Quality control of synthetic biomolecules using rapid methods with serial coupling of UV and MS detectors

Joachim Weiss[1], Sylvia Grosse[2], Katherine Lovejoy[2], Frank Steiner[2]; Thermo Fisher Scientific, [1]Idstein, Germany [2]Germering, Germany

ABSTRACT

Quality control is an important step during the production of synthetic biomolecules, both in peptide and oligonucleotide synthesis. HPLC-UV is commonly used to determine the purity of the API product. This technique is simple and efficient when the API is known. However, impurity analysis of the synthetic biomolecules may be quite challenging due to a wide variety of possible related impurities, quite often present at low concentration.

Peptide synthesis distinguishes between process and product related impurities. In particular, the determination of the product related impurities is difficult as they are mostly not known. Therefore, mass spectrometry is needed for reliable identification and confirmation. Oligonucleotide synthesis requires confident confirmation of oligonucleotide mass, as well as rough quantification of yield and impurity levels. Quantification of yield can easily be performed by UV. Rough estimation of impurities requires a mass spectrometer, as aborted sequences (N-1) are not usually chromatographically separated from complete sequences (N) during a quick QC method. Mass confirmation of oligo identity also requires a mass spectrometer.

Mass spectrometry is often considered to be too complex and too difficult to use in routine quality control applications. The Thermo Scientific™ ISQ™ EM single quadrupole mass spectrometer is developed for operation by chromatographers. Its full integration into the Thermo Scientific™ Chromeleon™ 7.2 chromatography data system (CDS) and the Thermo Scientific™ AutoSpray smart method set-up make LC-MS operation and data analysis straightforward and intuitive. The ISQ EM mass spectrometer has an extended mass range from 50 to 2000 m/z, allowing chromatographers to collect data over many charge states. The orthogonal source design provides high levels of instrument robustness, even for the challenging conditions posed by the ion pairing eluents used for oligonucleotides. In the current work, peptide impurity profiling and mass-based compound confirmation were demonstrated for the antimicrobial human LL-37 peptide, a compound of high medical importance due to its antibacterial, antimycotic, antiviral, wound healing, anticancer and immunomodulatory activity [1]. The sequence consists of 37 amino acids with a peptide mass of about 4500 Da. Also in this work, a quality control method for DNA oligomers of 29, 31, 37, and 40 base pairs in length was augmented with MS to provide mass-based compound confirmation and impurity profiling.

MATERIALS AND METHODS

Instrumentation

A Thermo Scientific™ Vanquish™ Flex Binary UHPLC system equipped with an ISQ EM single quadrupole mass spectrometer was used for the analysis.

- Thermo Scientific™ Vanquish™ System Base Vanquish Horizon/Flex (P/N VH-S01-A-02)
- Thermo Scientific™ Vanquish™ Binary Pump F (P/N VF-P10-A-01)
- Thermo Scientific™ Vanquish™ Split Sampler FT (P/N VF-A10-A-02)
- Thermo Scientific™ Vanquish™ Column Compartment H (P/N VH-C10-A-02)
- Thermo Scientific™ Vanquish™ Variable Wavelength Detector F (P/N VF-D40-A)
- Flow Cell Semi-Micro, 2.5 µL, 7 mm light path (SST) (P/N 6077.0360)
- ISQ EM Mass Spectrometer (P/N ISQEM-ESI)

Thermo Scientific™ Chromeleon™ 7.2.9 Chromatography Data System (CDS) was used for data acquisition and processing.

Table 1. Chromatographic Conditions

Parameter	Peptide Application		DNA Application	
Column	Thermo Scientific™ Acclaim™ RSLC 120 C18, 50x2.1 mm, 2.2 µm		Thermo Scientific™ DNAPac™ RP 2.1 x 10 mm, 4 µm (P/N 088925)	
Mobile Phase	A: water + 0.1% formic acid B: acetonitrile + 0.1% formic acid		A: 200 mM HFIP, 8.0 mM TEA, pH 8.0 B: methanol	
Flow rate	0.5 mL/min		0.5 mL/min	
Gradient	Time [min]	%B	Time [min]	%B
	0	20	0	15
	2	50	0.5	15
	2.1	20	1	60
	5.5	20	1.3	60
			1.4	15
			6	15
Mixer vol.	10 + 25 µL		10 + 25 µL	
Column T	50 °C (forced air mode, fan speed 5, active preheater)		50 °C (forced air mode, fan speed 5, active preheater)	
Sampler T	4 °C		4 °C	
UV	l = 214 nm, data collection rate = 10 Hz, response time = 0.5 s		l = 260 nm, data collection rate = 10 Hz, response time = 0.5 s	
Injection vol.	1 µL		1 µL	

Table 2. Mass Spectrometry Settings

Source Parameters	Both Applications HESI source, easy ESI settings for 0.5 mL/min
Sheath gas pressure	49.9 psig
Aux gas pressure	5.7 psig
Sweep gas pressure	0.5 psig
Vaporizer temperature	282 °C
Ion transfer tube temperature	300 °C
Source voltage	3000 V for peptides, -2000 for DNA
Method Parameters	Peptide Application Setting / DNA Application Setting
Method type	Full Scan / Full Scan
Ion polarity	positive / Negative
Mass range	m/z 500-2000 / m/z 600-2000
Dwell Time	0.2 s / 0.2 s
Source CID voltage	10 V / 15 V

RESULTS

Peptide Analysis

The peptide amino acids with sequences given in Table 3, reconstituted at 1 mg/mL in mobile phase A, were combined into a single sample with 150 µg/mL fragment 1 (LL-37 RKS), 50 µg/mL fragment 2 (LL-37 SKE) and 500 µg/mL of the active pharmaceutical agent (LL-37). Figure 1 shows the UV and TIC chromatogram used to confirm peak purity for LL-37. The spectrum obtained shows no significant additional mass peaks and therefore no indication of co-elution of the API with other impurities (Figure 4). Figure 2 shows the custom columns in the Chromeleon CDS sequence list that were used to calculate the theoretical mass of peptides by entering the peptide sequence or chemical formula. Charge states and adduct masses are calculated automatically to assist peak identification. Figure 3 shows the simulated product impurity profiling, for which two peptide fragments are spiked into LL-37 and are baseline separated with a short gradient. Peak assignment, not possible by UV, is enabled by the ISQ EM mass spectrometer. Figure 4 shows the use of the extended mass range of the ISQ EM mass spectrometer to detect charge states between 2 and 8. As is standard practice for the inherently lower resolution that characterizes single quadrupole instruments, average mass is used instead of monoisotopic mass for the high charge states.

Table 4 summarizes the measured and theoretical m/z values (averaged and most abundant isotope) of all charge states obtained for the LL-37 peptide and the two fragments. The mass deviation between the averaged theoretical and the measured mass (Delta m/z) was found to be ≤ 0.4 for all charge states. The uncharged averaged mass was calculated for each charge state using the following equation:

$$M = z(m/z - H^+)$$

and averaged, resulting in the intact deconvoluted average mass. The mass deviation for the intact molecule was ≤ 0.5 Da for all peptides.

UV is preferred for quantitative impurity analysis because the MS signal depends on analyte ionization efficiency and is sometimes affected by ion suppression. MS is sometimes used, especially with isotopically labeled standards, to quantify at concentrations below the UV limit of detection. Table 5 shows a comparison of relative peak areas for the UV-, TIC- and XIC-traces, based on ICH guidelines [2], and shows how challenging the quantitation of product related impurities can be. Due to the different sequences, the UV signal response differs from the theoretically calculated values (30% for fragment 1, 10% for fragment 2). For accurate UV quantification would require external calibration for each analyte or determination of response factors.

Table 3. Peptide amino acid sequences and theoretical molecular mass

Compound	Amino acid sequence	Theoretical average mass [Da]	Theoretical monoisotopic mass [Da]	Theoretical most abundant isotope mass [Da]
LL-37 (API)	LLGDFFRKSKEKIGKEFKRIVQRIKDFLRNLVPRTES	4493.299	4490.565	4492.582
LL-37 RKS (fragment 1)	RKSKEKIGKEFKRIVQRIKDFLRNLVPRTES	3800.488	3798.211	3799.225
LL-37 SKE (fragment 2)	SKEKIGKEFKRIVQRIKDFLR	2619.130	2617.544	2618.558

Figure 1. a) UV chromatogram at 214 nm and b) TIC chromatogram of the LL-37 peptide (*m/z 500 – 2000)* with zoom into the baseline

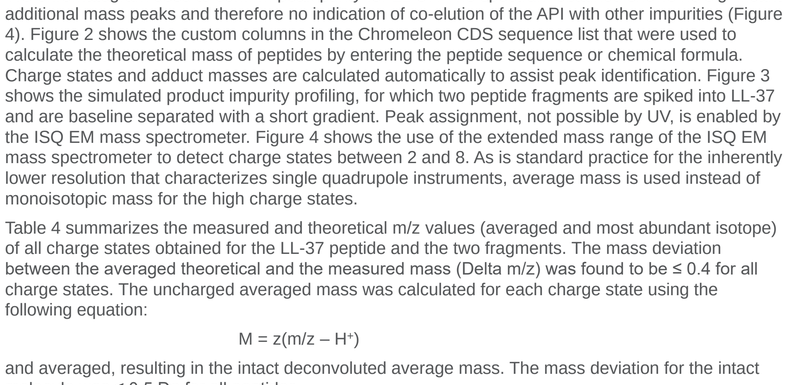
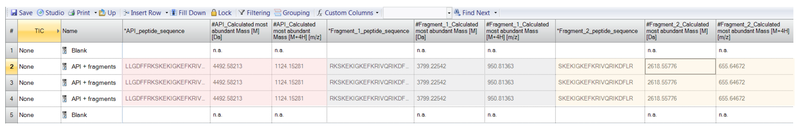


Figure 2. Screenshot of the Chromeleon CDS injection list with added custom columns for theoretical mass calculations of the target components (light blue: API; light green: fragment 1; light red: fragment 2)



#	TIC	Name	*API_peptide_sequence	#API_Calculated most abundant Mass [M] [Da]	#API_Calculated most abundant Mass [M+8H] [m/z]	*Fragment_1_peptide_sequence	#Fragment_1_Calculated most abundant Mass [M] [Da]	#Fragment_1_Calculated most abundant Mass [M+4H] [m/z]	*Fragment_2_peptide_sequence	#Fragment_2_Calculated most abundant Mass [M] [Da]	#Fragment_2_Calculated most abundant Mass [M+4H] [m/z]
1	None	Blank		n.a.	n.a.		n.a.	n.a.		n.a.	n.a.
2	None	API + fragments	LLGDFFRKSKEKIGKEFKRIV...	4492.58213	1124.15281	RKSKEKIGKEFKRIVQRIKDF...	3799.22542	950.81363	SKEKIGKEFKRIVQRIKDFLR	2618.55776	655.64672
3	None	API + fragments	LLGDFFRKSKEKIGKEFKRIV...	4492.58213	1124.15281	RKSKEKIGKEFKRIVQRIKDF...	3799.22542	950.81363	SKEKIGKEFKRIVQRIKDFLR	2618.55776	655.64672
4	None	API + fragments	LLGDFFRKSKEKIGKEFKRIV...	4492.58213	1124.15281	RKSKEKIGKEFKRIVQRIKDF...	3799.22542	950.81363	SKEKIGKEFKRIVQRIKDFLR	2618.55776	655.64672
5	None	Blank		n.a.	n.a.		n.a.	n.a.		n.a.	n.a.

Figure 3. UV chromatogram at 214 nm of simulated peptide impurity profiling

Figure 4. Obtained mass spectra for the API and the two fragments

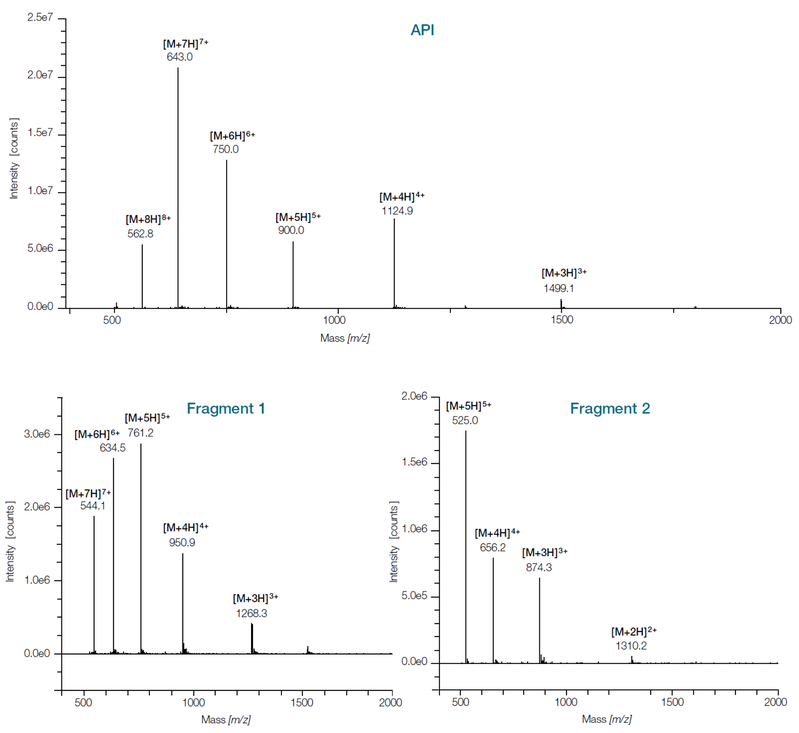


Table 4. Data for RT, Resolution and Rel. Area % of simulated impurities comparing UV, TIC and XIC trace with m/z 525.0 for LL-37 SKE (fragment 2), m/z 634.5 for LL-37 RKS (fragment 1) and m/z 643.0 for LL-37 (API)

Compound	RT [min]	Resolution	UV trace Rel. Area [%]	TIC trace Rel. Area [%]	XIC trace Rel. Area [%]
LL-37 SKE (fragment 2)	0.9	3.54	4.4	4.3	5.9
LL-37 RKS (fragment 1)	1.2	4.16	16.5	21.9	25.1
LL-37 (API)	1.7	—	100	100	100

Table 5. Comparison of theoretical and measured masses with corresponding absolute mass deviation of all assigned charge states in each peptide spectrum for Figure 4.

Charge state	Measured average mass [*m/z*]	Measured most abundant mass [*m/z*]	Theoretical average mass [*m/z*]	Delta [*m/z*]	Calculated average mass [Da]
LL-37 (API)					
[M]			4493.3		
$[M+1H]^{1+}$	—	—	4494.3	—	—
$[M+2H]^{2+}$	—	—	2247.7	—	—
$[M+3H]^{3+}$	1498.4	1499.1	1498.8	0.4	4492.2
$[M+4H]^{4+}$	1124.2	1124.9	1124.3	0.1	4492.8
$[M+5H]^{5+}$	899.5	900.0	899.7	0.2	4492.5
$[M+6H]^{6+}$	750.0	750.0	749.9	-0.1	4494.0
$[M+7H]^{7+}$	643.0	643.0	642.9	-0.1	4493.9
$[M+8H]^{8+}$	562.8	562.8	562.7	-0.1	4494.3
Deconvoluted average mass (n=6) [Da]					4493.3
Mass deviation to theoretical mass [Da]					-0.02
LL-37 RKS (Fragment 1)					
[M]			3800.5		
$[M+1H]^{1+}$	—	—	3801.5	—	—
$[M+2H]^{2+}$	—	—	1901.3	—	—
$[M+3H]^{3+}$	1267.5	1268.3	1267.8	0.3	3799.5
$[M+4H]^{4+}$	950.9	950.9	951.1	0.2	3799.6
$[M+5H]^{5+}$	761.2	761.2	761.1	-0.1	3801.0
$[M+6H]^{6+}$	634.5	634.5	634.4	-0.1	3801.0
$[M+7H]^{7+}$	544.1	544.1	543.9	-0.2	3801.6
$[M+8H]^{8+}$	—	—	476.1	—	—
Deconvoluted average mass (n=5) [Da]					3800.5
Mass deviation to theoretical mass [Da]					0.04
LL-37 SKE (Fragment 2)					
[M]			2619.1		
$[M+1H]^{1+}$	—	—	2620.1	-	-
$[M+2H]^{2+}$	1310.2	1310.2	1310.6	0.4	2618.4
$[M+3H]^{3+}$	873.7	874.3	874.1	0.4	2618.1
$[M+4H]^{4+}$	655.5	656.2	655.8	0.3	2618.0
$[M+5H]^{5+}$	525.0	525.0	524.8	-0.2	2620.0
$[M+6H]^{6+}$	—	—	437.5	—	—
$[M+7H]^{7+}$	—	—	375.2	—	—
$[M+8H]^{8+}$	—	—	328.4	—	—
Deconvoluted average mass (n=4) [Da]					2618.6
Mass deviation to theoretical mass [Da]					-0.5

DNA Oligomer Analysis

Oligonucleotide with sequences shown in Table 6, provided by Thermo Fisher Scientific, (Pleasanton, CA) were analyzed by reverse phase ion pairing chromatography using a short quality control method. The source was tuned once at the beginning of experiments using the operational qualification wizard in Chromeleon. CID voltage was optimized by maximizing the peak area associated with the most abundant mass for the charge state, but an alternative would be to choose a CID that leads to minimal depurination (see Table 9 and Figure 5D).

The concentration of the 37mer oligonucleotide was determined by a UV calibration curve over the values 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10, 20 and 50 µM with a coefficient of determination of 0.9996, a y-intercept of 0.099 and a slope of 0.539.

Table 7 shows the mass confirmation experiments for three injections of 5 pmol of each of the four oligomers. Mass accuracy is excellent over various oligonucleotide lengths at a relatively low mass on column. The observed masses are all within the mass accuracy specification of < + 0.1 Da of the ISQ EM mass spectrometer. The repeatability of the observations is excellent, with an RSD of 0.02% or less.

For mass confirmation of all charge states with m/z 600 – 2000, injections of 50 pmol of the 37mer oligonucleotide were made in quadruplicate. The charge states for the 37mer are shown in Table 8. The average molecular weight of each oligonucleotide was used to calculate the expected mass of each charge state. The observed masses, except for the -13 charge state, are within the mass accuracy specification. The repeatability of the observations is excellent, with an RSD of 0.02% or less except for the -14 charge state, which had an RSD of 0.07%. The increased instrument error associated with higher charge states, explains the slightly worse results for charge states 13, 14 and 15.

Based on peak area, rough amounts of adducts and impurities relative to the sum of the areas of all adducts identified for the -9 charge state are shown in Table 9. Potential impurities and adducts were identified using peer reviewed literature sources [3,4]. Extracted ion chromatograms (XICs) for the ninth charge state of the 37mer and some impurities are shown in Figure 2.

When using ion pairing reagents in MS, longer equilibration time improves peak area reproducibility. As shown in Figure 6, the RSD for the peak area of the full scan of six 1 µL injections (5 µM 37mer sample) improved from 15.4% to 8.1% when the equilibration time was increased from 1 to 4.5 minutes.

Table 6. Oligomer samples analyzed

Length	Sequence	Average Mass (Da)
40	CTCTCTGACACAATTAAGGGATAAAATCTCTGACGGAATG	12,306.1
37	CAGGAAACAGCTATGACCCGCGCTCACCTCGCCTCTG	11,234.9
31	ATGATATTATGATTAGGAGCCGCGCAGGGAG	9659.6
29	TGAAGGAITGCACTGAAAGGCAGGCTAAT	9024.9

Table 7. Mass confirmation for 29mer, 31mer, 37mer and 40mer using the relatively high-abundance $[M-9H]^{9-}$ charge state. The mass on column was 5 pnol

Length	Average Mass (Da)	Average Mass, $[M-9H]^{9-}$	Found Mass, Trial 1, $[M-9H]^{9-}$	Found Mass, Trial 2, $[M-9H]^{9-}$	Found Mass, Trial 3, $[M-9H]^{9-}$	Mean (n=3) Observed	Mass Accuracy	Std. Dev. (n = 3)	RSD
40	12312.0	1367.0	1367.1	1366.9	1366.9	1367.0	0.0	0.12	0.01%
37	11240.3	1247.9	1247.9	1247.8	1247.8	1247.8	-0.1	0.06	0.00%
31	9664.3	1072.8	1072.9	1073.0	1072.8	1072.9	0.1	0.10	0.01%
29	9024.9	1001.8	1001.6	1001.7	1002.0	1001.8	0.0	0.21	0.02%

Table 8. Charge states calculated and found for 37mer (TIC, scan m/z 600 – 2000, CID 15, 0.2 s dwell time, negative mode, 50 pmol on-column). Found values and standard deviations (Std. Dev.) are from the average of four injections.

	Expected and Found Masses for Charge States $[M-xH]^{x-}$									
	6	7	8	9	10	11	12	13	14	15
Expected	1872.4	1604.8	1404.0	1247.9	1123.0	1020.9	935.7	863.6	801.9	748.4
Found	1872.4	1604.8	1403.8	1247.9	1122.9	1020.7	935.7	864.0	802.1	748.6
Mass Accuracy	-0.0	-0.0	-0.2	-0.0	-0.1	-0.2	-0.0	+0.4	+0.2	+0.2
Std. Dev. (n = 4)	0.16	0.20	0.15	0.10	0.15	0.10	0.10	0.21	0.56	0.17
RSD	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.07%	0.02%

Table 9. Identification and rough estimation of impurity. Percent area is relative to the sum of the area of all adducts identified for the -9 charge state of the 37mer oligonucleotide.

Ion and Masses (m/z) Found	Percent Area (standard deviation, n = 3)	Characterization
$[M-9H]^{9-}$ 1247.9	21.1% (1.2%)	base peak
$[M-9H+xHFIP]^{9-}$, (x = 1-5) 1266.6, 1285.3, 1303.9, 1322.6, 1341.3	39.3% (3.9%)	solvent adduct
$[M-9H+xMeOH]^{9-}$, (x = 1-2) 1251.5, 1255.0	2.2% (1.7%)	solvent adduct
$[M-10H+TEA]^{9-}$ 1259.2	0.7% (0.6%)	solvent adduct
$[M+xK^+-yH+zHFIP]^{9-}$, (x = 1 – 5, y = 10 – 14, z = 0 – 2) 1252.2, 1270.8, 1289.5, 1308.2, 1326.8, 1255.7, 1256.4, 1275.1, 1293.7, 1260.6, 1264.9, 1269.1	14.4% (3.4%)	potassium counterion
$[M+Na^+-10H]^{9-}$ 1250.4	2.0% (0.9%)	sodium counterion
depurination products 1231.1, 1249.8, 1234.7, 1235.4, 1232.9, 1251.6, 1236.5, 1237.2	8.5% (2.1%)	impurity or possible MS-related decomposition
N-1 and adducts 1211.3, 1230.0, 1214.9, 1215.6, 1220.2	6.1% (2.5%)	impurity, minor
2-cyanoethyl impurity, 1253.5	2.1% (0.4%)	impurity, trace
isobutyryl impurity, 1255.4	0.8% (0.2%)	impurity, trace
chloral N-mer impurity, 1263.8	0.6% (0.3%)	impurity, trace
4,4-dimethoxytrityl N-mer impurity, 1281.2	0.5% (0.3%)	impurity, probably not present

Figure 5. Extracted ion chromatograms for the $[M-9H]^{9-}$ charge state of the 37mer and some of the identified impurities using a full scan in negative mode from m/z 600 – 2000, a dwell time of 0.2 s, a CID of 15 and smoothing using the Gaussian function and setting of 11 points in Chromeleon. A) $[M-9H]^{9-}$, m/z 1247.9. B) the -9 charge state for the aborted sequence, N-1, m/z 1211.3. C) $[M+K^+-10H]^{9-}$, m/z 1252.2. D) the -9 charge state for the depurination resulting in loss of one adenine base, m/z 1232.9. E) the -9 charge state for the 2-cyanoethyl impurity, m/z 1253.5.

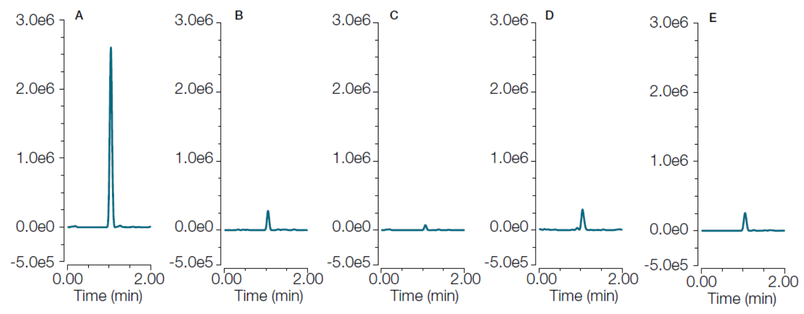


Figure 6. Overlay of six injections with long (left side) and short (right side) equilibration times. The RSD for peak area is 8.1% for the long equilibration time and 15.4% for the short equilibration time. The trace shown is the full scan (m/z 600-2000, 0.2 s dwell time, CID 15 V, negative mode).

CONCLUSIONS

- A rapid gradient method allowed the separation of the LL-37 peptide impurity profiling within 2 min.
- Adding MS to existing UV workflows for quality control provides peak purity, mass-based identity confirmation and impurity identification.
- The extended mass range up to m/z 2000 enables the detection of low-charge peptide species and allows the detection of the complete charge state profiles of mid-sized biomolecules. Ten charge states are detectable for a 37mer oligomer.
- The orthogonal source design did not require cleaning over 100 injections of “sticky” HFIP and TEA reagents. A long equilibration time improved peak area reproducibility when using these ion pairing reagents

REFERENCES

1. Yufei Chen, Sidi Yang, and Emmanuel A. Ho, Development of an Analytical Method for the Rapid Quantitation of Peptides Used in Microbicide Formulations (**2014**), *Chromatographia*, 77(23-24), 1713–1720
2. ICH Guidelines, Quality Guidelines, Q3A-Q3D Impurities
3. Apffel, A.; Chakel, J.A.; Fischer, S.; Lichtenwalter, K.; Hancock, W.S. Analysis of oligonucleotides by HPLC-electrospray ionization mass spectrometry. Anal. Chem. 1997, 69, 1320-1325.
4. Nikcevic, I.; Wyrzykiewicz, T.K.; Limbach, P.A. Detecting low-level synthesis impurities in modified phosphorothioate oligonucleotides using liquid chromatography—high resolution mass spectrometry. Int. J. Mass Spectrom. 2011, 304, 98-104.

TRADEMARKS/LICENSING

PO72846-EN 0918S