

Deep HCP profiling from short-gradient DIA

DIA-NN 2.7 Enterprise identified 604 host-cell protein groups in native digests of the NISTmAb reference antibody.

Orbitrap Astral

DIA acquisition

19.9 min

LC gradient

PXD067004

Public antibody data

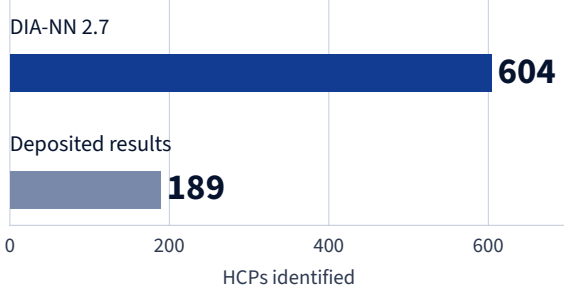
Host-cell proteins (HCPs) must be detected and quantified against the much higher abundance of the therapeutic antibody.

We reanalysed public data-independent acquisition (DIA) data for NISTmAb, a reference monoclonal antibody, prepared by native digestion or SP3 (single-pot, solid-phase-enhanced sample preparation). [1] We compared the DIA-NN results with identifications deposited by the original study.

HCPs detected in NISTmAb

Six injections per method

Native digestion



SP3 digestion

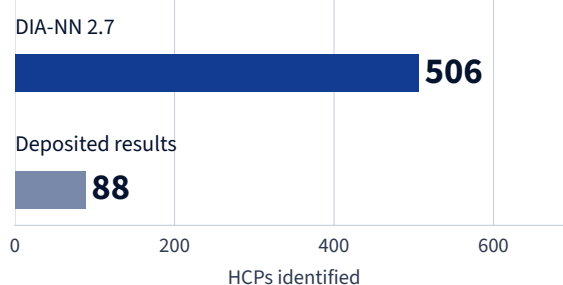
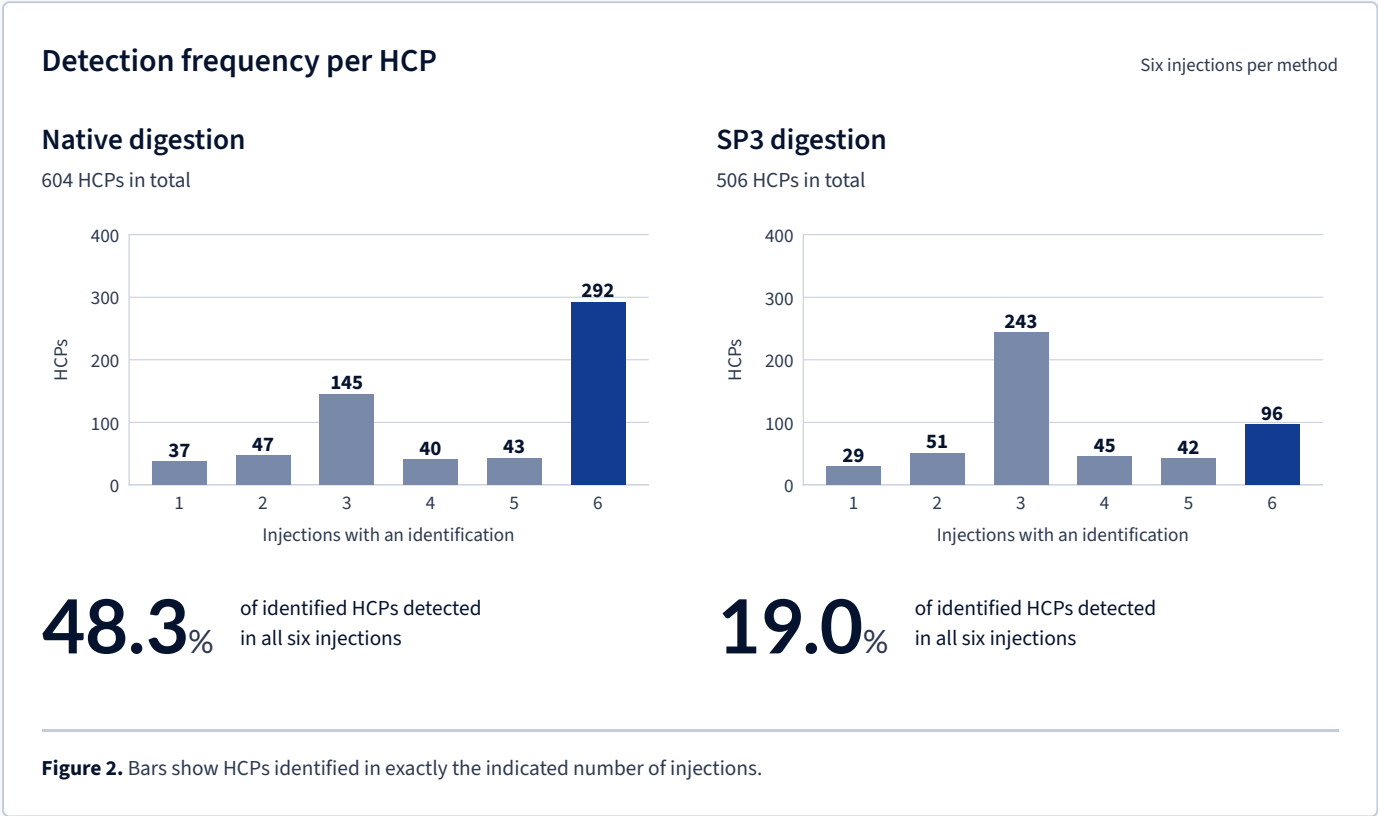


Figure 1. HCP counts from DIA-NN (protein groups) and the original study's deposited results (protein accessions). [1]

HCP detection across preparations

Two NISTmAb digests were prepared independently by each method, and each digest was injected three times.



Within and between digests

For 241 of the 243 SP3 HCPs detected in exactly three injections, all three detections came from digest 2.

Detection in both digests was more common after native digestion: 64.7% of identified HCPs, compared with 36.6% after SP3.

SP3 detection by digest

Detected in	DIA-NN groups	Deposited accessions
Digest 1	187	42
Digest 2	504	88
Both digests	185	42
All six injections	96	28
	19.0%	31.8%

02 / CHROMATOGRAMS

Peptide evidence for an HCP identification

We examined a peptide from protein disulfide-isomerase A6 (Pdia6; Q922R8) across the six native NISTmAb injections.

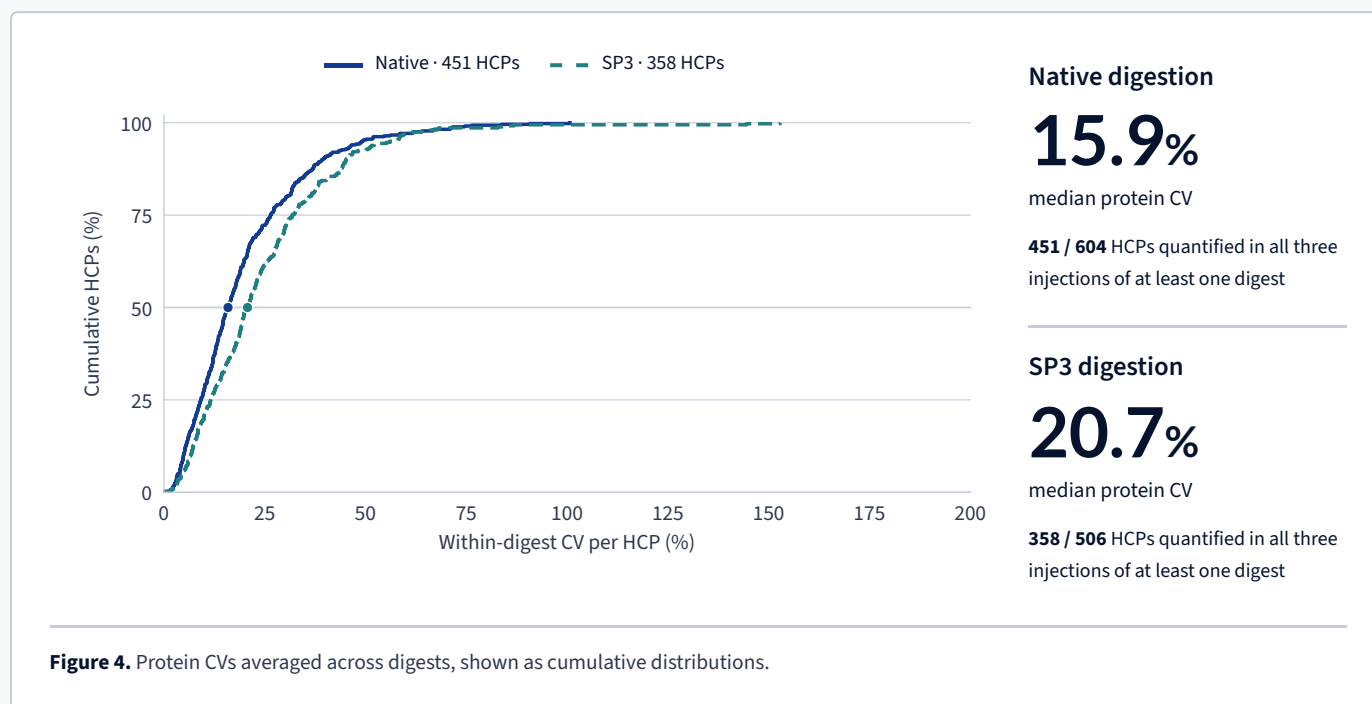


Coeluting fragment ions and consistent relative intensities across all six injections support the peptide identification.

03 / QUANTUMS

Repeatability of HCP quantification

We assessed QuantUMS quantification using protein coefficients of variation (CVs) across three injections of each digest. [3]



Injection precision by digest

SP3 digest 1 yielded fewer HCP identifications and a higher median CV than digest 2.

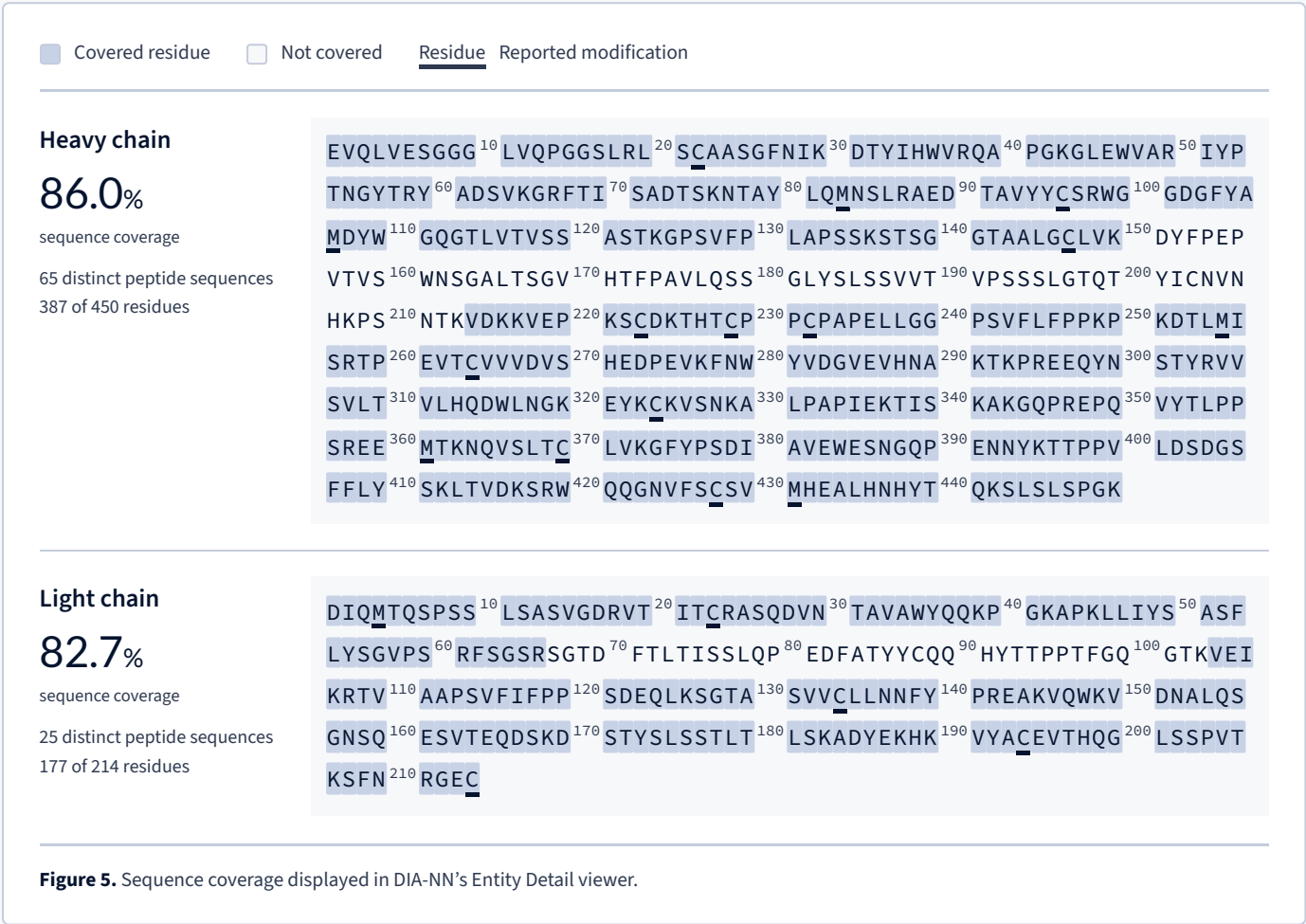
285 of 451 native-digest HCPs and 174 of 358 SP3 HCPs had CVs below 20%.

Protein CVs calculated within each digest

Preparation	HCPs	Median CV
Native · digest 1	448	14.7%
Native · digest 2	272	12.9%
SP3 · digest 1	90	30.0%
SP3 · digest 2	356	16.0%

Antibody sequence coverage in Herceptin

Herceptin (trastuzumab, lot H4759H01-A025) was analysed in three SP3 digests, each injected twice. [1] The same heavy- and light-chain residues were covered in all six injections.



05 / METHODS

Analysis details

Samples and acquisition

We analysed 18 files from PXD067004. [1] NISTmAb native digestion and SP3 each comprised two independent digests with three injections per digest. Herceptin comprised three SP3 digests with two injections each.

Orbitrap Astral DIA data were acquired with a 19.9 min gradient in a 24 min LC method and nominal loading of 500 ng per injection. Native digestion used 1 mg NISTmAb; SP3 used 50 µg.

DIA-NN analysis

Each six-injection experiment was analysed separately with DIA-NN 2.7.0 Enterprise, using Generic scoring, MBR, Knowledge base and QuantUMS (precision). [2] [3] We used MS1 and MS2 mass tolerances of 4 and 10 ppm, respectively, and a calibration library derived from public K562 data. Proteotypicity was set to Isoform IDs.

Library prediction used Trypsin/P, up to two missed cleavages, peptide lengths of 7–30 residues, charges 1–4, precursor m/z 380–980 and fragment m/z 150–2000. Carbamidomethyl (C) was fixed; N-term M excision, Ox (M) and Acetyl (N-term) were enabled, with a maximum of one variable modification allowed.

Search databases contained reviewed mouse sequences (UniProt 2023_02) for NISTmAb or Chinese hamster sequences (UniProt 2026_03) for Herceptin, plus the corresponding antibody-chain sequences, the *E. coli* ClpB sequence and DIA-NN's built-in contaminants.

HCP identification

To obtain HCP counts and calculate CVs, we applied 1% run-specific and global protein-group q-value filters (PG.Q.Value and Global.PG.Q.Value). We counted groups containing only mouse proteins for NISTmAb or Chinese hamster proteins for Herceptin.

For both DIA-NN and deposited HCP counts, we excluded antibody chains, ClpB, database contaminants, keratins, immunoglobulin-related proteins, desmosomal contaminants and proteins listed in the original study's Tables S6/S7.

Protein quantification

Protein CVs were calculated within each digest using at least three positive quantities and averaged across digests.

Deposited results

We counted mouse protein accessions linked to peptide identifications in the deposited native and SP3 mzIdentML files. SP3 injection assignments came from the corresponding mzML file.

The analyses used different databases and processing settings. We did not independently validate the search-engine outputs.

Herceptin coverage

Herceptin sequence coverage was calculated using 1% run-specific and global precursor q-value filters (Q.Value and Global.Q.Value).

References

1. Smith J et al. *Quantitative Host Cell Protein Analysis of Antibody-Based Protein Therapeutics Using the Orbitrap Astral Mass Spectrometer*. J Am Soc Mass Spectrom. 2026;37:1077–1094. doi:10.1021/jasms.5c00272. Data: PXD067004.
2. Demichev V et al. *DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput*. Nat Methods. 2020. doi:10.1038/s41592-019-0638-x.
3. Grossmann JL et al. *Accurate quantification in proteomics with QuantUMS*. Nat Biotechnol. 2026. doi:10.1038/s41587-026-03131-2.