Advanced LC-MS Capabilities for Protein, Peptide, Glycan, and Oligonucleotide Analysis at the Emory EGMIC Core

Yi Lasanajak[1], Afzal Yacoob[1], Xuezheng Song[1,2], Blaine Roberts[1,2]

1. Emory Glycomics and Molecular Interaction Core, Emory University, Atlanta, GA 30322
2. Department of Biochemistry,Emory University, Atlanta, GA 30322

Emory Glycomics and Molecular Interactions Core
Emory Integrated Core Facilities

Abstract

Emory Glycomics and Molecular Interactions Core (EGMIC) is equipped with top-of-the-line LC-MS instruments. Our core provides extensive support to researchers through a broad range of mass spectrometry tools and techniques. This includes the options to learn to conduct LC-MS analysis for trainees or to have our dedicated staff run the samples. We are unique in the world by offering ion mobility mass spectrometry with electron capture dissociation. These techniques allow characterization of the shape and quaternary structure of proteins, and resolve glycan, and lipid isomers. In addition, we conduct validation of reagents especially recombinant proteins and synthetic peptides. This allows the user to confirm they received what they ordered or made the correct protein. We find 10% or more of synthetic peptides from reputable suppliers are incorrect and recombinant proteins contain surprising additions or truncations ~25% of the time. LC-MS offers precise molecular mass measurements and comprehensive analyses of proteins, peptides, glycans, and oligonucleotides. For antibodies, proteins and peptides, LC-MS provides insights into purity, folding stages, dimer configurations, protein interactions, post-translational modifications, and top-down sequencing. In glycan analysis, LC-MS enables the study of various glycoforms and glycosylation in monoclonal antibodies, utilizing workflows such as Agilent InstantPC tagging for rapid and high-throughput N-glycan composition analysis. Validation of synthetic oligonucleotides, LC-MS delivers critical data on sequence confirmation, purity, and structural modifications. Our team is dedicated to helping you leverage LC-MS to gain deeper insights into your samples, ensuring high-quality data and impactful scientific outcomes.

Our Team

Xuezheng Song, *PhD*; Scientific Director
Blaine Roberts, *PhD*; Scientific Director
Yi Lasanajak, *MS, MSPH*; Technical Director
Afzal Yacoob, *BS;* Lead Research Specialist

Whitehead Biomedical Research Building
Room 665, 615 Michael St, Atlanta, GA 30322
Website: cores.emory.edu/egmic
Email: EGMIC@emory.edu
Scheduling and Invoicing:

PPMS for the Emory Glycomics and Molecular Interactions Core - EGMIC

Application 1: Protein. Replace your Westerns, ELISA, study protein folding.

➢ Example -1: Rapid and reproducible native protein analysis

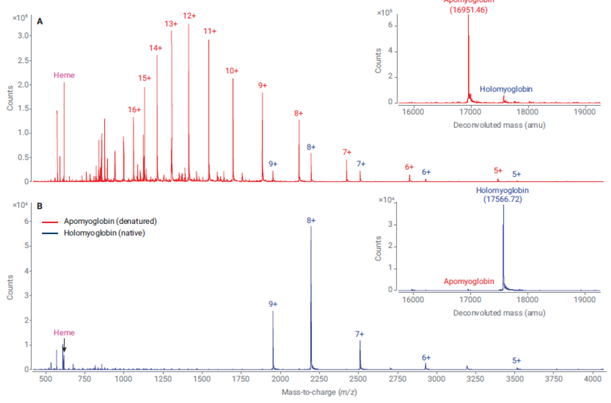
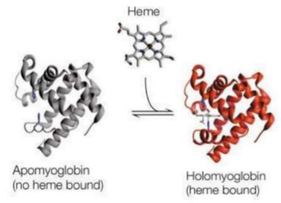


Figure 2. LC/MS analysis of intact myoglobin sample. A) Myoglobin sample was analyzed under denatured LC/MS conditions (previous studies). The heme group was dissociated from the protein complex and the majority of the protein was apomyoglobin (inset figure). B) Native MS analysis of myoglobin. The holomyoglobin (with heme) structure was preserved and only trace amount of heme was detected.



- Native MS analysis of noncovalent interaction of myoglobin. Heme is noncovalently attached to the globin through hydrogen bonds and hydrophobic interactions.
- Monitor the charge distribution of myoglobin ions in mass spectra to study folding/unfolding.
- Apomyoglobin with high charge states indicated the disruption of the native heme-protein interaction.
- Native holomyoglobin was denatured into apomyoglobin under harsh condition.

Source: Agilent Application Note#5994-1739EN

➢ Example -2: Native mass spectrometry to characterize intact protein and protein complex

- Native mass spectrum of the tetrameric pyruvate kinase (PK) showed two major charge envelopes. The deconvoluted spectrum revealed two multi-proteoform complex of PK tetramers: full-length PK and truncated PK tetramer.
- Native MS spectrum of Glutamate dehydrogenase (GDH). The molecular mass of the intact hexameric GDH was determined to be 334,754 Da.
- Native MS spectrum of beta-galactosidase. The molecular mass of the intact tetrameric beta-galactosidase was determined to be 465,788 Da with 1 to 10 microgram sample injection.

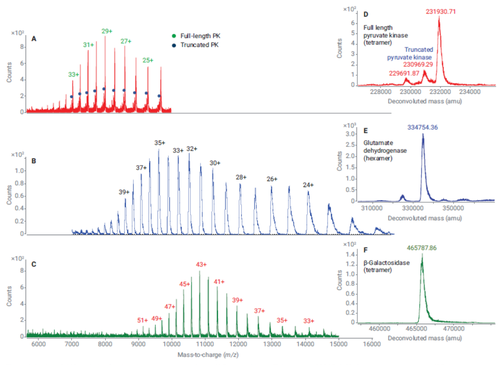
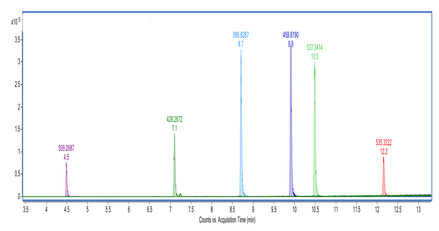


Figure 6. Native LC/MS analysis of various intact protein complexes. A) Pyruvate kinase (PK, tetramer, 232 kDa), B) glutamate dehydrogenase (GDH, hexamer, 335 kDa) and C) β-galactosidase (tetramer, 466 kDa). The deconvoluted spectra are shown in D) to F), respectively. The raw MS spectrum in Figure 6B was smoothed using the mMass open-source MS software tool.

Source: Agilent Application Note#5994-1739EN

Application 2: Peptide. Post-translational modifications (PTMs).

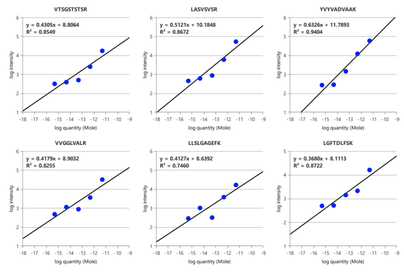
➢ Example-3: Sensitive peptide detection



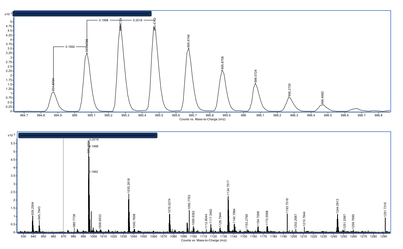
Peptide standards

Peptide Sequence	Mass-to-Charge Ratio
VTSGSTSTSR	509.3 m/z
LASVSVSR	428.2 m/z
YVYVADVAAK	566.8 m/z
VVGGLVALB	459.8 m/z
LLSLGAGEFK	537.3 m/z
LGFTDLFSK	535.3 m/z

Source: EGMIC internal



➢ Example-4: Is your peptide a dimer? Do you have a pure peptide product?



Expected mass: 4969.328

Monomer		Dimer	
Z	M/Z	Z	M/Z
4	1243.332	7	1420.808
5	994.8656	8	1243.332
6	829.2213	9	1105.295
7	710.904	10	994.8656
8	622.166	11	904.5142

- Target peptide identified, expected mass: 4969.328
- Observed mass: Z=5, 1st isotope mass=994.8734, calculated mass: 4969.311
- This peptide is not a dimer since no presence of mass expected from a dimer (1420.808, 1105.295, 904.5142).
- This sample contained a mixture of peptides.

Source: EGMIC internal

➢ Example-5: Top-down amino acid sequences

- High-throughput workflow of peptide mapping provides complete amino acid sequences of mAb
- Automatic data analysis for complete sequence mapping
- Provides information on post-translational modifications (PTMs) and location

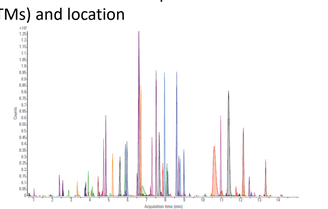


Figure 2. ECC of peptides from Trypsin/Lys-C digested NIST mAb standard RM 8671, separated using an Agilent AdvanceBio Peptide Mapping column.

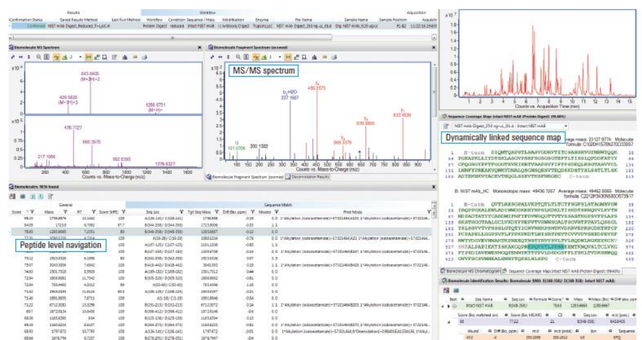


Figure 4. Screenshot of Agilent MassHunter BioConfirm B.08 software with representative peptide mapping results and protein sequence coverage.

Source: Agilent Application Note#5991-7815EN

Application 3: Glycan, glycoform, glycopeptide in Glycosylation study

Four levels of LC/MS workflows

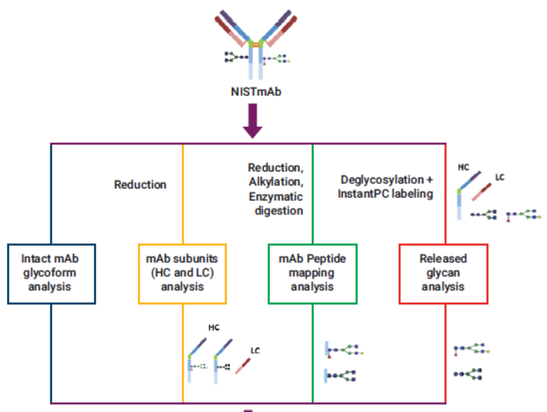


Figure 1. Various glycoforms/glycans quantitative analysis workflows.

➢ Example-6: Characterization glycoforms of intact NISTmAb

- Relative quantification of mAb with 5 glycoforms

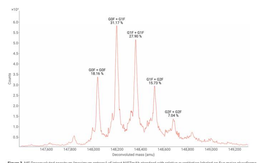


Figure 3. MS Deconvoluted spectrum (maximum entropy) of intact NISTmAb standard with relative quantitation labeled on five major glycoforms.

- Subunits analysis. Three major glycoforms were observed in heavy chain (HC)

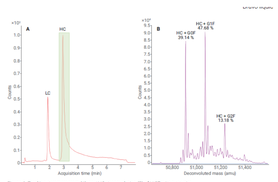
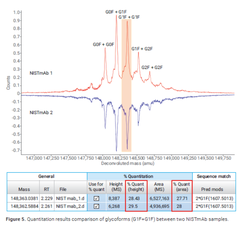


Figure 6. Total ion chromatogram (A), and MS deconvolution (B) of NISTmAb sub-units.

- Quantification results comparison of glycoforms from two NISTmAb samples



General				% Quantitation				Sequence match
Mass	RT	File	Use for % quant	Height (MS)	% Quant (height)	Area (MS)	% Quant (area)	Pred mods
148,063.0381	2.229	NIST mab_1.d	☐	8,997	28.49	6,527,160	27.71	2*G1F(1607.5013)
148,063.5884	2.261	NIST mab_2.d	☐	9,268	29.5	4,936,489	28	2*G1F(1607.5013)

Figure 5. Quantitation results comparison of glycoforms (G1F+G1F) between two NISTmAb samples.

- InstantPC labeled N-glycans from NISTmAb

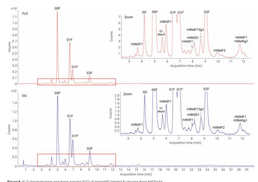


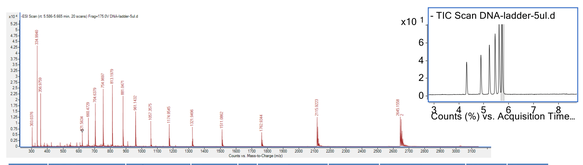
Figure 8. Full chromatogram and mass spectra (ESI) of instantPC labeled N-glycans from NISTmAb.

Source: Agilent Application Note#5991-8796EN

Application 4: Oligonucleotides (DNA/RNA/ASO). Impurity, truncations, additions.

Figure 1. Analytical components of the oligonucleotide analysis - Target Plus Impurities (TPI) workflow.

➢ Example-8: Oligonucleotides detection



Peak Index	LC-RT	m/z	Gap	z	Mass from BioConfirm	Seq	Expected mass	Mass Diff	Error PPM
1	4.245-4.436	1498.5697	0.3334	3	4500.9	15mer	4500.9	0	0
2	4.804-4.941	2005.3126	0.3341	3	6022.1	20mer	6021.9	0.2	33.2
3	5.197-5.293	1883.7857	0.25	4	7543.0	25mer	7542.9	0.1	13.3
4	5.421-5.5512	2264.0989	0.25	4	9064.0	30mer	9063.9	0.1	11.0
5	5.586-5.665	2644.4051	0.25	4	10584.9	35mer	10584.8	0.1	9.4
6	5.677-5.797	2419.5669	0.20	5	12105.8	40mer	12105.8	0	0

Source: EGMIC internal

➢ Example-9: Impurities analysis

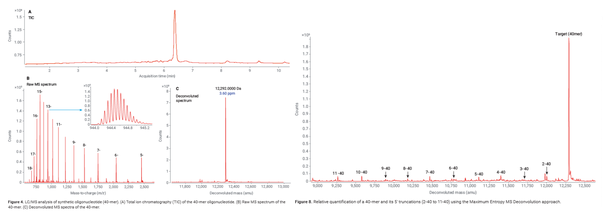


Figure 4. LC/MS analysis of synthetic oligonucleotide (40-mer). (A) Total ion chromatography (TIC) of the 40-mer oligonucleotide. (B) Raw MS spectrum of the 40-mer. (C) Deconvoluted MS spectra of the 40-mer.

Figure 5. Impurity analysis summary on the top 15 oligonucleotide impurities of a 40-mer synthetic oligonucleotide.

Oligo Length	RT (min)	Measured Mass	%Quant	Sequence
16–40	5.861	7,666.0915	2.58	ASCCAATGAATCGGAGTCGAGATCCAT
15–40	5.878	7,988.2299	2.26	CASCCAATGAATCGGAGTCGAGATCCAT
14–40	5.994	8,301.3760	1.24	ACASCCAATGAATCGGAGTCGAGATCCAT
13–40	5.990	8,631.5440	1.35	GACASCCAATGAATCGGAGTCGAGATCCAT
12–40	6.048	8,935.0243	1.60	TGACASCCAATGAATCGGAGTCGAGATCCAT
11–40	6.073	9,263.8039	2.39	GTGACASCCAATGAATCGGAGTCGAGATCCAT
10–40	6.148	9,577.3936	2.54	AGTGACASCCAATGAATCGGAGTCGAGATCCAT
9–40	6.193	9,891.1191	2.94	AAGTGACASCCAATGAATCGGAGTCGAGATCCAT
8–40	6.185	10,179.3068	1.18	CAAGTGACASCCAATGAATCGGAGTCGAGATCCAT
7–40	6.214	10,468.9206	2.36	CCAAGTGACASCCAATGAATCGGAGTCGAGATCCAT
6–40	6.289	10,782.0817	1.08	ACCAAGTGACASCCAATGAATCGGAGTCGAGATCCAT
5–40	6.276	11,111.0793	2.03	GACCAAGTGACASCCAATGAATCGGAGTCGAGATCCAT
4–40	6.280	11,400.2121	2.82	GGACCAAGTGACASCCAATGAATCGGAGTCGAGATCCAT
3–40	6.334	11,712.5835	0.65	AGGACCAAGTGACASCCAATGAATCGGAGTCGAGATCCAT
2–40	6.339	12,002.3602	2.68	CAGGACCAAGTGACASCCAATGAATCGGAGTCGAGATCCAT
Target	6.384	12,292.1749	72.80	CCAGGACCAAGTGACASCCAATGAATCGGAGTCGAGATCCAT

- Targeted approach: Target plus impurities (TPI) workflow
- Untargeted approach: Maximum Entropy Deconvolution

Source: Agilent Application Note#5994-4817EN

➢ Example-10: Rapid mRNA capping analysis

- Sensitive and efficient method for process optimization and quality control of nucleic acid therapies

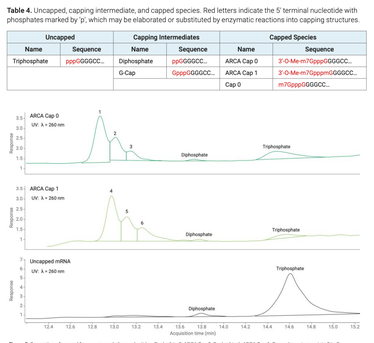


Table 4. Uncapped, capping intermediate, and capped species. Red letters indicate the 5′ terminal nucleotide with phosphates marked by 'p', which may be elaborated or substituted by enzymatic reactions into capping structures.

Uncapped		Capping Intermediates		Capped Species	
Name	Sequence	Name	Sequence	Name	Sequence
Triphosphate	pppGGGGCC…	Diphosphate	ppGGGGCC…	ARCA Cap 0	3′-O-Me-m7GpppGGGGCC…
		G-Cap	GpppGGGGCC…	ARCA Cap 1	3′-O-Me-m7GpppmGGGGCC…
				Cap 0	m7GpppGGGGCC…

Figure 5. Separation of capped from uncapped oligonucleotides. Peaks 1 to 3: ARCA Cap 0, Peaks 4 to 6: ARCA Cap 1. Capped species contain 0 to 2 non-templated G nucleotides as slippage sequence variants.

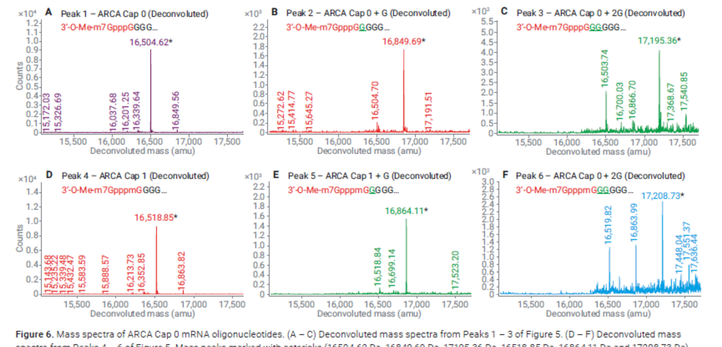


Figure 6. Mass spectra of ARCA Cap 0 mRNA oligonucleotides. (A – C) Deconvoluted mass spectra from Peaks 1 – 3 of Figure 5. (D – F) Deconvoluted mass spectra from Peaks 4 – 6 of Figure 5. Mass peaks marked with asterisks (16504.62 Da, 16849.69 Da, 17195.36 Da, 16518.85 Da, 16864.11 Da and 17208.73 Da) were matched to their putative identities (sequences inset) with < 20 ppm mass error (Table 5). Underlined letters in green indicate non-templated nucleotides likely due to T7 transcriptional slippage.

Source: Agilent Application Note#5994-3984EN

➢ Example-7: N-glycan profiling from glycoprotein using InstantPC tag labeling workflow

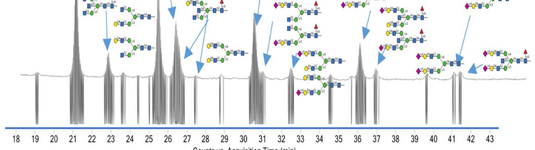
- High sensitivity
- Confident assignment of glycan structures
- Quick: 1 hour
- High throughput: 96 well format

Figure 1. Flow chart of the sample preparation, LC column, and instruments used for identification and quantification of InstantPC labeled N-glycans.

Glycan Composition detected from therapeutic mAb

ID	Observed [M+2H]2+	Proposed composition	Theoretical [M+2H]2+	Proposed structure	
				Oxford	Alternative
1	639.7	F1+Core	639.774	F1M3	G0F-2GlcNAc
2	688.4	N1+Core	688.284	A1	G0-GlcNAc
3	748.9	H2+Core	748.798	M5	Man5
4	761.5	N1F1+Core	761.313	FA1	G0F-GlcNAc
5	790.9	N2+Core	789.824	A2	G0
6	842.4	H1N1F1+Core	842.340	FA1G1	G1F-GlcNAc
7	864.0	N2F1+Core	862.853	FA2	G0F
8	871.9	H1N2+Core	870.851	A2G1G1	G1
9	871.9	H1N2+Core	870.851	A2G2S1	G1'
10	923.6	H2N1F1+Core	923.366		
11	944.0	H1N2F1+Core	943.879	FA2G1G1	G1F
12	944.0	H1N2F1+Core	943.879	FA2G2S1	G1F'
13	952.0	H2N2+Core	951.877	A2G2	G2
14	964.7	N2F1+Core	964.350	FA2B	G0FB
15	991.9	H5+Core	991.877	M6	Man6
16	996.1	H1N1S1F1+Core	995.885	FA1G1S1	G1S1F-GlcNAc
17	1025.1	H2N2F1+Core	1024.906	FA2G2	G2F
18	1025.1	H2N2F1+Core	1024.906	FA2G1Ae1	G1Fa1ae
19	1045.6	H1N2F1+Core	1045.419	FA2BG1S1	G1FB
20	1045.7	H1N2F1+Core	1045.419	FA2B2G1S1	G1FB'
21	1089.3	H1N1S1F2+Core	1089.427	FA1G1S1	G1S1F-GlcNAc
22	1097.7	H1N2F1S1+Core	1097.425	FA2G1S1g1	G1Sg1F
23	1106.1	H2N2F1+Core	1105.932	FA2G2Ae1	G2Fa1ae
24	1106.1	H2N2F1+Core	1105.932	FA2G2Ae1 iso	G2Fa1ae'
25	1126.5	H2N2F1+Core	1126.446	FA2BG2	G2FB
26	1170.4	H2N2F1S1+Core	1170.454	FA2G2S1	A1F
27	1170.4	H2N2F1S1+Core	1170.454	FA2G2S1 iso	A1F iso
28	1178.7	H2N2F1Sg1+Core	1178.451	FA2G2Sg1	Ag1F
29	1187.1	H4N2F1+Core	1186.959	FA2G2Ae2	G2F+2a6Ae2
30	1207.7	H3N2F1+Core	1207.472	FA2BG2Ae1	G2FB+eAe2
31	1209.4	H3N2F1Sg1+Core	1209.477	FA2G2Sg1Ae1	Ag1F+eAe2
32	1216.2	H2N2F1S2+Core	1216.501	FA2G2S2	A2F

Source: Agilent Application Note#5991-8071EN



Source: EGMIC internal

LC-MS Mass Spectrometry

<< Agilent LC/Q-TOF Ion Mobility 6560 IMS >>

- With ECD, CID, mass range of m/z 5-3,200
- The gold standard in direct collision cross section measurement and accuracy
- Intact proteins, native and denatured, calculate molecular size directly
-
- Separation of complex isobaric classes such as peptides, lipids and glycans. Used to resolve structural isomers, peptide cross-links, post-translational modifications and characterize protein shape and protein-protein interactions.

<< Agilent AdvancedBio LC/Q-TOF 6545XT >>

- High Mass Capable Q-TOF, mass range of m/z 50-30,000
- Unique ability for the top-down sequencing of proteins and synthetic peptides
- Unique fragmentation modes provides characterization of post translational modifications of peptides and structural characterization of lipids.
- Used for metabolomics, lipidomics, and qualitative flux analysis, protein-interaction mapping.

<< ThermoFisher Scientific TSQ ALTIS >>

- Triple-stage quadrupole mass spectrometer
- High sensitivity to quantify all types of molecules, such as metabolites, illicit drug, contaminants, at ultra-low levels in complex biological matrices.
- Extra selectivity for peptide quantification with Highly Selected Reaction Monitoring (H-SRM) capability.
- Used for Pharma, Biopharmaceutical, clinical research, forensic toxicology testing, food and environmental safety testing.

Molecular Interactions

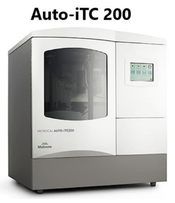
<< SPR: Surface Plasmon Resonance >>

Biacore X100

Cytiva

<< ITC: Isothermal Titration Calorimeter >>

Auto-iTC 200



Malvern Panalytical

<< Functional Glycan Microarray >>

InnoScan AL1100

Innopsys

Why us?

✓ Extensive support on project planning and technology consultation.

✓ Expert support on training for instrument operation, assay development and data interpretation.

✓ Meticulously maintained instruments.

✗ No confusion - we are here to help!

✗ No surprises - full transparency on costs!

✗ No delays - reserve and run today!