QUANTUM SI

Identification and characterization of protein complexes using a benchtop single-molecule protein sequencer

David Sloan,[1] Bonnie Lun,[1] Kendrick Nguyen,[1] Anthony Oppedisano,[2] Esteban Lenoir,[2] Elizabeth Stephens,[3] Ajay Vashisht,[1] John Vieceli,[1] Phil Hieter,[3] Nigel O'Neil[2]

1. Quantum-Si Incorporated, 29 Business Park Drive, Branford, CT 06405, USA
2. Arrowsmith Genetics, Vancouver, BC, Canada
3. Michael Smith Laboratories, University of British Columbia, BC, Canada

Abstract

Introduction/methods

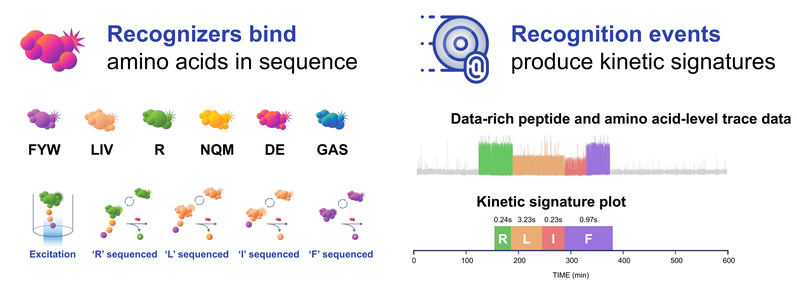
Accurate and efficient protein identification remains a central challenge in modern biochemistry and molecular biology. Traditional methods often depend on antibody-based detection, which can be limited by antibody availability and specificity, or on complex and costly instrumentation that restricts accessibility for many laboratories. Consequently, there is a growing need for innovative approaches that enable direct and unbiased protein analysis without these constraints. Recent advances in single-molecule protein sequencing have opened new avenues for proteomic analysis, allowing for precise identification of proteins and characterization of protein assemblies. In this study, we performed next-gen protein sequencing (NGPS) on the Platinum®/Platinum® Pro benchtop instruments. These single-molecule protein sequencers utilize fluorescently labeled recognizer proteins that bind to and identify amino acids in sequence, generating peptide sequences for individual molecules attached to a semiconductor chip. We applied NGPS to several proteomic analyses. We first used NGPS to QC protein purifications, detecting host cell protein contaminants and highlighting the effectiveness of this method for quality assurance and risk mitigation in downstream applications. We also leveraged NGPS to identify an unknown protein band on an SDS-PAGE gel, demonstrating the system's capacity for direct protein identification without the need for antibodies or affinity enrichment. Finally, we employed the system to analyze the composition of multi-component protein complexes in yeast cells heterologously expressing human cohesin subunits.

Results/conclusions

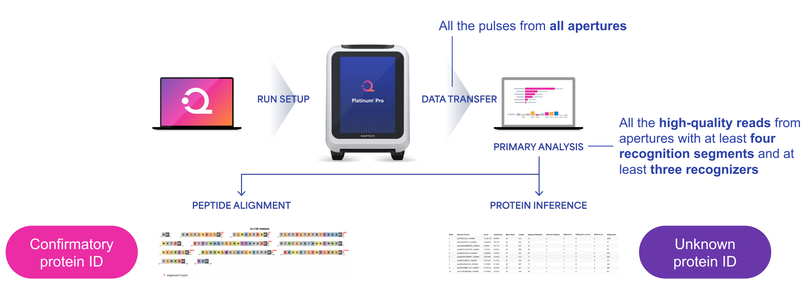
NGPS could detect the presence of hybrid cohesin complexes comprised of both yeast and human subunits, even in the absence of targeted antibodies. NGPS can also be used for proteoform characterization, enabling interrogation of post-translational modifications and sequence variants. These findings establish NGPS on Platinum Pro as a versatile tool for comprehensive and accessible proteomic analysis.

Introduction to NGPS on Platinum Pro

- **Simple sample prep process:** Proteins are digested, functionalized, conjugated, and immobilized on the surface of a semiconductor chip.
- **Unique *kinetic signature* mechanism:** Fluorescently labeled N-terminal amino acid (NAA) recognizers and aminopeptidases are added to the semiconductor chip.
- **Single-molecule level data:** Fluorescent intensity and duration of each NAA binding event generate a unique kinetic signature.
- **Automated analysis:** Kinetic signatures are automatically analyzed to align reads to reference peptides and compute false discovery rate (FDR).



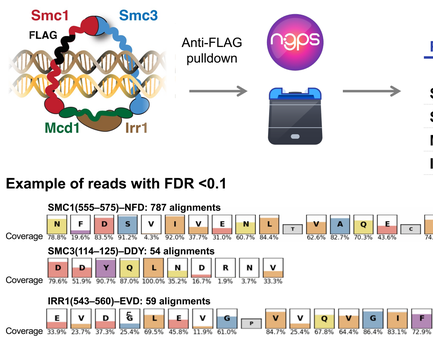
Overview of NGPS workflows



Characterizing cohesin model system protein complexes

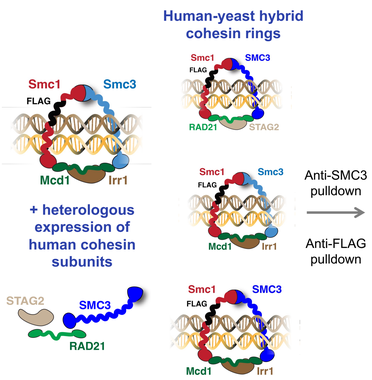
Figure 1. Human cohesin complex (left) and yeast homolog (right). Human cohesin complex composes four core subunits, namely SMC1, SMC3. RAD21, and STAG. Yeast cohesin complex composes four core subunits, including Smc1, Smc3, Mcd1, and yIrr1.

The project goal was to develop a humanized yeast model for drug development and testing. The human-yeast cohesin hybrid complex was cloned and expressed in *S. cerevisiae*. Heterologous expression of human cohesin in yeast resulted in a dominant negative phenotype. To investigate whether dysfunctional hybrid complexes of human and yeast cohesin were forming, we co-immunoprecipitated human and yeast cohesin subunits followed by NGPS. NGPS was necessary to identify yeast subunits as there are no available antibodies for yeast cohesin proteins.



Pulldown	Smc1-FLAG Anti-FLAG	
	Reads	**Peptides** (FDR <0.1)
Smc1	2,735	40 (10)
Smc3	535	46 (6)
Mcd1	78	13 (0)
Irr1	272	22 (3)

Figure 2. Human cohesin subunits were engineered and expressed in *S. cerevisiae*. Hybrid cohesin complexes could be composed of any combination of the four human and yeast subunits. Complexes were co-immunoprecipitated with either anti-FLAG or anti-SMC3 and subjected to NGPS. Both co-immunoprecipitation experiments identified both human and yeast cohesin subunits supporting the presence of inter-species hybrid rings. The presence of human cohesin subunits in the yeast Smc1-FLAG pulldown was confirmed by Western blotting (lower).



Pulldown	Smc1-Flag; hSMC3, hRAD21, hSTAG2 neochromosome: Anti-FLAG (2×)		Anti-hSMC3	
	Reads	**Peptides** (FDR <0.1)	**Reads**	**Peptides** (FDR <0.1)
Human				
SMC3	509	22 (4)	164	20 (3)
RAD21	39	2 (0)	12	3 (0)
STAG2	293	8 (2)	78	9 (1)
Yeast				
Smc1	1,682	24 (11)	215	17 (3)
Smc3	862	28 (6)	222	22 (4)
Mcd1	124	8 (1)	24	5 (0)
Irr1	277	8 (2)	167	13 (1)

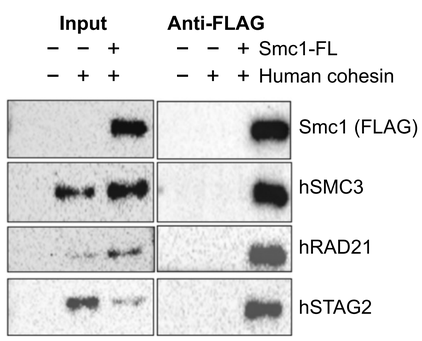
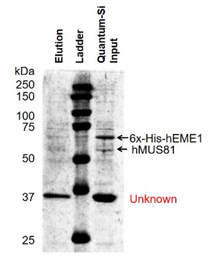


Figure 3. Western blots confirm human and yeast cohesin subunits co-immunoprecipitated.

Using NGPS as quality control of a purified protein sample

- hMUS81/hEME1-6xHis expressed in yeast and purified on nickel column.
- SDS-PAGE protein gel revealed bands from multiple unknown contaminants.
- Eluted sample analyzed via NGPS to identify contaminants.



NGPS inference results

Rank	Inferred protein	Score (rounded)	Likelihood (rounded)	Mass (kDa)	Length (aa)	Abundance* (molecules/cell 10^3)
1	DPH1	1	66%	48	425	**2.7** +/- 0.7
2	TDH1	0.4	33%	36	332	**80.4** +/- 58.3
3	ICP55	0.4	33%	58	511	**16.8** +/- 4.3
4	RKM2	0.2	20%	56	479	**0.6** +/- 0.2
5	DAN1	0.2	19%	30	298	N/A
6	Dubious ORF	0.2	19%	13	109	N/A
7	ACH1	0.2	16%	59	526	**8.5** +/- 3.8

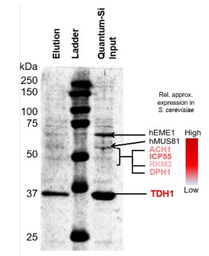


Figure 4. NGPS successfully identified both major and minor contaminants in the elution.

Rapid identification of unknown bands on a gel

- Researchers observed multiple bands on an SDS-PAGE gel spanning the proteasome complexes and other proteins of interest.
- The goal is to identify the unknown bands by NGPS.
- Software identified proteasome complex subunits as top hit in protein inference workflow.

Inference

Rank	Inferred protein	Score	Likelihood	Mass (kDa)	
1	sp|P20618|PAB1_HUMAN	0.343661	29.08%	26	« Beta type 1
2	sp|P49720|PSB3_HUMAN	0.247431	21.91%	23	« Beta type 3
3	sp|Q6IPT2|GAR5A_HUMAN	0.107635	10.2%	28	
4	sp|Q9GZP9|DERL2_HUMAN	0.101229	9.62%	28	
5	sp|A6NHS7|MANS4_HUMAN	0.087937	8.41%	38	
6	sp|Q9282IN|NJ1_HUMAN	0.081039	7.78%	16	
7	sp|Q96AX2|RAB37_HUMAN	0.081039	7.78%	25	
8	sp|Q8TDN7|ACER1_HUMAN	0.066704	6.45%	31	
9	sp|Q96L2|IRL10L_HUMAN	0.063196	6.12%	25	
10	sp|P27635|RL10_HUMAN	0.063196	6.12%	25	

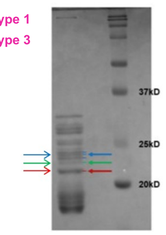


Figure 5. NGPS protein inference analyses infer both major and minor contaminants from sample elution.

Conclusions

Next-gen protein sequencing is an accessible, benchtop workflow that can be used for multiple protein identification, analysis, and QC applications, including:

- Discriminating between yeast and human cohesin complex proteins, even in the absence of specific antibodies.
- Clearly identifying the presence of multiple subunits in human/yeast hybrid cohesin protein complexes.
- Analyzing a purified protein sample to identify the presence of host cell protein contaminants.
- Sequencing bands excised from a Coomassie-stained protein gel and identifying the bands from a reference genome.

References

1. Stephens, E. et al. In preparation (2026)

© 2026 Quantum-Si. All rights reserved.