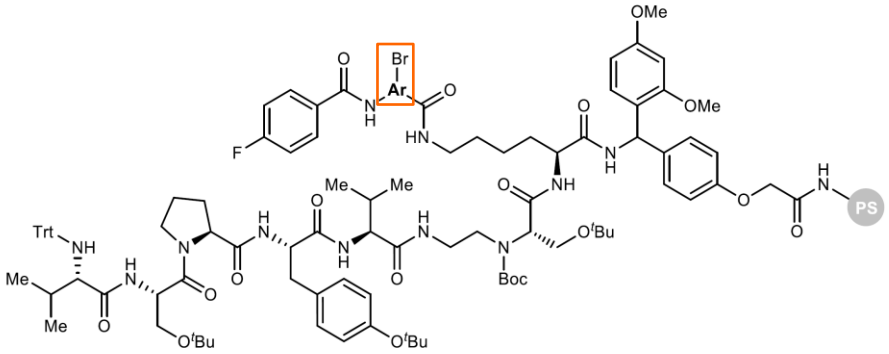


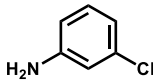


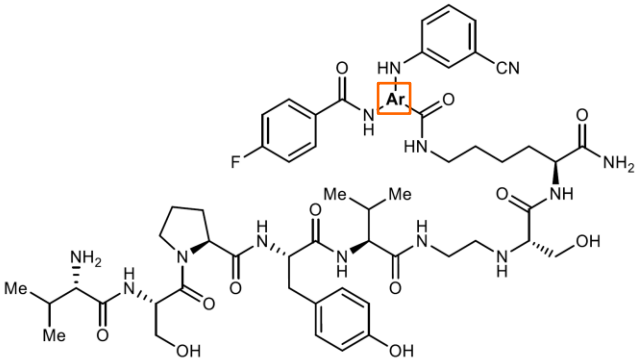
Ad = adamantyl

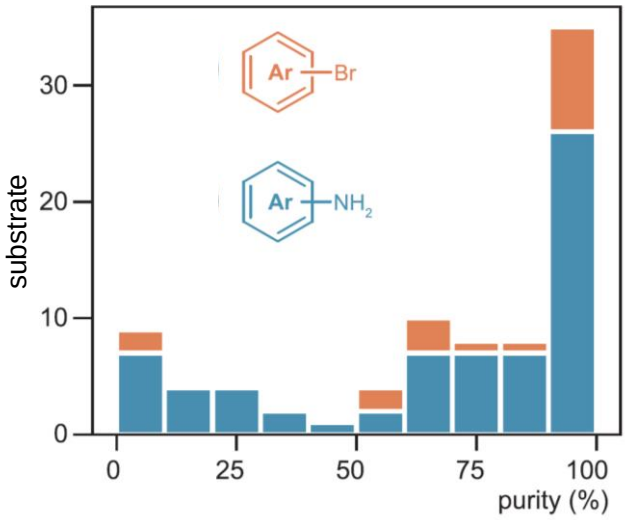
1. Ar-NH_2 (3 eq)
   $[(AlPhos)Pd]_2(cod)$ (0.6 eq)
   DBU (5 eq), 2-MeTHF, 50 °C, N_2, 2 h
2. $CF_3CO_2H/HS(CH_2)_2SH/H_2O/i\text{-}Pr_3SiH$ (94/2.5/2.5/1)





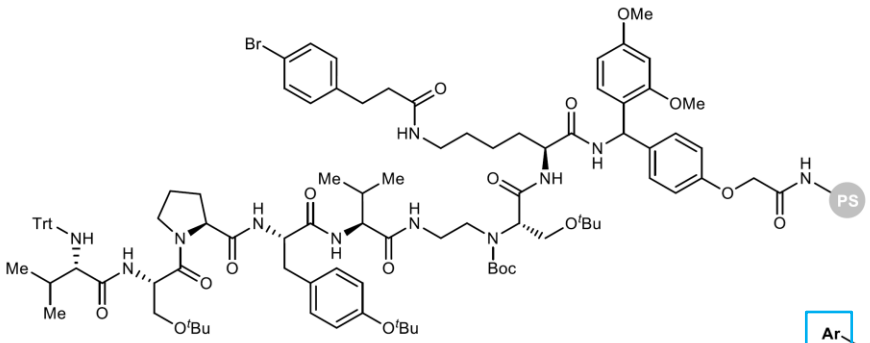
1.  (3 eq)
   $[(AlPhos)Pd]_2(cod)$ (0.6 eq)
   DBU (5 eq), 2-MeTHF, 50 °C, N_2, 2 h
2. $CF_3CO_2H/HS(CH_2)_2SH/H_2O/i\text{-}Pr_3SiH$ (94/2.5/2.5/1)

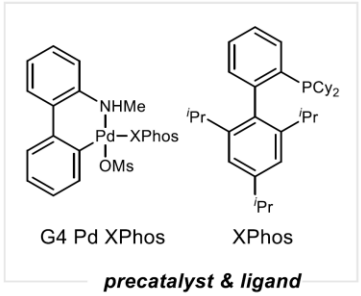




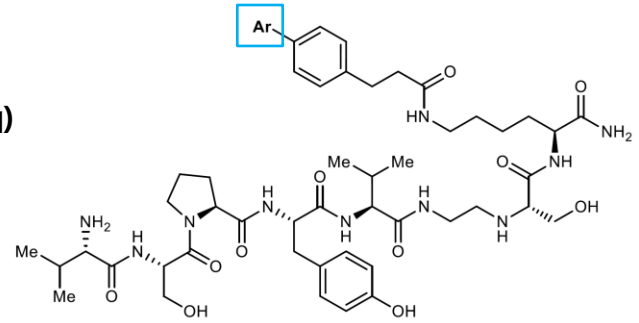
LC-MS spectra were automatically integrated to afford peak area percent as purity (For full list of Ar-NH_2 and Ar-Br, see p.30–33 in Appendix)

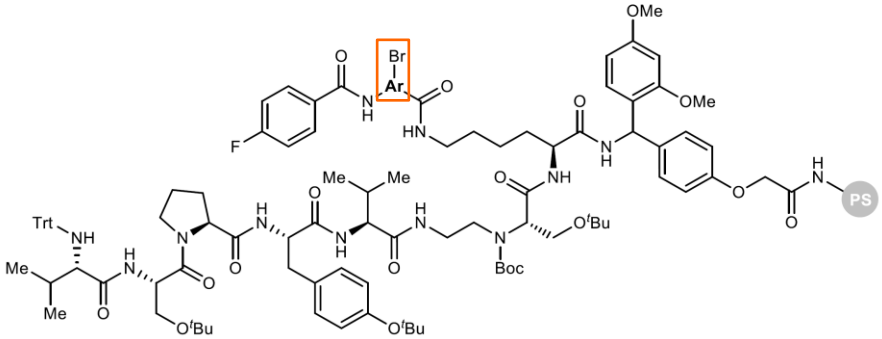
Efficiency of C-C Cross Coupling

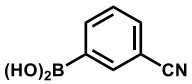


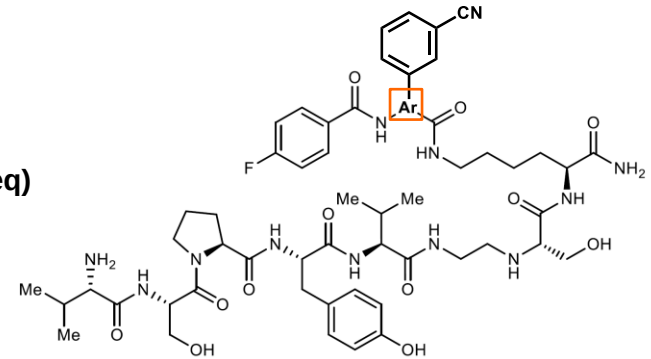


1. Ar-B(OH)$_2$ (5 eq)
   G4 Pd XPhos (1.2 eq), XPhos (1.2 eq)
   K_3PO_4 (5 eq), THF/H_2O, rt, 24 h
2. CF_3CO_2H/$HS(CH_2)_2SH$/H_2O/*i*-Pr_3SiH
   (94/2.5/2.5/1)

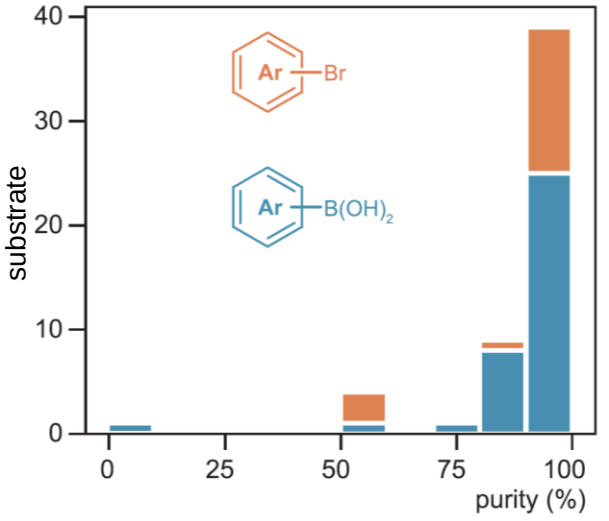




1.  (5 eq)
   G4 Pd XPhos (1.2 eq), XPhos (1.2 eq)
   K_3PO_4 (5 eq), THF/H_2O, rt, 24 h
2. CF_3CO_2H/$HS(CH_2)_2SH$/H_2O/*i*-Pr_3SiH
   (94/2.5/2.5/1)



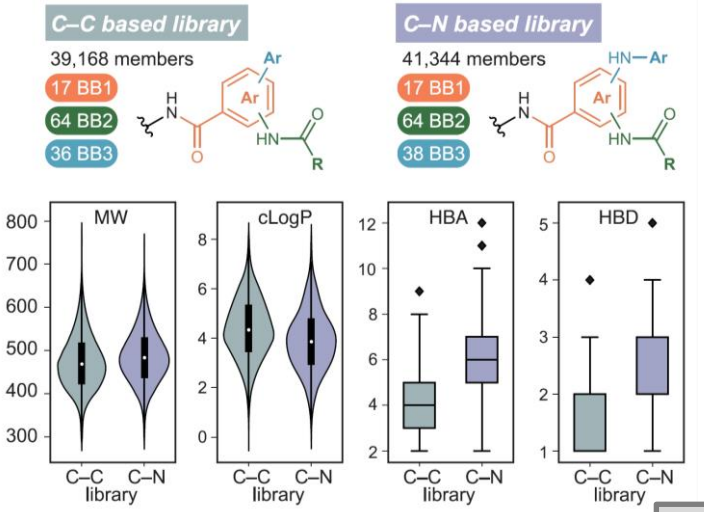
C–C coupling conditions
boronic acid (5 equiv)
XPhos Pd G4 (1.2 equiv), XPhos (1.2 equiv)
aq. K_2PO_4 (5 equiv), THF, rt, 24 h



LC-MS spectra were automatically integrated to afford peak area percent as purity (For full list of Ar-B(OH)$_2$ and Ar-Br, see p.34–36 in Appendix)

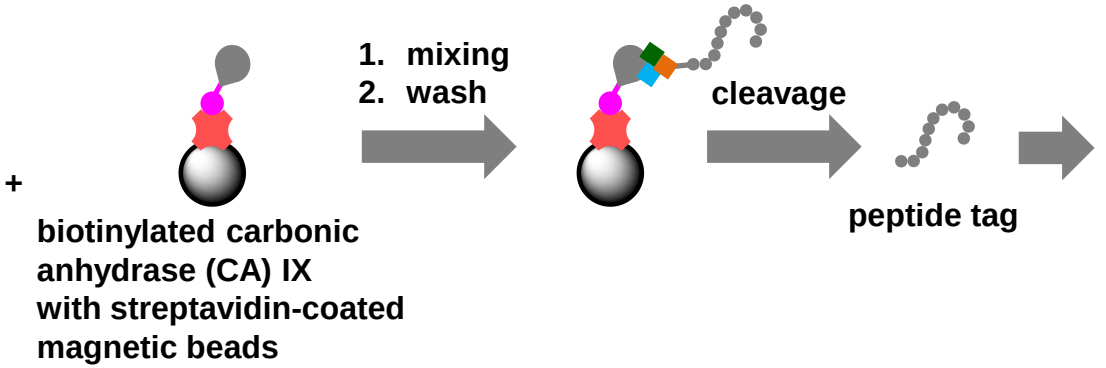
Proof of Concept: Selection of CA IX Binders

For library members, see p.37–40



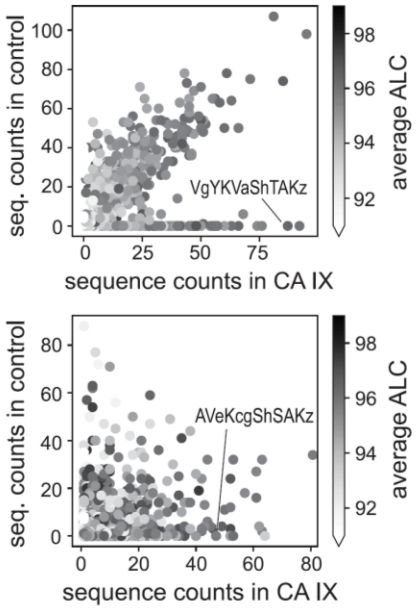
c.f. Lipinski's rule of five
MW ≤ 500, LogP ≤ 5, HBA ≤ 10, HBD ≤ 5

1. affinity selection (on automated Kingfisher Duo Prime Purification system)



Considering drug likeness, molecular weight > 650 Da or clogP > 5.0 were excluded from nomination

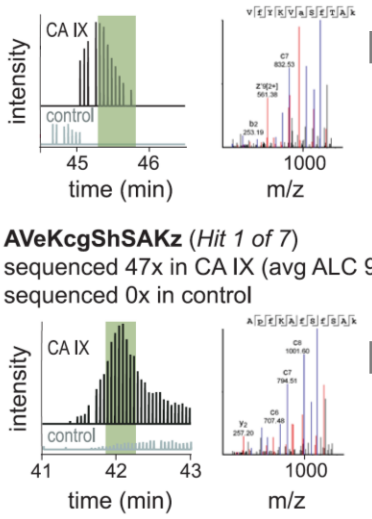
2. selection analysis



3. hit nomination

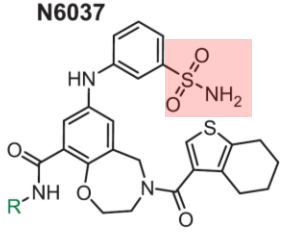
VgYKVaShTAKz (*Hit 1 of 8*)
sequenced 87x in CA IX (avg ALC 97.3)
sequenced 0x in control

AVeKcgShSAKz (*Hit 1 of 7*)
sequenced 47x in CA IX (avg ALC 96.6)
sequenced 0x in control

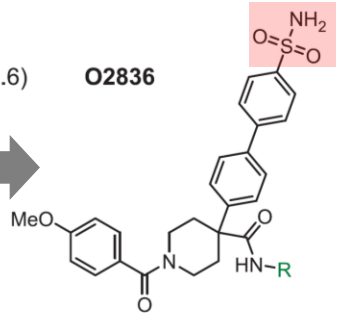


4. decoding

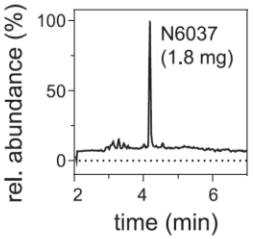
N6037



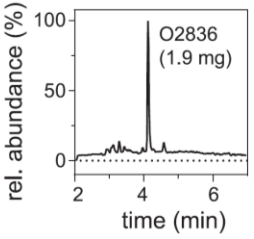
O2836



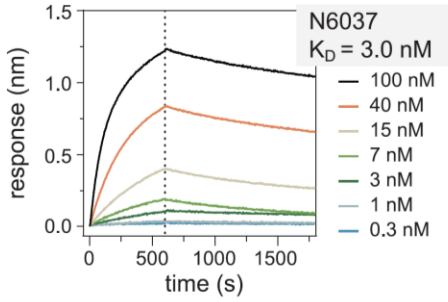
5. resynthesis

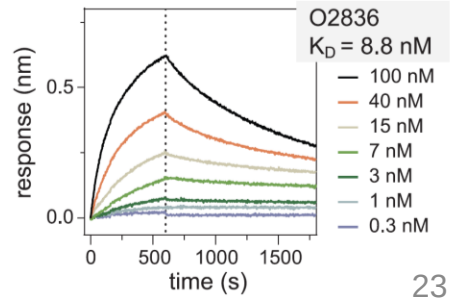


SPS enables hit synthesis in ~2 days



6. affinity measurement (BLI)

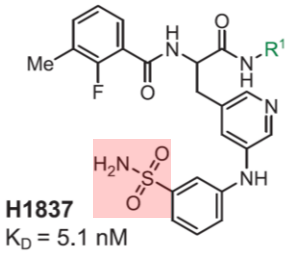




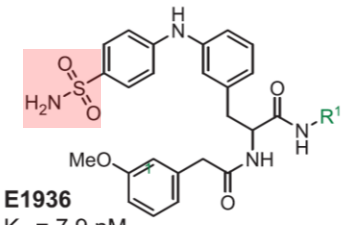
Hit Molecules from Library

C-N based library

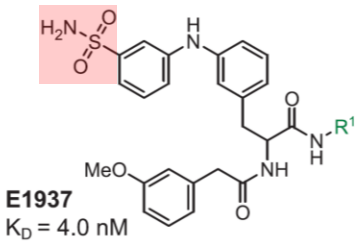
N6037 (see page 24), K_D = 3.0 nM



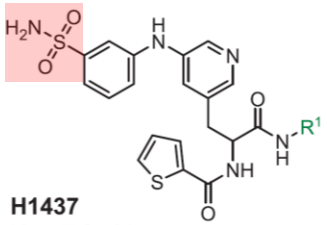
H1837
K_D = 5.1 nM



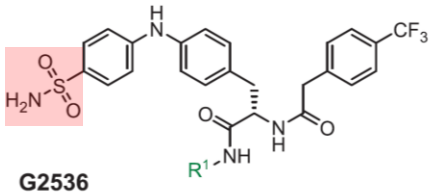
E1936
K_D = 7.9 nM



E1937
K_D = 4.0 nM



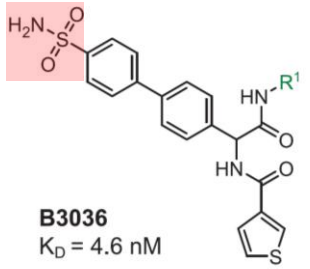
H1437
K_D = 7.8 nM



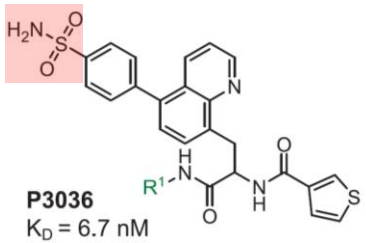
G2536
K_D = 8.1 nM

C-C based library

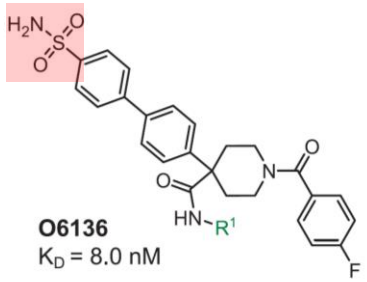
O2836 (see page 24), K_D = 8.8 nM



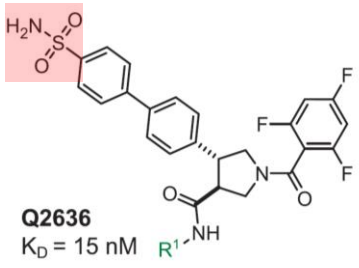
B3036
K_D = 4.6 nM



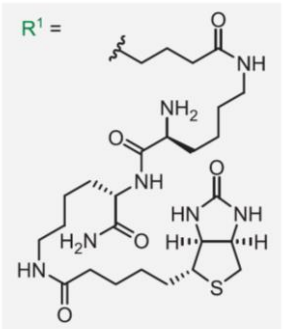
P3036
K_D = 6.7 nM



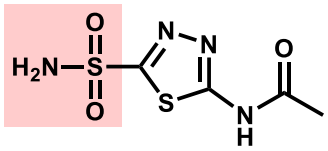
O6136
K_D = 8.0 nM



Q2636
K_D = 15 nM



sulfonamide moiety: binding motif for Zn^{2+} center of carbonic anhydrase



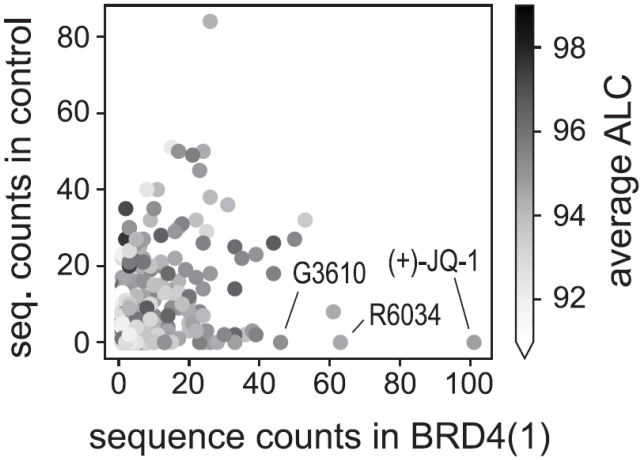
c.f. acetazolamide (inhibitor of CA IX)

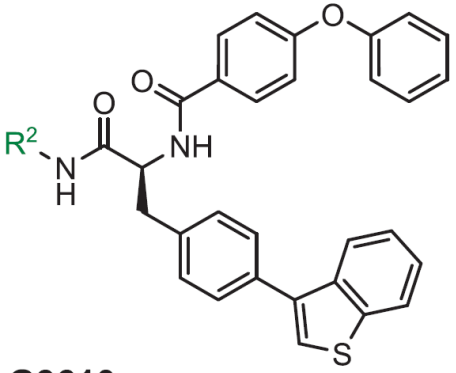
for conjugating compound to streptavidin probes of biolayer interferometry (BLI)

- **Hit molecules were found to exhibit affinities to CA IX in the 10^8–10^9 M range (All the compounds possessed sulfonamide moiety as expected)**
- **Encoding peptides were separately synthesized and confirmed to have no contribution to the binding**

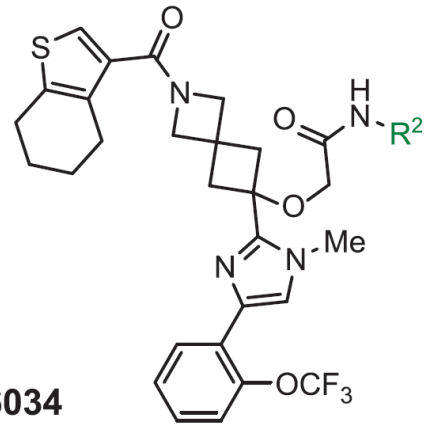
De novo Discovery of Binders for BRD4(1)

- **bromodomain 1 of BRD4 [BRD4(1)]**



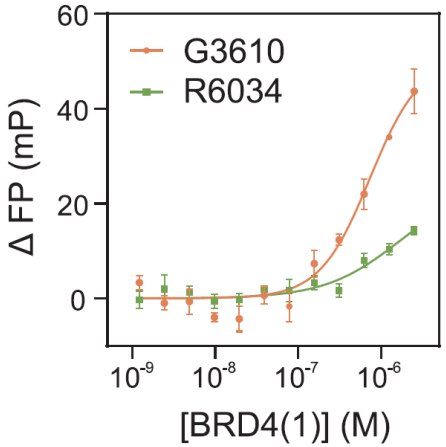


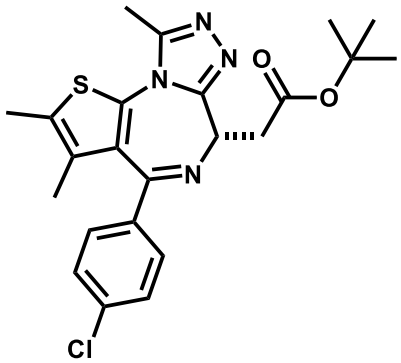
G3610
K_D = 0.7 μM



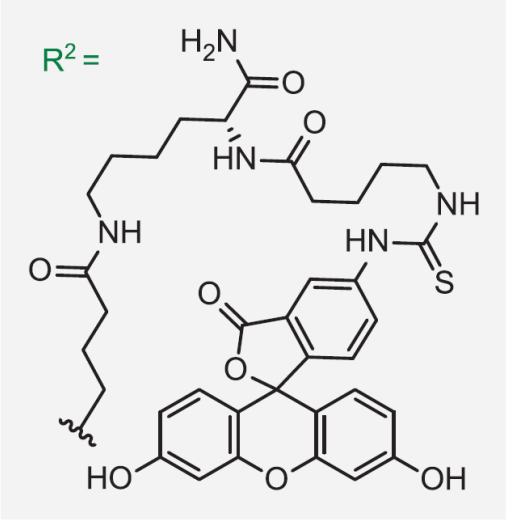
R6034
K_D = 1.7 μM

- **fluorescence polarization assay**





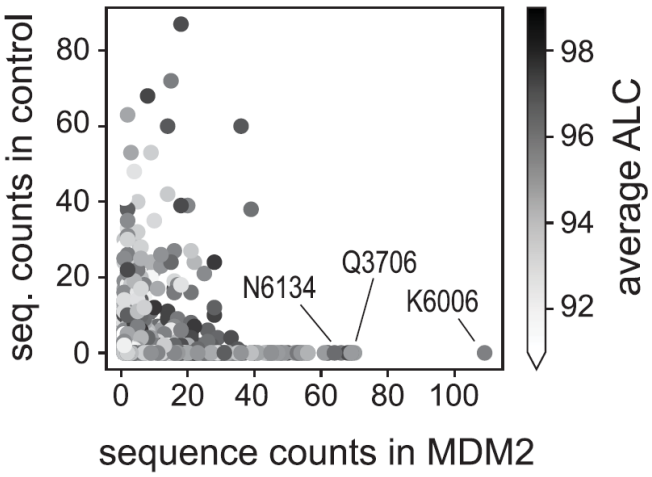
(+)-JQ-1(BRD4 inhibitor)
K_D = 0.52 μM (BLI)

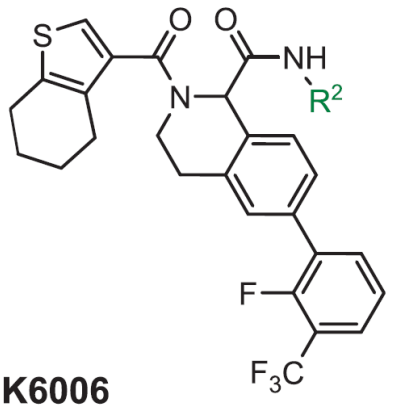


It was demonstrated that affinity selection realized de novo discovery of small molecules with approx. 1-digit micromolar affinity for proteins involved with oncogenic pathways

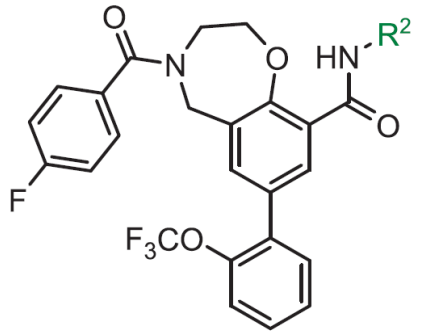
De novo Discovery of Binders for MDM2

- **MDM2**



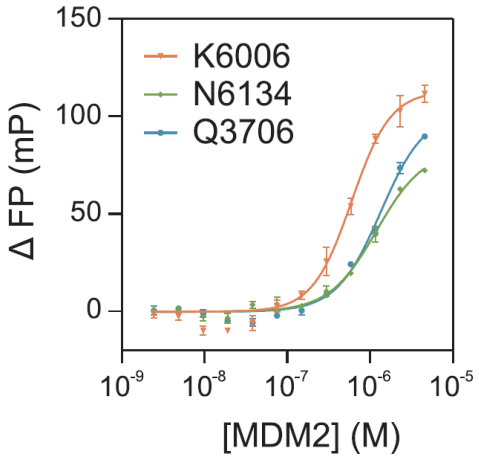


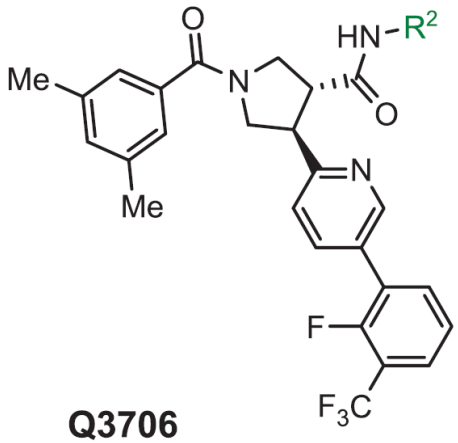
K6006
K_D = 0.6 μM



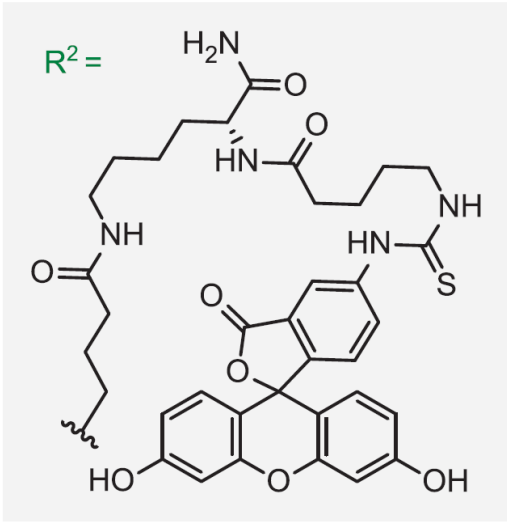
N6134
K_D = 1.3 μM

- **fluorescence polarization assay**





Q3706
K_D = 1.5 μM



It was demonstrated that affinity selection realized de novo discovery of small molecules with approx. 1-digit micromolar affinity for proteins involved with oncogenic pathways

Summary and Outlook

- **robustness of tags**

  high chemical stability, low sequencing bias, defined charge states

- **reaction efficiency of both tags and small molecules**
  (a large excess amount of reagent is acceptable, filtration as a simple workup /purification procedure owing to the solid-phase system)

- **range of application**

  The peptide tag is applicable to DNA binding proteins

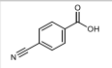
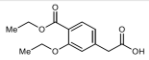
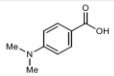
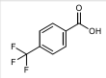
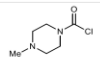
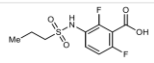
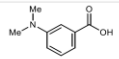
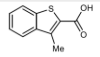
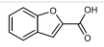
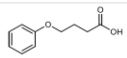
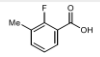
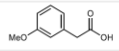
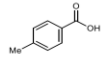
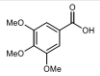
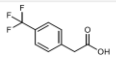
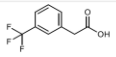
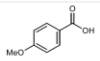
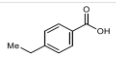
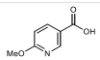
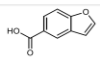
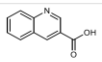
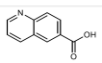
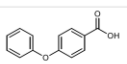
- **detection limit/library size**
  (peptide tags cannot be amplified)

- **orthogonality of the reactions**
  Fmoc, Tr, and Alloc groups are utilized for peptide tags

Improvement of mass spectrometric instrumentation and methods is a key factor for further development
(which could also expand the possibility of “self-encoded” library)

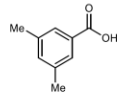
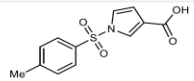
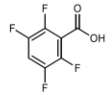
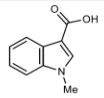
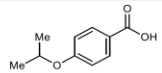
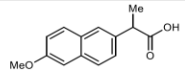
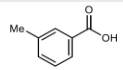
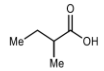
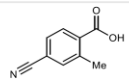
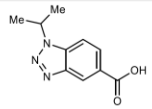
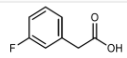
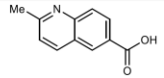
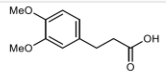
Appendix

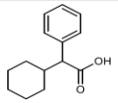
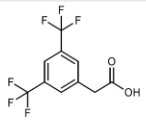
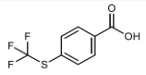
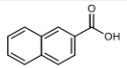
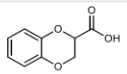
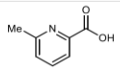
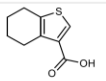
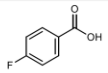
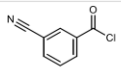
Carboxylic Acids (1)

Carboxylic acid	Structure	Purity (%)	Mass calculated $[M+Na]^+$ (Da)	Mass found (Da)
CA01		94	1692.9124	1692.8227
CA02		89	1797.9802	1797.8882
CA03		85	1710.9593	1710.8667
CA04		88	1735.9045	1735.8117
CA05		85 [a]	1689.9702	1689.8792
CA06		38	1824.9180	1824.8212
CA07		90	1710.9593	1710.8687
CA08		91	1685.9077	1685.8487
CA09		79 [a]	1690.9906	1690.9001
CA10		91	1737.9048	1737.8173
CA11		93	1707.9120	1707.8243
CA12		91	1685.9077	1685.8159
CA13		91	1673.9641	1673.8750
CA14		90 [a]	1673.8735	1673.7834
CA15		87	1657.8964	1657.8088
CA16		78	1699.9045	1699.8148
CA17		84	1725.9590	1725.8695
CA18		91	1699.9234	1699.9064
CA19		94	1711.9433	1711.8520
CA20		92	1681.9328	1681.8384
CA21		n/a [b]	1757.9488	n/a
CA22		93	1709.9641	1709.8738
CA23		88	1701.9202	1701.8298
CA24		75	1669.9076	1669.8220
CA25		79	1749.9202	1749.8288
CA26		93	1721.8889	1721.7968
CA27		92	1749.9202	1749.8266
CA28		91	1697.9277	1697.8393
CA29		90	1695.9484	1695.8585
CA30		93	1673.8735	1673.7840
CA31		95	1698.9229	1698.8342
CA32		92	1695.9484	1695.8604
CA33		94	1707.9119	1707.8239
CA34		93	1718.9280	1718.8402
CA35		88	1718.9280	1718.8401
CA36		91	1759.9433	1759.8520

Appendix

Carboxylic Acids (2)

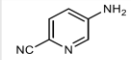
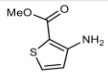
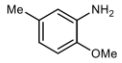
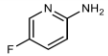
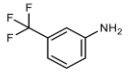
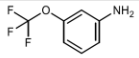
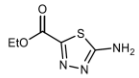
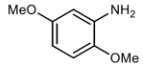
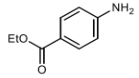
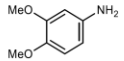
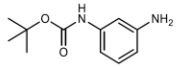
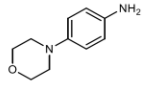
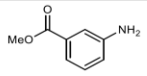
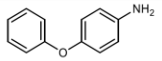
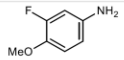
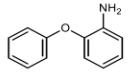
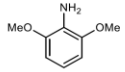
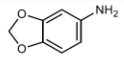
CA37		92	1695.9484	1695.8616
CA38		91	1810.9212	1810.8306
CA39		91	1739.8794	1739.7884
CA40		90	1661.9277	1661.8389
CA41		92	1720.9437	1721.0080
CA42		93	1725.9590	1725.8699
CA43		93	1775.9746	1775.8836
CA44		n/a [b]	1681.9328	n/a
CA45		92	1647.9484	1647.8622
CA46		91	1706.9280	1706.8367
CA47		91	1750.9655	1750.8744
CA48		91	1699.9234	1699.9021
CA49		84	1732.9437	1732.8494
CA50		87	1755.9696	1755.8781

CA51		87	1645.9328	1645.8423
CA52		92	1703.8983	1703.8105
CA53		94	1764.0110	1763.9822
CA54		39	1817.9075	1817.8203
CA55		93	1767.8766	1767.7869
CA56		92	1717.9328	1717.8446
CA57		91	1725.9226	1725.8336
CA58		94	1668.9124	1668.8251
CA59		91	1682.9280	1682.8386
CA60		94	1727.9205	1727.8331
CA61		92	1685.9077	1685.9563
CA62		87 [a]	1692.9124	1692.8264

a) Reaction conducted without HATU. b) not evaluated.

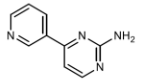
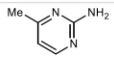
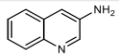
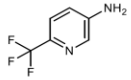
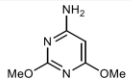
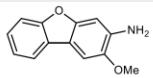
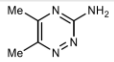
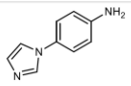
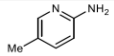
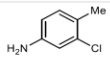
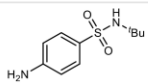
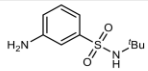
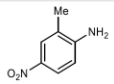
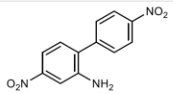
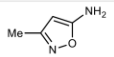
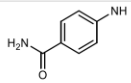
Appendix

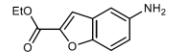
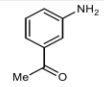
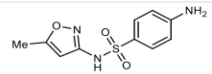
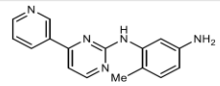
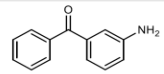
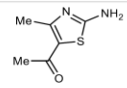
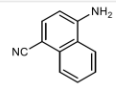
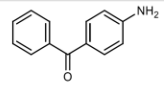
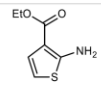
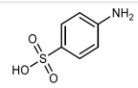
Ar-NH_2 for C-N Cross Coupling (1)

Aniline	Structure	Purity (%)	Mass calculated $[M+Na]^+$ (Da)	Mass found (Da)
A01		92	1067.5648	1067.555
A02		100	1092.5601	1092.5489
A03		73	1130.5315	1130.5186
A04		100	1066.5696	1066.5595
A05		100	1110.5958	1110.5839
A06		100	1073.5212	1073.5125
A07		100	1085.5554	1085.5449
A08		100	1134.557	1134.5428
A09		100	1150.5519	1150.5372
A10		88	1108.5662	1108.552
A11		81	1146.5376	1146.5218
A12		100	1126.5907	1126.5775
A13		100	1138.5907	1138.5762
A14		76	1068.5601	1068.5493
A15		100	1126.5907	1126.5775
A16		100	1097.5754	1097.5613
A17		100 [a]	1081.5805	1081.5678
A18		100	1151.6223	1151.6048
A19		100	1124.5751	1124.5598
A20		100	1158.5958	1158.5787
A21		100	1114.5707	1114.5563
A22		100	1084.5602	1084.5486
A23		100	1158.5958	1158.5781
A24		100	1126.5907	1126.5744
A25		89	1110.5594	1110.5560

Appendix

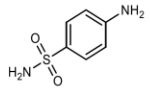
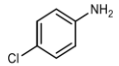
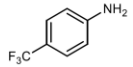
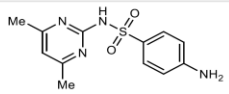
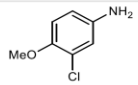
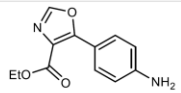
Ar-NH_2 for C-N Cross Coupling (2)

A26		78	1145.5866	1145.5846
A27		75	1082.5757	1082.5736
A28		91	1117.5805	1117.5810
A29		90	1135.5522	1135.5520
A30		80	1128.5812	1128.5825
A31		93	1186.5907	1186.5914
A32		90	1097.5866	1097.5881
A33		70	1132.5914	1132.5912
A34		89	1081.5805	1081.5819
A35		100	1114.5463	1114.5472
A36 [b]		88	1693.8746 [c]	1693.8770
A37 [b]		88	1693.8746 [c]	1693.8771
XA01		21	1125.5703	1125.5553
XA02		0	1232.571	Not found
XA03		0	1071.5597	Not found
XA04		100 [d]	1109.5754	1109.5633

XA05		69	1091.5648	1091.5523
XA06		0	1178.5856	not found
XA07		0	1134.557	not found
XA08		100 [d]	1082.5645	1082.553
XA09		20	1108.5801	1108.5678
XA10		10	1226.5638	1226.5422
XA11		15	1084.555	1084.5437
XA12		60	1250.6445	1250.6402
XA13		66	1170.5958	1170.5971
XA14		38	1129.5475	1129.5487
XA15		0	1068.5601	not found
XA16		30	1141.5805	1141.5822
XA17		62	1170.5958	1170.5965
XA18		0	1144.5471	not found
XA19		49	1124.5445 [e]	1124.5434
XA20		74	1098.5165	1098.5177

Appendix

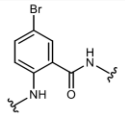
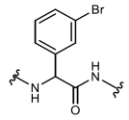
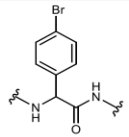
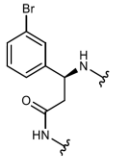
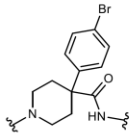
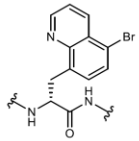
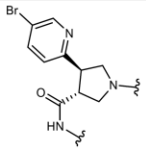
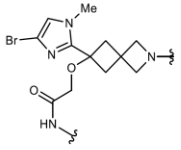
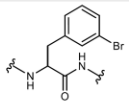
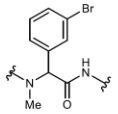
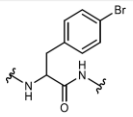
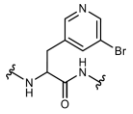
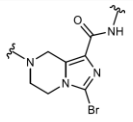
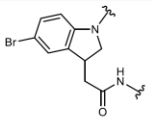
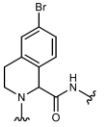
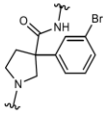
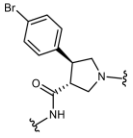
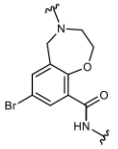
Ar-NH_2 for C-N Cross Coupling (3)

XA21		0	1069.5553	not found
XA22		50	1092.5601	1092.5605
XA23		57	1145.5424	1145.5438
XA24		13	1100.5306	1100.5313
XA25		20	1100.5306	1100.5308
XA26		62	1057.5441	1057.5449
XA27		67	1134.557	1134.5556
XA28		0	1251.5955	not found
XA29		71	1130.5412	1130.5412
XA30 [b]		64	1753.9288	1753.9319

a) Boc deprotection during resin cleavage; b) model peptide tag Trt-Lys-Dmf-Ala-Ser-Pro-Tyr-Val-Lys-Smx-R; c) *t*Bu deprotection during resin cleavage; d) not included in library due to potential side reactions, e) $[M+H]^+$ ion.

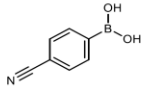
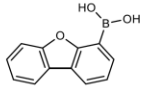
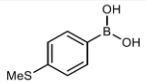
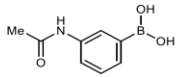
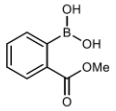
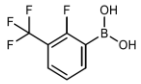
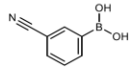
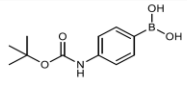
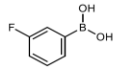
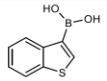
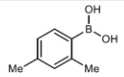
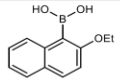
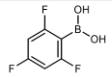
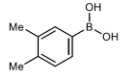
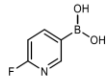
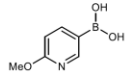
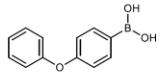
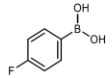
Appendix

Ar-Br for C-N Cross Coupling

triBB	Structure	Purity (%)	Mass calculated $[M+Na]^+$ (Da)	Mass found (Da)
A		68	1200.5612	1200.5557
B		100	1214.5769	1214.5766
C		100	1214.5769	1214.5774
D		100	1228.5925	1228.594
O		90	1268.6238	1268.6262
P		0	1279.6034	not found
Q		65	1255.6034	1255.6053
R		68	1314.6405	1314.6415
E		100	1228.5925	1228.5919
F		52	1228.5925	1228.5932
G		100	1228.5925	1228.5934
H		73	1229.5878	1229.5875
I		5	1230.583	1230.5837
J		51	1240.5925	1240.5929
K		84	1240.5925	1240.593
L		93	1254.6082	1254.6081
M		100	1254.6082	1254.6088
N		90	1256.5874	1256.5884

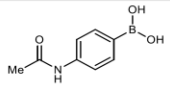
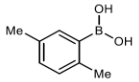
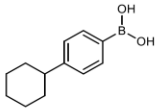
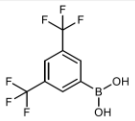
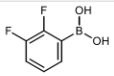
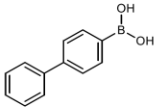
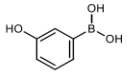
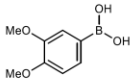
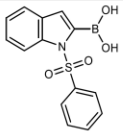
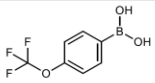
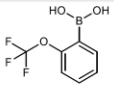
Appendix

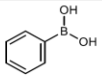
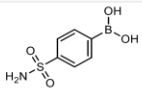
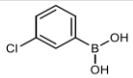
Ar-B(OH)$_2$ for C-C Cross Coupling (1)

Boronic acid	Structure	Purity (%)	Mass calculated [M+Na]$^+$ (Da)	Mass found (Da)
BA01		90	1076.5539	1076.5356
BA02		90	1141.5692	1141.5574
BA03		80	1097.5464	1097.5282
BA04		88	1108.5801	1108.5640
BA05		94	1109.5642	1109.5452
BA06		100	1137.5366	1137.5306
BA07		100	1076.5539	1076.5423
BA08		100	1066.5696 [a]	1066.5596
BA09		93	1069.5493	1069.5391
BA10		90	1107.5308	1107.5209
BA11		90	1079.59	1079.5785
BA12		91	1145.6005	1145.5948
BA13		89	1095.5849	1095.5743
BA14		73	1105.5304	1105.5186
BA15		92	1137.5366	1137.5261
BA16		92	1079.59	1079.5795
BA17		100	1070.5445	1070.5402
BA18		90	1082.5645	1082.5526
BA19		96	1143.5849	1143.5749
BA20		87	1069.5493	1069.5429
BA21		92	1081.5692	1081.5694

Appendix

Ar-B(OH)$_2$ for C-C Cross Coupling (2)

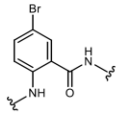
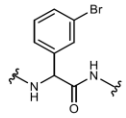
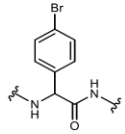
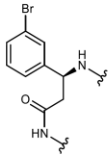
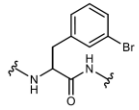
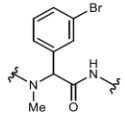
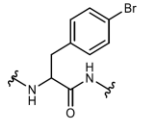
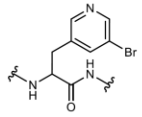
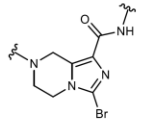
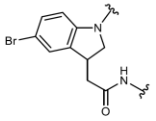
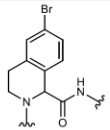
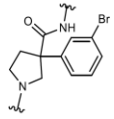
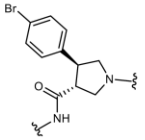
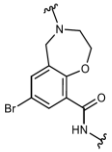
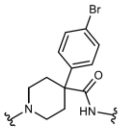
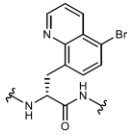
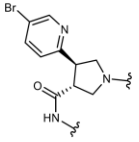
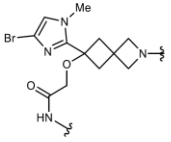
BA22		93	1108.5801	1108.5797
BA23		100	1079.59	1079.5913
BA24		98	1133.6369	1133.6374
BA25		100	1187.5334	1187.5316
BA26		90	1087.5398	1087.5375
BA27		93	1127.59	1127.5895
BA28		80	1067.5536	1067.5525
BA29		91	1070.5645	1070.5635
BA30		n/a [b]	1111.5798	n/a
BA31		100	1052.5539	1052.5522
BA32		95	1230.5628	1230.5632
BA33		89	1135.541	1135.5402
BA34		86	1135.541	1135.5410

BA35		n/a [b]	1051.5587	n/a
BA36		100	1678.8637 [c]	1678.8656
SBA1		52	1085.5197	1085.5027
SBA2		0 (91%) [d]	1041.5379	not found

a) Boc deprotection during resin cleavage; b) not confirmed; c) model peptide tag Trt-Lys-Dmf-Ala-Ser-Pro-Tyr-Val-Lys-Smx-R; d) reduction product of aryl bromide

Appendix

Ar-Br for C-C Cross Coupling

triBB	Structure	Purity (%)	Mass calculated [M+Na]⁺ (Da)	Mass found (Da)
A		68	1105.5441	1105.55
B		100	1199.566	1199.5674
C		100	1199.566	1199.5676
D		95	1213.5816	1213.5833
E		100	1213.5816	1213.5808
F		56	1213.5816	1213.5905
G		100	1213.5816	1213.5843
H		97	1214.5769	1214.5773
I		94	1215.5721	1215.5812
J		52	1225.5816	1225.5914
K		97	1225.5816	1225.5847
L		100	1239.5973	1239.5979
M		100	1239.5973	1239.5980
N		100	1241.5765	1241.5773
O		100	1253.6129	1253.6144
P		92	1264.5925	1264.5987
Q		100	1240.5925	1240.6050
R		56	1299.6296	1299.6330

Appendix

Components for C-N Based Library (1)

Building block type	BB-Nr.	encoding monomer 1	encoding monomer 2
linker	γ-aminobutyric acid	Ala	none
triBB	B	Tyr	Thr
triBB	C	Tyr	Ser
triBB	D	Tyr	Hyp
triBB	E	Tyr	Mox
triBB	F	Tyr	Pro
triBB	G	Tyr	Abu
triBB	H	Phe	Thr
triBB	I	Phe	Ser
triBB	J	Phe	Hyp
triBB	K	Phe	Mox
triBB	L	Phe	Pro
triBB	M	Phe	Abu
triBB	N	Dmf	Thr
triBB	O	Dmf	Ser
triBB	P	Dmf	Hyp
triBB	Q	Dmf	Mox
triBB	R	Dmf	Pro
CA	1	Ala	Tyr
CA	2	Thr	Tyr
CA	3	Ser	Tyr
CA	4	Hyp	Tyr
CA	5	Pro	Tyr
CA	6	Val	Tyr
CA	7	Ala	Phe
CA	8	Thr	Phe
CA	9	Ser	Phe
CA	10	Hyp	Phe
CA	11	Pro	Phe
CA	12	Val	Phe
CA	13	Ala	Mox
CA	14	Thr	Mox
CA	15	Ser	Mox
CA	16	Hyp	Mox
CA	17	Pro	Mox
CA	18	Val	Mox
CA	19	Ala	Ala
CA	20	Thr	Ala
CA	21	Ser	Ala
CA	22	Hyp	Ala
CA	23	Pro	Ala
CA	24	Val	Ala
CA	25	Ala	Cha
CA	26	Thr	Cha
CA	27	Ser	Cha
CA	28	Hyp	Cha
CA	29	Pro	Cha
CA	30	Val	Cha
CA	31	Ala	Cba
CA	32	Thr	Cba
CA	33	Ser	Cba
CA	34	Hyp	Cba
CA	35	Pro	Cba
CA	36	Val	Cba
CA	37	Ala	Leu
CA	38	Thr	Leu
CA	39	Ser	Leu
CA	40	Hyp	Leu
CA	41	Pro	Leu
CA	42	Val	Leu
CA	43	Ala	Cpa
CA	44	Thr	Cpa
CA	45	Ser	Cpa
CA	46	Hyp	Cpa
CA	47	Pro	Cpa
CA	48	Val	Cpa
CA	49	Ala	Val
CA	50	Thr	Val
CA	51	Ser	Val
CA	52	Hyp	Val
CA	53	Pro	Val
CA	54	Val	Val
CA	55	Ala	Abu
CA	56	Thr	Abu
CA	57	Ser	Abu
CA	58	Hyp	Abu
CA	59	Pro	Abu
CA	60	Val	Abu
CA	61	Ala	Dmf
CA	62	Thr	Dmf
CA	63	Ser	Dmf
CA	64	Hyp	Dmf
A	1	Cha	Thr
A	2	Val	Thr
A	3	Leu	Thr
A	4	Abu	Thr
A	5	Ala	Thr
A	6	Cba	Thr
A	7	Cha	Ser

Appendix

Components for C-N Based Library (2)

A	8	Val	Ser
A	9	Leu	Ser
A	10	Abu	Ser
A	11	Ala	Ser
A	12	Cba	Ser
A	13	Cha	Hyp
A	14	Val	Hyp
A	15	Leu	Hyp
A	16	Abu	Hyp
A	17	Ala	Hyp
A	18	Cba	Hyp
A	19	Cha	Mox
A	20	Val	Mox
A	21	Leu	Mox
A	22	Abu	Mox
A	23	Ala	Mox
A	24	Cba	Mox
A	25	Cha	Pro
A	26	Val	Pro
A	27	Leu	Pro
A	28	Abu	Pro
A	29	Ala	Pro
A	30	Cba	Pro
A	31	Cha	Ala
A	32	Val	Ala
A	33	Leu	Ala
A	34	Abu	Ala
A	35	Ala	Ala
A	36	Cba	Ala
A	37	Cha	Tyr
XA	30	Val	Tyr
A	None	Leu	Tyr

Appendix

Components for C-C Based Library (1)

Building block type	BB-Nr.	encoding monomer 1	encoding monomer 2
linker	γ-aminobutyric acid	Ala	
triBB	B	Tyr	Thr
triBB	C	Tyr	Ser
triBB	D	Tyr	Hyp
triBB	E	Tyr	Mox
triBB	F	Tyr	Pro
triBB	G	Tyr	Abu
triBB	H	Phe	Thr
triBB	I	Phe	Ser
triBB	J	Phe	Hyp
triBB	K	Phe	Mox
triBB	L	Phe	Pro
triBB	M	Phe	Abu
triBB	N	Dmf	Thr
triBB	O	Dmf	Ser
triBB	P	Dmf	Hyp
triBB	Q	Dmf	Mox
triBB	R	Dmf	Pro
CA	1	Ala	Tyr
CA	2	Thr	Tyr
CA	3	Ser	Tyr
CA	4	Hyp	Tyr
CA	5	Pro	Tyr
CA	6	Val	Tyr
CA	7	Ala	Phe
CA	8	Thr	Phe
CA	9	Ser	Phe
CA	10	Hyp	Phe
CA	11	Pro	Phe
CA	12	Val	Phe
CA	13	Ala	Mox
CA	14	Thr	Mox
CA	15	Ser	Mox
CA	16	Hyp	Mox
CA	17	Pro	Mox
CA	18	Val	Mox
CA	19	Ala	Ala
CA	20	Thr	Ala
CA	21	Ser	Ala
CA	22	Hyp	Ala
CA	23	Pro	Ala
CA	24	Val	Ala
CA	25	Ala	Cha
CA	26	Thr	Cha
CA	27	Ser	Cha
CA	28	Hyp	Cha
CA	29	Pro	Cha
CA	30	Val	Cha
CA	31	Ala	Cba
CA	32	Thr	Cba
CA	33	Ser	Cba
CA	34	Hyp	Cba
CA	35	Pro	Cba
CA	36	Val	Cba
CA	37	Ala	Leu
CA	38	Thr	Leu
CA	39	Ser	Leu
CA	40	Hyp	Leu
CA	41	Pro	Leu
CA	42	Val	Leu
CA	43	Ala	Cpa
CA	44	Thr	Cpa
CA	45	Ser	Cpa
CA	46	Hyp	Cpa
CA	47	Pro	Cpa
CA	48	Val	Cpa
CA	49	Ala	Val
CA	50	Thr	Val
CA	51	Ser	Val
CA	52	Hyp	Val
CA	53	Pro	Val
CA	54	Val	Val
CA	55	Ala	Abu
CA	56	Thr	Abu
CA	57	Ser	Abu
CA	58	Hyp	Abu
CA	59	Pro	Abu
CA	60	Val	Abu
CA	61	Ala	Dmf
CA	62	Thr	Dmf
CA	63	Ser	Dmf
CA	64	Hyp	Dmf
BA	1	Thr	Cha
BA	2	Ser	Cha
BA	3	Hyp	Cha
BA	4	Mox	Cha
BA	5	Pro	Cha
BA	6	Val	Cha
BA	7	Thr	Val

Appendix

Components for C-C Based Library (2)

BA	8	Ser	Val
BA	9	Hyp	Val
BA	10	Mox	Val
BA	11	Pro	Val
BA	12	Val	Val
BA	13	Thr	Leu
BA	14	Ser	Leu
BA	15	Hyp	Leu
BA	16	Mox	Leu
BA	17	Pro	Leu
BA	18	Val	Leu
BA	19	Thr	Abu
BA	20	Ser	Abu
BA	21	Hyp	Abu
BA	22	Mox	Abu
BA	23	Pro	Abu
BA	24	Val	Abu
BA	25	Thr	Ala
BA	26	Ser	Ala
BA	27	Hyp	Ala
BA	28	Mox	Ala
BA	29	Pro	Ala
BA	30	Val	Ala
BA	31	Thr	Cba
BA	32	Ser	Cba
BA	33	Hyp	Cba
BA	34	Mox	Cba
BA	35	Pro	Cba
BA	36	Val	Cba