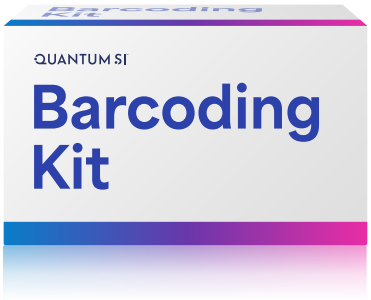
Barcoding Kit



A multiplexed protein screening solution that quantifies relative expression of peptide barcodes for protein research, engineering, and drug development.

- **Screen eight protein barcodes in multiplex** with relative quantitation across tenfold dynamic range
- **Accelerate findings with a two-day workflow** from protein to insights with less than one hour of hands-on time
- **Maximize insights with precious samples** with sensitivity down to 5 pmol per protein barcode

INTRODUCTION

Protein barcodes, with their capacity to encode extensive information in short sequences, present cutting-edge opportunities. There is a critical need for methods to directly read protein barcode sequences and to identify protein barcodes with single-molecule resolution.

Next-Generation Protein Sequencing™ (NGPS™) offers researchers the ability to directly sequence protein barcodes with single-molecule resolution for the first time. The combination of protein barcodes and NGPS enables robust, multiplexed functional protein screening and characterization for a number of applications, including screening mRNA vaccine candidates, optimizing drug delivery systems, tracking protein subcellular localization, engineering proteins, and studying protein-protein interactions.

PROTEIN BARCODING WORKFLOW

The Barcoding Kit provides methods for expression of unique peptide barcodes in proteins of interest for downstream sequencing and analysis on Platinum. The Barcoding Kit contains reagents to prepare and immobilize barcode tagged peptides for sequencing.

Unique barcode sequence designs were optimized for peptide sequencing to reliably differentiate among an 8-plex mixture. Researchers are provided with proprietary barcode and tag sequences for expression with their protein of interest.

BARCODE DESIGN

Barcode construct was designed for researchers to easily adopt to C-terminus or N-terminus of the protein of interest, leveraging readily available tags for a single-step reaction and single step-cleavage in less than four hours. Protein barcodes are cloned into a plasmid containing an

N-terminal FLAG tag, a polyG linker (GS Linker), the protein barcode sequence, a sortase recognition motif (LPETG), and an optional C-terminal His tag. A representation of the barcode construct is as follows:

Protein of Interest – Affinity Tag – GS Linker – LysC Cleavage – Barcode – Sortase Tag – His Tag

Eight barcode designs were computationally generated and empirically validated for optimal sequencing on Platinum.

Barcode ID	Sequence
BC028	RFEQIANFAELPETG
BC032	RQAELFRDYSLPETG
BC049	FQRLAELEQALPETG
BC051	FALRQDYVAQLPETG
BC067	QRESFLFLNELPETG
BC075	NDYRLSQRYLLPETG
BC079	ALQRFEQDYSLPETG
BC096	ELFNRALNAFLPETG

Table 1. Eight validated barcode sequences for use with the Barcoding Kit.

Protein purification from cell lysates depend on the protein and expression system. FLAG enrichment purification is leveraged to isolate barcoded proteins of interest.

The Barcoding Kit includes reagents necessary to prepare peptide barcode libraries for sequencing with the Platinum Sequencing Kit. Peptide barcodes are functionalized for immobilization on the sequencing chip with a sortase reaction and cleaved from FLAG beads and the protein of interest with LysC digestion. Recovered protein libraries can be loaded for sequencing or stored at -20°C.

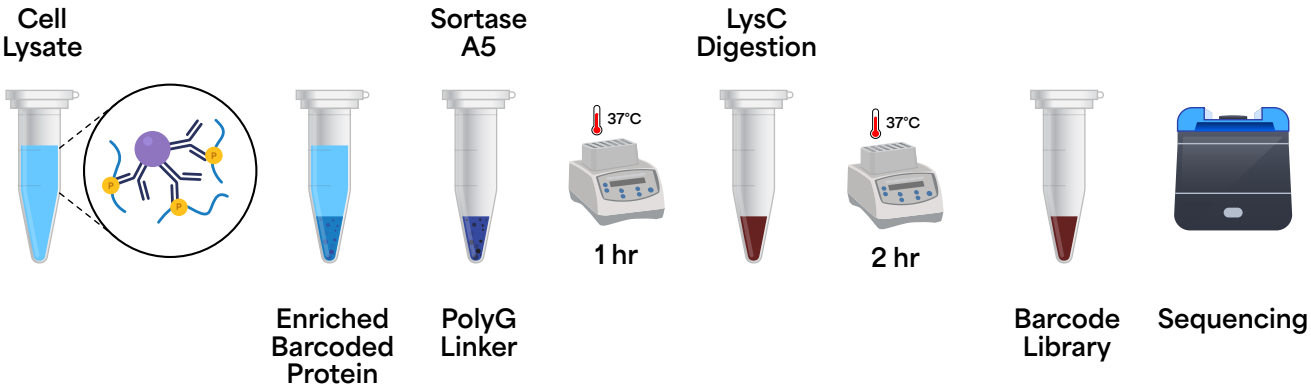


Figure 1. Barcoding workflow from cell lysate to sequenceable barcode library.

BARCODE SEQUENCING PERFORMANCE ON PLATINUM

Multiplexed analysis of barcodes on Platinum offers efficiency saving researchers time and resource cost. The Barcode Kit workflow was optimized to support an 8-plex peptide barcode run. Peptide alignments provided by the software graphs the reads aligned to all the peptides identified in a run and ranked from highest to lowest number of alignments. Each peptide has information on the number of reads successfully aligned to the reference with a false discovery rate (FDR) of less than 0.05, as shown in figure 2.

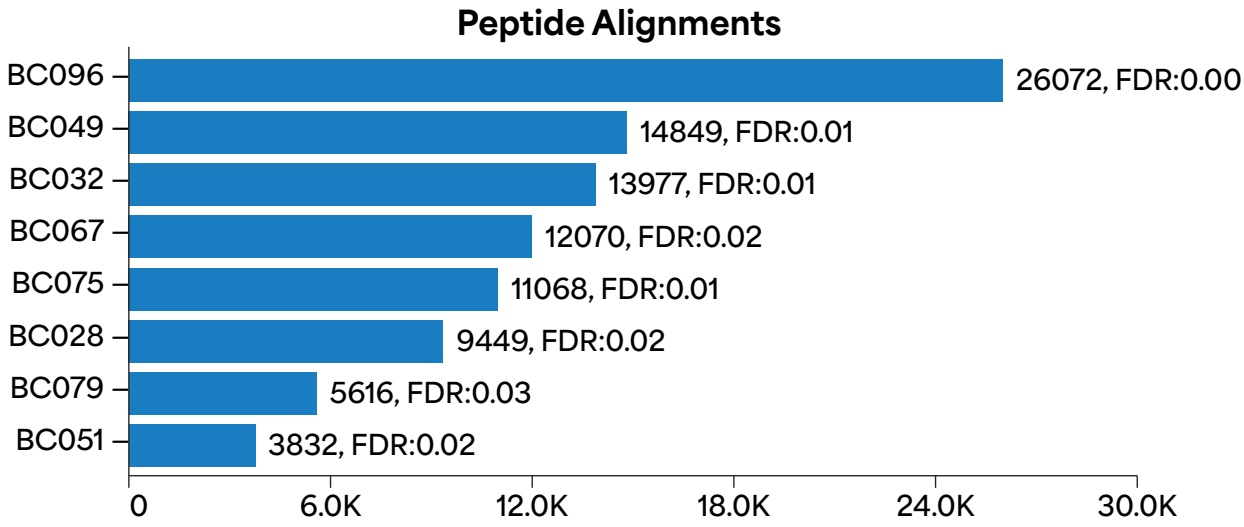


Figure 2. Peptide alignments across eight barcodes multiplexed on Platinum.

The Barcoding Kit calls for standard sample input of 5 pmol per peptide barcode and 40 pmol per 8-plex mixture. Peptide barcodes, together with the Platinum Sequencing Kit, V3, offer a highly sensitive solution for low expressing barcodes, where the lowest sample input validated for an individual barcode was 50 fmol. The limit of detection for each barcode in the complexity of 8 barcodes are in the sub-pmol range (~400 fmol) sample input (just 1% of mixture) and on-chip about 1 fmol, as shown in figure 3.

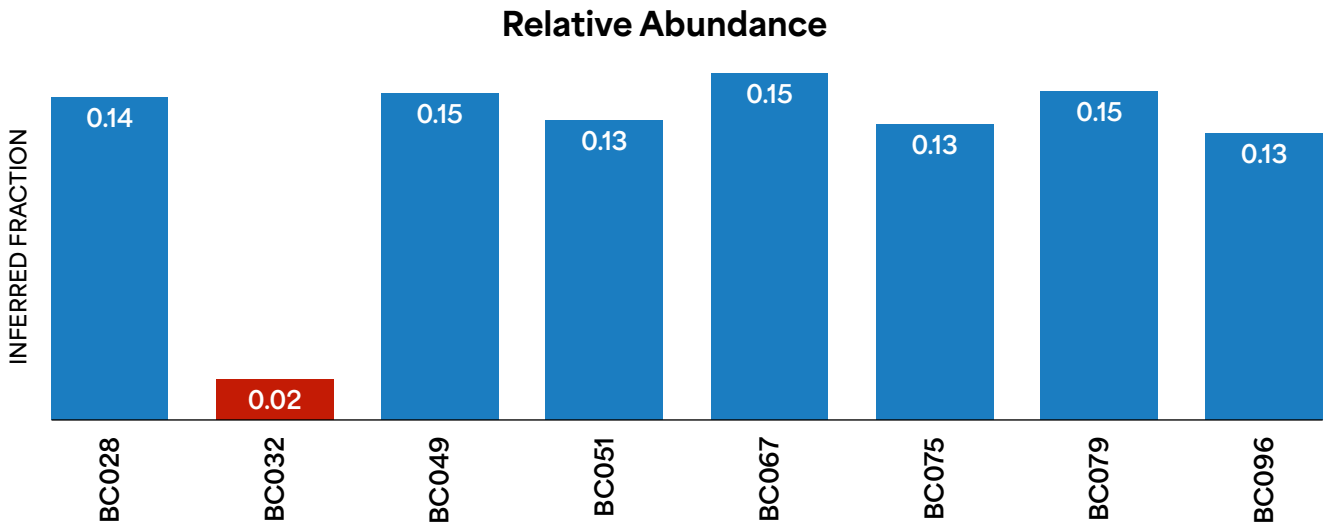


Figure 3. Relative abundance across eight barcodes with BC032 at the limit of detection.

RELATIVE QUANTITATION ACROSS TENFOLD DYNAMIC RANGE

Relative quantitation of barcode is critical to evaluate protein expression, optimize drug delivery vehicles, and to evaluate protein trafficking and other protein characteristics.

Barcodes provided in the Barcoding Kit were evaluated empirically to determine normalization factors for 8-plex sequencing on Platinum with the Sequencing Kit, V3. As demonstrated in figure 4, the normalization factors yield the expected relative fraction around 12.5% for an 8-plex mixture.

Barcodes	Sequence	Normalization Factors
BC028	RFEQIANFAELPETG	0.0939
BC032	RQAELFRDYSLPETG	0.1185
BC049	FQRLAELEQALPETG	0.1424
BC051	FALRQDYVAQLPETG	0.0314
BC067	QRESFLFLNELPETG	0.1448
BC075	NDYRLSQRYLLPETG	0.1029
BC079	ALQRFEQDYSLPETG	0.0590
BC096	ELFNRALNAFLPETG	0.3070

Table 2. Normalization factors for eight unique barcodes.

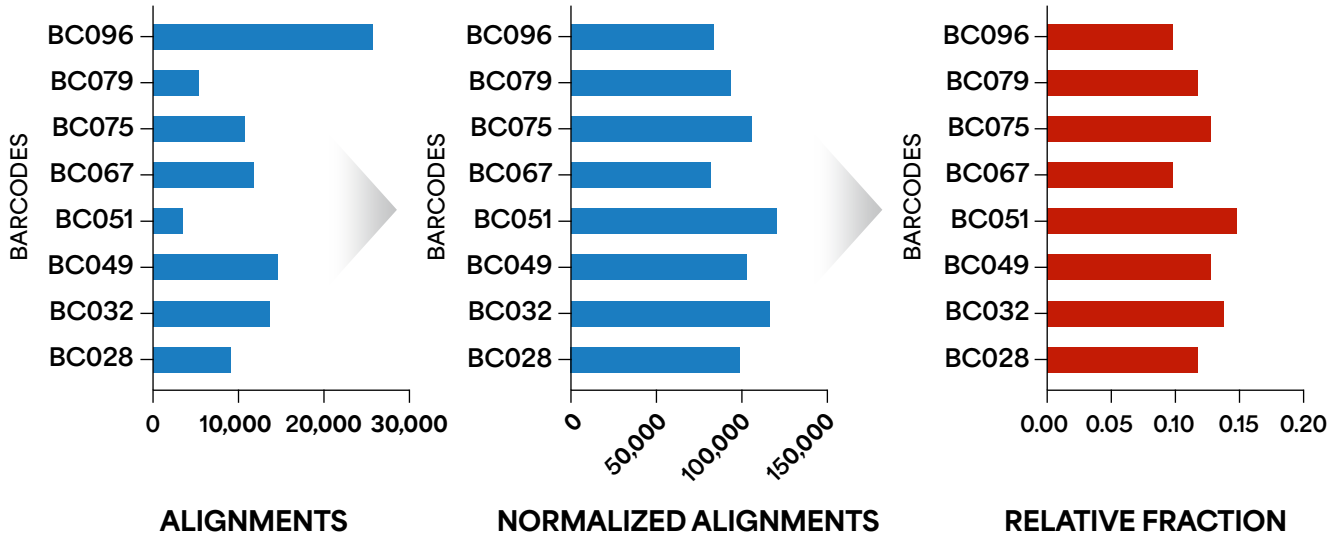


Figure 4. Normalized relative fraction of eight equimolar barcodes.

Multiplexed barcode analysis relative quantitation was verified at a tenfold dynamic range, as shown in figure 5. In applications where researchers are seeking the strongest expressors, the ability to reproducibly differentiate between the highest- and lowest-abundance proteins is supported within a defined rate of error.

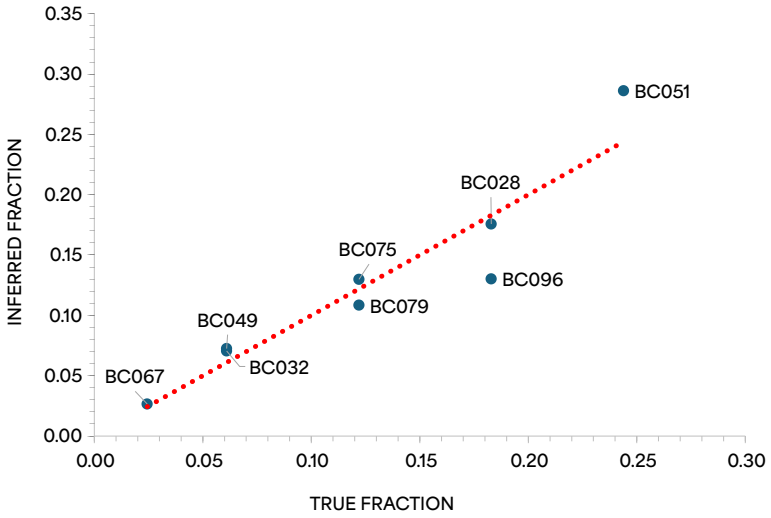


Figure 5. Eight multiplexed barcodes across tenfold dynamic range.

FLEXIBLE SOLUTION ENABLES A BROAD RANGE OF APPLICATIONS

The greatest advantage of the Barcoding Kit is its application across a broad range of protein research areas. The kit supports evaluation of drug delivery systems, relative expression of mRNA therapeutics, and assessment of multiplexed protein characteristics. The multiplex capability of the Barcoding Kit has utility across protein engineering, cell and gene therapy, mRNA therapeutics, and primary research.

mRNA Therapeutics	Cell and Gene Therapy	Protein Engineering and Research
Optimize drug delivery systems like LNPs and AAVs by integrating protein barcodes	Screen relative expression of mRNA gene therapy targets with protein barcodes *in vitro* and *in vivo*	Assess protein characteristics and cell trafficking by integrating protein barcodes

SUMMARY

The Barcoding Kit, together with the Platinum Sequencing Kit, V3, enables robust, multiplexed functional protein screening and characterization for a number of applications, including screening mRNA vaccine candidates, optimizing drug delivery systems, tracking protein subcellular localization, engineering proteins, and studying protein-protein interactions.

PRODUCT SPECIFICATIONS

Attribute	Detail
Barcoding Kit catalog number	910-00047-01
Description	Reagents for protein barcoding; generates eight libraries for sequencing
Samples per kit	Eight reactions per kit
Workflow	Two days expressed barcode to results with <1 hour hands-on time
Multiplexed barcodes per sample	8-plex barcodes per sequencing chip
Relative quantitation dynamic range	Tenfold dynamic range with >5 pmol protein input
Peptide barcode input	Down to 5 pmol per peptide barcode