Published in final edited form as:

J Pharm Sci. 2017 November ; 106(11): 3242–3256. doi:10.1016/j.xphs.2017.07.012.

The botanical drug substance crofelemer as a model system for comparative characterization of complex mixture drugs

Peter A. Kleindl^{1,‡}, Jian Xiong^{1,2,‡}, Asha Hewarathna¹, Olivier Mozziconacci¹, Maulik K. Nariya³, Adam C. Fisher⁴, Eric J. Deeds^{3,5,6}, Sangeeta B. Joshi^{1,2}, C. Russell Middaugh^{1,2}, Christian Schöneich¹, David B. Volkin^{1,2}, and M. Laird Forrest^{1,*}

¹Department of Pharmaceutical Chemistry, University of Kansas, Lawrence, Kansas, USA, 66047

²Macromolecule and Vaccine Stabilization Center, University of Kansas, Lawrence, Kansas, USA, 66047

³Center for Computational Biology, University of Kansas, Lawrence, Kansas, USA, 66047

⁴Center for Drug Evaluation and Research, Office of Pharmaceutical Quality, U.S. Food and Drug Administration. The views expressed in this publication by the authors do not necessarily reflect the views or policies of the Food and Drug Administration (FDA)

⁵Department of Molecular Biosciences, University of Kansas, Lawrence, Kansas, USA, 66047

⁶Santa Fe Institute, Santa Fe, New Mexico, USA, 87501

Abstract

Crofelemer is a botanical polymeric proanthocyanidin that inhibits chloride channel activity and is used clinically for treating HIV-associated secretory diarrhea. Crofelemer lots may exhibit significant physicochemical variation due to the natural source of the raw material. A variety of physical, chemical, and biological assays were utilized to identify potential critical quality attributes (CQAs) of crofelemer, which may be useful in characterizing differently sourced and/or processed drug products. Crofelemer drug substance was extracted from tablets of one commercial drug product lot, fractionated, and subjected to accelerated thermal degradation studies to produce derivative lots with variations in chemical and physical composition potentially representative of manufacturing and raw material variation. Liquid chromatography, UV absorbance spectroscopy, mass spectrometry, and NMR analysis revealed substantial changes in the composition of derivative lots. A chloride channel inhibition cell-based bioassay suggested that substantial changes in crofelemer composition did not necessarily result in major changes to bioactivity. In two companion papers, machine learning and data mining approaches were applied to the analytical and biological data sets presented herein, along with chemical stability data sets derived from forced degradation studies, to develop an integrated mathematical model that can identify CQAs which are most relevant in distinguishing between different populations of crofelemer.

*Correspondence to Author: M. Laird Forrest: lforrest@ku.edu; phone, (785) 864 4388.

‡PAK and JX contributes equally to this work

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

polymeric drugs; polymer chemical degradation; UV-visible spectroscopy; FTIR; mass spectrometry; NMR spectroscopy; circular dichroism; intestinal secretion/transport; complex mixture; biopharmaceuticals; analytical characterization; formulation; stability

Introduction

Crofelemer is a botanical drug substance currently approved by the U.S. Food and Drug Administration (FDA) for the treatment of noninfectious secretory diarrhea in HIV/AIDS patients undergoing antiretroviral therapies. Crofelemer is a polymeric proanthocyanidin isolated from the red latex of the South American tree *Croton lechleri*. Crofelemer's chemical structure consists primarily of a random, linear sequence of (+)-catechin, (+)-gallocatechin, (–)-epicatechin, and (–)-epigallocatechin units. Polymer chains contain 1 to 28 repeating units and a number average of 5 to 7.5 units (Figure 1). Hence, there is the potential for significant drug substance variation arising from the processing and purification of the raw material and the botanical nature of the crude source material itself, e.g. growth conditions and specimen biodiversity. This inherent heterogeneity makes rigorous characterization of crofelemer a difficult task.

Accurate and sufficient analytical characterization of therapeutic complex molecules such as crofelemer, as well as biopharmaceuticals such as mAbs, is a crucial part of ensuring product quality, safety, and efficacy between product batches during scale-up and manufacturing changes (i.e., comparability)^{1,2} as well as between drug product lots across different manufacturers (i.e., pharmaceutical equivalence or similarity).^{3,4} Structural heterogeneity within individual batches (intra-lot) and between separate batches (inter-lot) must be characterized and evaluated against clinical batches deemed to be safe and effective. Compared to traditional homogeneous small molecule pharmaceuticals, there is a typically a larger number of analytical techniques required to characterize different attributes of heterogeneous drug products such as botanicals and biotherapeutics. Further, the state of product knowledge and analytical technology may preclude comprehensive characterization of every component of a mixture. To streamline quality assurance, critical quality attributes (CQAs) are identified during drug product development. CQAs are specific structural and biological aspects of complex molecules that provide the most pertinent information for ensuring batch consistency and acceptable clinical outcomes.⁵ CQAs are product specific and variations in these attributes must be related to the inherent heterogeneity and clinical performance of these complex drugs. Selecting relevant CQAs is a difficult but necessary step in the development of complex drug products and biologics. Table 1 summarizes the analytical methods applied in this work for crofelemer characterization.

The FDA has issued a draft guidance on the development of botanical drugs that describes a “totality of evidence” approach that takes into account multiple lines of evidence to support therapeutic consistency.⁶ Analytical characterization is an important element in botanical drug product development. Comparative characterization is necessary to support any post-approval changes to the manufacturing process (e.g., process parameters, raw materials). This works to ensure there is no difference in safety and efficacy after a process change. To

date, two botanical drugs have been approved by the FDA, crofelemer, approved in 2012, and sinecatechins (Veregen®), an ointment approved in 2006.

This article is the first in a series of three papers in this issue that uses crofelemer as a model complex mixture drug to develop integrated mathematical models for comparative characterization using a machine learning approach. This approach has recently been demonstrated in proof-of-concept studies from various analytical data sets from four different well-defined IgG1-Fc glycoforms.^{14–17} Multiple lots of a drug product are usually necessary to determine the critical quality tests and their acceptable limits, i.e. the critical quality attributes (CQA). However, early in product development few lots may be available, especially from botanical drugs that may be dependent on harvest size and season. In this study, only a single in-date lot of crofelemer was available from the manufacturer at the initiation of the study. Thus, “virtual lots” of crofelemer were produced from a single lot through a combination of dialysis, to simulate the batch to batch variability in molecular weight and chemical composition, and thermal treatment, to simulate chemical and physical degradation of the drug product. These samples are “like,” in that all had similar activity by the biological assay, but subtle variations from unadulterated crofelemer based on various chemical and physical tests.

Machine learning can be a powerful tool to identify structural signatures and physicochemical properties that clearly distinguish between samples with only subtle variations. These methods could be especially valuable in the development of biosimilar pharmaceuticals, where CQA of the originator are unknown or proprietary. After generating a large analytical dataset from a single lot of crofelemer and multiple artificially degraded sublots, data mining and machine learning approaches were applied to identify assays and/or chemical and physical features (i.e., potential CQAs) of the crofelemer biopolymer that are most relevant in distinguishing between different crofelemer samples. The techniques used and initial results are described in this first article. The analytical results from a variety of forced chemical degradation studies of crofelemer are presented in the second companion article.¹⁸ In the third article¹⁹, a machine learning approach is applied to the entirety of this data set to identify which individual and/or combinations of assays which can distinguish between subtle variations in highly similar samples of crofelemer.

Materials and Methods

API Extraction

Crofelemer was extracted from commercially available FulyzaqTM tablets (Salix Pharmaceuticals, Lot# 3117608. Fulyzaq was recently acquired by Napo Pharmaceuticals and rebranded as MytesiTM). These delayed-released tablets are white, oval enteric-coated and should be stored at 20–25 °C. The coating of each tablet was physically removed using a clean razor blade and tablets were then crushed and dissolved in ultrapure water. The solution was centrifuged at $4,500 \times g$ at 4°C for 20 minutes and the reddish-brown colored supernatant was collected as the unfractionated crofelemer extract with an approximate concentration of 25 mg/mL, based on expected mass extract.

Membrane Fractionation of Crofelemer Extract

The crofelemer extract was size fractionated using Amicon Ultra centrifugal filters (Millipore, Billerica, MA) of 3 and 10 kDa MWCO by centrifuging at 4,000 ×g and 4°C for 30 minutes. Red colored fractions were retained in the insert (3 kDa Top and 10 kDa Top) and diluted with ultrapure water to recover the original concentration. Colorless fractions were collected as the filtrate (3 kDa Bottom and 10 kDa Bottom) and no further dilutions were performed (Figure 2).

Unfractionated Sample

The unfractionated samples were prepared from crofelemer stock solution. To conserve stocks and maximize the replicates, the stock solution was diluted and individually aliquoted for MS, HPLC/SEC and the cell-based *in vitro* biological assay. The individually sealed aliquots were incubated at 25 and 40°C for the accelerated stability studies.

Low MW (3 kDa MWCO) Filtered Sample

Aliquots of the 3 kDa Top fraction were made as described for the unfractionated material. Low concentration aliquots (ca. 0.5 mg/mL) were made for all assays with the exception of NMR studies. Due to the low concentration of the 3 kDa Bottom fraction, no further dilution was applied. Individual aliquots of the 3 kDa Bottom fraction were incubated for the accelerated stability study.

High MW (10 kDa MWCO) Filtered Sample

The 10 kDa Top and Bottom fractions were prepared and aliquoted in the same manner as the 3 kDa fractions, using a 10 kDa MWCO spin filter to achieve separation. Each individual vial was labeled for certain assays and then incubated at either 25 or 40°C for the accelerated stability study. Due to the large number of vials and the limited instrumental throughput, some samples were frozen and stored at –80°C after the desired time points were reached and until measurements were taken.

UV-Visible Spectroscopy

The UV-Visible absorption spectra of crofelemer were collected using an HP-8453 photodiode array spectrometer (Agilent Technologies, Santa Clara, CA) equipped with deuterium (D₂) and tungsten (W) lamps. Water background signal was recorded before performing the experiment with crofelemer samples. Crofelemer samples were diluted to concentrations within the linear range of the instrument. Spectra were collected from 190–1100 nm with 1 cm path length quartz cuvettes. The UV-Visible absorbance spectra were corrected for light scattering using a technique included in the manufacturer's data analysis software (Chemstation UV-Vis analysis software, Agilent Technologies): first, the spectra data, where the optical density values are only due to light scattering (350–400 nm), are fitted to an equation, then this curve is extrapolated across the entire sample spectrum and subtracted from the original spectrum, to produce the light-scatter corrected absorbance spectra.

Variable Path Length Spectroscopy

Variable path length absorption spectra of crofelemer were collected using a Solo VPE instrument (C Technologies, Bridgewater, NJ) coupled with a Cary 50 Bio spectrometer (Agilent Technologies, Santa Clara, CA). Spectra were collected in a path length range from 1.4 to 5.0 mm and corrected for light scattering from 350–400 nm. All data analysis was performed using Solo VPE analysis software where the extinction coefficient (ϵ) of crofelemer was determined by software algorithms.

NMR Spectroscopy

NMR samples were prepared as mentioned above except that crofelemer was dissolved into D₂O during product extraction. Crofelemer fractions were also prepared using 3 and 10 kDa Amicon Ultra-15 ml centrifugal filters. Bottom fractions, however, were of too low a concentration for NMR analysis. No dilutions were performed in the NMR samples prepared. Accelerated stability samples were sealed in clean, sterilized crimp top glass vials and stored at their respective temperatures (25 and 40°C) in a controlled environment. Due to limited instrument throughput, some samples were frozen and stored at –80°C after the desired time points were reached and remained frozen until measurements were taken.

¹H NMR spectra were collected with an Avance 400 MHz NMR (Bruker, Billerica, MA). A total of 128 scans were collected per ¹H NMR sample. Quantitative ¹³C NMR (qCNMR) spectra of crofelemer samples were collected using an Avance AVIII 500 MHz NMR spectrometer with a dual carbon/proton cryoprobe (Bruker, Billerica, MA), using a zgig pulse program, 3200 scans, with a 25 second interscan delay. Thirty day incubated samples (25 and 40°C) showed visual evidence of precipitation and were centrifuged for 15 min at 14,000 × g to exclude particulates before spectra were recorded. Spectra were manually corrected for phase and baseline using MestRe Nova 9.0 (Mestrelab Research, Santiago de Compostela, Spain) before peak integration.

Peak assignments for crofelemer in the qCNMR analysis were made based on shift values reported previously for the precursor to crofelemer, SP-303,²⁰ as well as structurally similar proanthocyanidin polymers²¹ and crofelemer monomeric units²² (Figure 3). Integration of the area under the curve (AUC) for carbon shifts of interest were used for determining certain aspects of different crofelemer samples. The internal/terminal unit ratio (IU/TU) of each sample was calculated through integration of the C-4 carbon δ 36.45 and δ 38.21 ppm signals of internal units and the δ 28.81 ppm signal of its terminal unit counterpart. The ratio of 2,3-trans and 2,3-cis moieties within samples was determined through comparison of the C-2 δ 83.07 ppm signal characteristic of the catechin and galocatechin (trans) moieties, and the C-2 δ 76.44 ppm signal characteristic of the epicatechin and epigallocatechin (cis) moieties (illustrated in blue, Figure 3). The procyanidin (PC) and prodelphinidin (PD) composition of crofelemer could be estimated due to slight variations in signal contributions to the δ 146.17 ppm and δ 132.44, 133.93 ppm signals attributed to the B-ring of repeating units (illustrated in red, Figure 3). Both catechin and galocatechin have two carbons that contribute to the δ 146.17 signal seen in the ¹³C-NMR spectra. Catechin, however, only possesses one carbon which contributes signal at around δ 133.93 ppm, while galocatechin units have both a carbon that contributes to the δ 133.93 signal as well as a carbon which has

a δ 132.44 ppm shift. This difference in signal contribution, along with the ratio of the $\sim$ 146 ppm signal ($Signal_{146ppm}$) and $\sim$ 132–134 ppm signal ($Signal_{133ppm}$) ranges, were used to determine the different contributions of both PC and PD moieties within the total crofelemer sample. When applied to the normalized function for PC and PD contributions, the percent composition of PC and PD could be calculated using the following formulas:

$$Signal_{146ppm} = 2PC + 2PD = 1 \text{ (normalized)}$$

$$PD = 0.5 - PC$$

$$Signal_{133ppm} = PC + 2PD = AUC \text{ relative to Signal at 146ppm}$$

$$\frac{Signal_{133ppm}}{Signal_{146ppm}} = \frac{PC + 2(0.5 - PC)}{2PC + 2(0.5 - PC)} = \frac{AUC \text{ relative to Signal at 146ppm}}{1 \text{ (normalized)}}$$

$$PC = 1 - AUC \text{ relative to Signal at 146ppm}$$

Circular Dichroism (CD) Spectroscopy

Circular dichroism spectra were collected with a Chirascan-plus spectrometer (Applied Photophysics, Beverly, MA) equipped with a 4-cuvette position Peltier temperature controller (Quantum Northwest, Liberty Lake, WA). Crofelemer was scanned from 350–200 nm in 1 nm steps with 0.5 sec accumulation in a 1 mm cuvette. A total of four scans were averaged for each result. The water background signal was subtracted from all measurements, and Matlab (MathWorks, Natick, MA) was used for data processing.

Fourier-Transform Infrared Spectroscopy

FTIR spectroscopy analysis of crofelemer was performed using a Tensor-27 FTIR spectrometer (Bruker, Billerica, MA). The system was cooled with liquid N₂ for 20 minutes prior to use and the interferometer was purged continuously with N₂ gas. FTIR Spectra were recorded from 4000 to 600 cm⁻¹ with a 4 cm⁻¹ resolution using a Bio-ATR cell and 256 scans. The background signal of water was collected and subtracted from the sample spectra. Atmospheric and baseline corrections were applied using OPUS V6.5 (Bruker, Billerica, MA) software.

Size Exclusion Chromatography (SEC)

A LC 2010C HT HPLC system (Shimadzu, Columbia, MD) equipped with a SPD-M20A diode array detector was used for SEC. Samples were injected onto a PolarGel L column (7.5 $\times$ 300 mm, Agilent, Santa Clara, CA). The column was operated with an isocratic mobile phase (94% DMF, 5% water, 1% HAc with 0.15 M LiBr) at 50°C with a flow rate at

0.5 mL/min.²³ Molecular weight standard chromatograms of DP-11 (Planta Analytica) (MW 3904), Punicalagin (MW 1084), Cinnamtannin B1 (MW 864), Epigallocatechin gallate (MW 458) and Catechin (MW 290) were collected under the same conditions, and a linear regression relationship was established using their elution volumes versus the log₁₀ of MW to estimate the average MW of crofelemer. LC solutions software was used for data collection, and Matlab (MathWorks, Natick, MA) was used for data processing.

Hydrophilic Interaction Chromatography (HILIC)

HILIC HPLC was performed with a Luna HILIC column (2 × 150 mm, Phenomenex, Torrance, CA) on a Shimadzu LC 2010C HT HPLC system with a Shimadzu SPD-M20A diode array detector.²⁴ The column was operated with mobile phase A: 90% MeCN, 10% water, 5 mM ammonia formate (AF), pH 3.2; and mobile phase B: 50% MeCN, 50% water, 5 mM AF, pH 3.2. A gradient of 100% A to 100% B over 45 minutes at 30°C with a 0.5 mL/min flow rate was used. LC solutions (Shimadzu) and Matlab (MathWorks, MA) were used for data collection and processing.

Thiolysis

Thiolysis was carried out according to a previously published protocol for the quantification of proanthocyanidins.²⁵ A 50 µL sample of 6 mg/mL unfractionated crofelemer was thiolysed in 1.2-mL low adsorption clear glass vials (Sigma Aldrich, St. Louis, MO) using 0.6-N HCl and 0.7-M β-mercaptoethanol at 95°C for 5 min, followed by vacuum concentration for 5 min. Catechin (C), epicatechin (epiC), gallocatechin (GC) and epigallocatechin (epiGC) standard solutions (50 µL of 2 mg/mL) were separately thiolysed using the same protocol to determine the epimerization efficiency of each monomer under these reaction conditions.

Reversed Phase (RP) HPLC Analysis

The C18 column (Vydac 218TP, 250 × 4.6 mm) was pre-heated to 40°C, and equilibrated with 5% mobile phase B (70% methanol, 30% water, 0.1 % formic acid and 0.5% 2-propanol) and 95% mobile phase A (0.1 % formic acid and 0.5% 2-propanol in water). The compounds were eluted at a flow rate of 0.8 mL/min. A linear gradient was used: 5–30% B from 0 to 55 min, and 30–90% B from 55–60 min. The eluates were monitored by UV detection at 280 nm.

HPLC-MS Analysis

The crofelemer samples were analyzed by a Micromass Q-ToF Premier mass spectrometer (Micromass Ltd., Manchester, U.K.) equipped with an Acquity UPLC system (Waters Corp., Milford, MA, USA). The instrument was operated in the negative ion ESI mode, and the following instrument parameters were used: capillary voltage, 2.8 kV; desolvation temperature, 250°C. Thiolysed crofelemer samples and authentic standards were injected onto a Agilent ZORBAX (+) phenyl-hexyl column (3×50 mm, 1.8 µ) using an Acquity UPLC (Waters Corporation, Milford MA) front end. The mobile phases were: A (0.1 % formic acid and 0.5% 2-propanol in water) and B (70% methanol, 30% water, 0.1 % formic acid and 0.5% 2-propanol). The material was eluted through the analytical column with a

linear gradient as follows: 15–50% B from 0 to 6.5 min, and 50–90% B from 6.5 to 9 min at a flow rate of 0.3 mL/min.

Non-thiolysed crofelemer was analyzed by the same MS method, except, the sample was eluted from a C18 column (250 × 4.6 mm, 5 μ) at a flow rate of 0.6 mL/min using the following method. The composition of mobile phase A and B was the same as above, and the compounds were eluted by a linear gradient as follows: 1–20% B from 0 to 40 min, 20–50% B from 40 to 45 min, and 50–80% B from 45 to 47 min. MassLynx software was used to acquire and interpret the data.

T84 Chloride Channel Inhibition Assay

The T84 chloride channel inhibition assays were performed using a method adapted from West and Molloy.²⁶ T84 cells were purchased from American Tissue Culture Collection (passage 55) and maintained in 1:1 DMEM/F12 culture media with L-Glutamine and 2.438 g/L sodium bicarbonate (Life Technologies, Grand Island, NY) supplemented with 5% (v/v) fetal bovine serum (Atlanta Biologicals, Inc., Atlanta, GA) in a humidified 5% CO₂ atmosphere at 37°C. Cells were initially plated on to T-75 tissue culture flasks (Corning Inc., Corning, NY) and passaged at least twice before use in fluorescence assays. Costar 3603 96-well plates (Corning Inc., Corning, NY) were seeded at a density of 2 × 10⁴ cells/well, and culture media were changed every two days until all wells reached confluency. A 10 mM solution of the Cl[−] quenched fluorescent probe, N-(ethoxycarbonylmethyl)-6-methoxyquinolinium bromide (MQAE) (Biotium, Inc., Hayward, CA) dissolved in 100 μ L of culture media was administered to each well and incubated at 37°C and 5% CO₂ in the dark for 10 hours prior to sample testing.

After MQAE incubation, culture media were removed and replaced with a Cl[−] containing buffer solution (buffer 1) containing 65 mM NaCl, 1.2 mM K₂HPO₄, 0.3 mM KH₂PO₄, 0.5 mM CaSO₄, 0.5 mM MgSO₄, 5 mM HEPES, 5 mM Dextrose, pH 7.4 (all buffer salts obtained from Fisher Scientific, Fair Lawn, NJ). Monolayers were washed twice for 10 minutes with 150 μ L/well of fresh buffer 1 on a ZD-9556-A Orbital Shaker (Madell Technology Corp, CA) at 65 rpm to establish intracellular/extracellular equilibrium Cl[−] concentrations. Monolayers were then exposed to either 100 μ M quercetin (Sigma Aldrich, St. Louis, MO) in buffer 1 as a positive control, 0.1 mg/ml crofelemer in buffer 1, or buffer 1 alone as a negative control for 5 min. After 5 min, buffer 1 solutions were removed and replaced with 0.1 mg/ml crofelemer samples, alone or in the presence of either 10 μ M forskolin or 10 μ M ionomycin (LC Laboratories, Woburn, MA) dissolved in Cl[−] free buffer (buffer 2), composed of 65 mM NaNO₃, 1.2 mM K₂HPO₄, 0.3 mM KH₂PO₄, 0.5 mM CaSO₄, 0.5 mM MgSO₄, 5 mM HEPES, 5 mM Dextrose, pH 7.4. Crofelemer samples were tested along with respective positive and negative controls. Quercetin, forskolin and ionomycin were insufficiently soluble in buffer solutions alone and were first dissolved in 26 μ L of 95% ethanol before addition to respective buffers. An equivalent volume of ethanol was present in all sample solutions for consistency. Fluorescence measurements were taken immediately in 25-second intervals for 25 min, λ_{ex} = 360 nm λ_{em} = 460 nm. Changes in fluorescence were determined by subtracting the time-zero relative fluorescence units (RFU) from subsequent measurements (Ft–F0), and curves were normalized between quercetin and

blank controls for sample comparison. Samples were tested on three separate 96 well plates, eight wells per sample group per plate, for a total of 24 wells per sample group. Samples were randomized between columns, with the outermost columns of each plate excluded. All assays were carried out at room temperature.

Results

Mass Recovery and Determination of the Extinction Coefficient (ϵ)

Crofelemer is reported to have an average molecular weight of approximately 2300 daltons (Da) with oligomers of 3 to 30 monomers, with an average chain length of approximately 7 units.²⁷ Upon membrane fractionation of crofelemer extracted from tablets, using either a 3 kDa or 10 kDa MWCO filter, the red colored material was retained and the filtrate was colorless (see Methods section). The red fraction presumably contains a substantial portion of polymer-like crofelemer molecules that could not pass through the MWCO membrane, whereas low MW molecules, such as monomers and proanthocyanidin oligomers, might be able to freely pass through the membrane. Although the compound that is responsible for the red color of crofelemer appears to have a higher MW than the cutoff of the membrane, it is possible that the retention by the membrane may also involve hydrophilic interactions between proanthocyanidin molecules and the membrane (regenerated cellulose), which could result in a lower than expected yield in the filtrate. Non-covalent aggregation between crofelemer species due to extensive hydrogen bonding and/or π - π stacking of the highly aromatic substituents could also result in retention of low MW chains below that of the cutoff limit, resulting in a lower than expected number average MW within the concentrate after membrane fractionalization.

Due to the heterogeneity of crofelemer, the extinction coefficient (ϵ) was determined by three independent methods. These results were then averaged to minimize the variations from each method. First, aliquots of crofelemer extract (2 trials with a total of 12 experiments) were evaporated to dryness and the sample weights were recorded. The residual water by Karl Fischer titration was 0.014% w/w. The individual samples were then reconstituted in water and the UV-Vis spectra were collected to calculate the ϵ . The mean ϵ was 7.64 mL/(mg*cm) at 280 nm. Second, serial dilutions were made of a weighted crofelemer extract. Absorption data at 280 nm were fit to Beers' law to obtain a ϵ of 7.65 mL/(mg*cm). Third, variable pathlength spectroscopy (Solo VPE) was applied to develop a relationship between the pathlength and absorbance of crofelemer samples with known concentrations, and then ϵ was determined by Beers' law to be 7.51 mL/(mg*cm) at 280 nm. The extinction coefficient values determined by the three methods were averaged to obtain an ϵ of 7.6 mL/(mg*cm) at 280 nm, which was used to prepare subsequent analytical samples and to estimate material recovery.

Analysis of Fractionated and Stability Samples for Biophysical Properties

UV-Visible Absorption Spectroscopy—The UV-Visible absorption spectra of the unfractionated, 10 kDa Top/Bottom and 3 kDa Top/Bottom samples are shown in Figure 4. Using the corrected Day 0 absorbance value at 280 nm and the previously determined ϵ , the concentrations of the incubated fractions were: 0.44 mg/mL, unfractionated; 0.54 mg/mL, 10

kDa top; 0.15 mg/mL, 10 kDa bottom, 0.48 mg/mL, 3 kDa top; 0.08 mg/mL, 3 kDa bottom. These results were used to normalize the biopolymer concentration from other experimental results whenever necessary. After incubation at 25°C for 30 days, there was no obvious peak shift, but the raised baseline in the uncorrected spectra and decreased intensity at 280 nm in the corrected spectra indicated the formation of aggregates in all samples. Compared to spectra obtained from the 25°C incubation, the 2nd derivative spectra at 40°C showed slightly up-shifted peak maxima around 294 nm. Significantly decreased peak intensities in the 40°C corrected spectra also suggest sample aggregation. Single wavelength UV-Visible absorption measurements may be insufficient to normalize mass quantification across different lots for the assays. Thus, UV-Visible spectra were collected from 190 to 800 nm for all samples. Although the unfractionated, 10 kDa Top and 3 kDa Top fractions showed red color, no obvious absorbance above 300 nm was observed due to dilution.

Circular Dichroism—The CD spectra of Day 0 samples of the unfractionated, 10 kDa Top and 10 kDa Bottom had negative minima at 215 nm and positive maxima at 245 nm (Figure 5). However, the intensities at 215 nm decreased over the 30-day incubation at 25°C for all three samples, and down-shifted peak minima were observed. Meanwhile, the 245 nm maxima increased over the same incubation period with down-shifted peak positions for both unfractionated and 10 kDa Top samples. Peak shifts occurred in the 10 kDa Bottom sample without significant changes in intensity. The results from the 40°C incubated samples were very similar to the 25°C for the 10 kDa Bottom fraction, while the unfractionated and 10 kDa Top fraction show slight variations in the peak position and intensity at 215 nm.

FTIR Analysis—Differences in peak position and intensity were observed in the FTIR spectra of crofelemer samples from the accelerated stability study (Supplemental Figure S1). First, all three samples at Day 0 contained a single peak located at 1440 cm^{-1} and 1607 cm^{-1} , which can be assigned to C=C groups. After 30 days of incubation at both 25 and 40°C, the initial single peaks at 1607 cm^{-1} split into two visible features. Similar splitting could be observed in the 1440 cm^{-1} region as well, indicating structural changes in the aromatic rings. Also, the intensities of the features around 1340 cm^{-1} decreased significantly, which can be attributed to the loss of C-O groups. However, oxidation of the phenol structure (C-O) should result in quinone (C=O) and cause peak intensity increase around 1750 cm^{-1} , but no experimental evidence is observed in FTIR.

Analysis of Fractionated and Stability Samples for MW and Average Degree of Polymerization

^1H and Quantitative ^{13}C NMR Spectroscopy—The NMR spectra of crofelemer are not published. However, there are published peak assignments for the investigational precursor to crofelemer, SP-303,²⁰ which is from the same botanical source but not processed as a clinical material. Proton NMR spectra for crofelemer show broad peaks in the ranges of δ 2.0–3.1, 3.1–5.5, and 5.5–7.3 ppm. Quantitative ^{13}C spectra for crofelemer possess broad peaks at δ 28.81, 36.45, 38.21, 66.71, 72.59, 73.69, 76.44, 78.58, 83.07, 96.33, 97.81, 103.83, 107.58, 109.08, 115.50, 117.11, 120.28, 121.96, 132.44, 133.93, 144.80, 146.17, and 155.20 ppm.

Proton NMR spectra for Day 0 unfractionated, 10 kDa Top, and 3 kDa Top samples were similar in peak position and relative signal intensity. The Day 30 unfractionated samples incubated at both 25 and 40°C displayed a broadening of any distinguishable characteristics present in the peaks of Day 0 samples, with a significant decrease in the signal intensity of 40°C samples compared to 25°C samples (Figure 6).

In the qCNMR analysis, there was no observable change in chemical shift values of the 25 and 40°C incubated samples when compared to their Day 0 counterparts. However, peak intensities decreased and peaks broadened over the course of incubation, indicative of chemical and physical changes occurring within the sample entirety over time (Figure 7). Peaks of interest for determining IU/TU, Trans/Cis, and PC% and PD% were integrated to capture variations between samples. Respective values for each sample are shown in Table 2.

Day 0 unfractionated crofelemer had an IU/TU ratio of 6.76, resulting in an estimated number average molecular weight of 2328 g/mol, given the crofelemer monomers are approximately 300 g/mol. An unexpected decrease in molecular weight was observed in the Top fractions of both 10 kDa and 3 kDa MWCO filtrates, with estimated average molecular weights of 1581 g/mol and 1437 g/mol, respectively. A decrease in the IU/TU ratio was observed in the unfractionated sample after 30 days of incubation at 25°C due to an increase in terminal unit signal relative to total signal spectra, suggesting formation of terminal units through degradation of higher molecular weight species. Accurate integration of the δ 36.45, 38.21 ppm and 28.81 ppm peaks in the 30 Day, 40°C unfractionated crofelemer sample was not possible, probably due to sample loss by precipitation and/or more extensive physical and chemical degradation. Both 3 and 10 kDa Top fractions showed slight increases in the IU/TU ratio for their 40°C, 30 Day counterparts when compared to time zero values.

The trans/cis ratio increased from time zero to 30 days in the unfractionated and 3 kDa Top samples (25 and 40°C). There was a statistically insignificant increase in the epicatechin and epigallocatechin population, reflected by a relative increase in the δ 76.44 ppm signal when normalized to total signal of the spectrum. However, in 10 kDa Top samples the trans/cis ratio was nearly unchanged over time, with relatively consistent signal contributions from both δ 76.44 and 83.07 ppm shifts.

Trends in the PC and PD percentages were consistent between all three samples. Increases in PC% in the unfractionated, 3 kDa and 10 kDa Top fractions were due to increases in δ 146.17 ppm relative signal contribution along with decreased contributions from the δ 132–134 ppm range. Almost every 30 Day sample incubated at either 25 or 40°C, with the exception of the 30 Day, 25°C 10 kDa Top sample, showed decreased relative signal from δ 132.44 and 133.93 ppm peaks. This was indicative of loss of the B-5' hydroxyl of gallocatechin repeating units.

Size Exclusion Chromatography—Crofelemer detection during SEC of the stability samples was achieved using UV spectra collected from 190 to 800 nm for all retention times, but only the 280 nm signal is shown in Supplemental Figure S2 to simplify the presentation. The unfractionated crofelemer (S2, **Top**) had a single wide major peak at 12.4 minutes, consistent with the known MW distribution of crofelemer. This peak was then

collected at 30 seconds intervals as Fraction 1, 2 and 3, and re-injected for SEC. The chromatogram of the three fractions indicated that they contained different MW species, suggesting that the broad peak resulted from different MW fractions within the sample and not non-specific peak broadening by interaction with the column. Using a linear regression of log MW vs. retention time for a set of standards (see Methods), the MW of crofelemer was estimated to be ~3300 Da and the average degree of polymerization could be estimated by:

$$DP = \frac{MW_{crofelemer}}{MW_{monomer}} = \frac{3300}{290} = 11.4$$

In Figure 8, the SEC profile of the crofelemer stability samples show the retention time values of the main peak of unfractionated, 10 kDa Top and 10 kDa Bottom were 12.4, 12.3, 13.4 min, respectively, indicating that the 10 kDa Bottom fraction contains the lowest mass distribution of the three while the unfractionated and 10 kDa Top fraction were similar in MW. After 30 days of incubation at 25°C, the retention times of all samples decreased to 11.9, 11.7 and 12.7 min, respectively. This suggests the formation of larger size molecules or, more probably, sample aggregation. The overall peak area also increased significantly, suggesting that the extinction coefficient of the samples may change during the incubation period. Compared to the increase at 280 nm, the area change at 460 nm is more pronounced, which could arise from the existence of more conjugated molecules derived from potential sample oxidation. The results of the 40°C samples are generally similar to the 25°C incubation with slightly lower retention times and smaller changes in the total peak area.

HILIC HPLC—The crofelemer stability samples were not well resolved by hydrophilic interaction liquid chromatography (HILIC) due to the heterogeneous nature of the samples (Supplemental Figure S3). However, general observations can be made. The 3 and 10 kDa Bottom fractions had shorter retention times than the other three, which could indicate these fractions contained less polar species.

Ratio and Monomer Distribution for Thiolyzed Fractionated and Stability Samples

Thiolysis of Crofelemer—Thiolysis has been used for the characterization and quantification of proanthocyanidins.²⁵ A detailed thiolysis mechanism, evaluated for oxidative degradation products of crofelemer, is described in the accompanying article. In the presence of a nucleophile, such as β -mercaptoethanol (β -ME), terminal crofelemer units are released as flavanol monomers, and the extension units are derivatized with β -ME to form flavan-3-ol adducts (See Supplemental Figure S4). The resultant terminal monomers and flavanol adducts were analyzed by RP-HPLC at 280 nm.

RP-HPLC separation of crofelemer following thiolysis revealed that crofelemer contains catechin (C), epicatechin (epiC), gallocatechin (GC) and epigallocatechin (epiGC) as terminal units (Figure 9). Thiolysis generates thioether adducts in addition to monomers. The more hydrophilic, trihydroxylated, GC elutes earlier than the corresponding dihydroxylated C. The 2,3-trans isomers, C and GC, elute earlier than the 2,3-cis isomers, epiC and epiGC. Monomers were identified by injecting the authentic standards. The peaks

at 7.5 min, 14.9 min, 15.5 min and 27 min were assigned to GC, C, epiGC and epiC, respectively. To quantify the terminal monomers, peak areas of the respective peaks were calculated. The standard curves generated (See Supplemental Figure S5 B, C, D and E) were then used to quantify the monomers generated by thiolysis of crofelemer. Table 3A shows the composition of terminal monomers of thiolysed crofelemer. Then, the actual yield of each monomer was determined (Table 3B) using the calculated epimerization factors.

Quantification of Terminal Monomers—The composition of terminal monomers after thiolysis was analyzed by RP-HPLC and UV detection at 280 nm. Epimerization of C and GC to epiC and epiGC, respectively, and vice versa, is inevitable during thiolysis,²⁸ which could lead to an overestimation or underestimation of individual species. Therefore, the extent of epimerization was first evaluated by subjecting individual monomers to thiolysis conditions (see Supplemental Figure S5 B–E). Calibration curves were generated using known amounts of standards and an epimerization factor was determined. These epimerization factors (Table 4) were used to correct the actual yields of C, epiC, GC and epiGC derived from the thiolysis of crofelemer. The terminal monomers for unfractionated, Day 0 crofelemer are primarily GC units (33.23 ± 6.47 epiGC and 10.28 ± 0.25 GC). In both catechin and gallocatechin groups, the cis isomer was the major contributor in the particular lot of crofelemer tested.

RP-HPLC-UV Analysis of Thiolysed Monomers—Supplemental Figure S5 A shows the epimerization of C and epiC during thiolysis as monitored by RP-HPLC analysis and UV detection at 280 nm. Thiolysis of C results in four major products and the peaks corresponding to C and epiC were identified by comparison with authentic standards. The peaks eluting at 17.5 min and 20.5 min represent mercaptothioether derivatives of C. Similarly, the epimerization of GC and epiGC was monitored by HPLC analysis and UV detection (data not shown).

ESI-MS Analysis of Thiolysed Monomers—Negative ion ESI-MS analysis of thiolysed C generated four major peaks in the total ion current (TIC) chromatogram (Figure 10A). The products eluting at 5.89 and 6.36 min were identified as thioether derivatives of C (Figure 10B and 10C). The peaks eluting at 5.31 and 7.13 min have a similar fragment ion distribution in their mass spectra and the peak at 5.31 min has the same retention time as authentic C. Hence, the compound eluting at 7.13 min can be assigned to epiC. Similarly, products formed by GC thiolysis were identified by ESI-MS analysis.

Negative-Ion ESI-MS Analysis of Crofelemer—The negative ion ESI-MS spectrum of the various crofelemer stability samples are shown in Figure 11. The sequential occurrence of groups of peaks reflects the characteristic polymeric nature of the proanthocyanidins. Monomers, dimers, trimers, etc. were well separated. The consecutive clusters of peaks were separated by the mass of a monomer unit. Within a cluster, the mass difference between subsequent peaks is 16 Da, which corresponds to the substitution of a GC with a C unit. The intensity of the peaks decreased as the degree of polymerization increased. Monomer ions with m/z of 287.1, 289.1, 303.1 and 305.1 were located in the low mass region of the MS spectrum. The highest molecular weight singly charged oligomer detected in our ESI-MS

analysis was a cluster with m/z 1793.7, which corresponds to a hexamer. The cluster consists of peaks at m/z 1745.6, 1761.6, 1777.6, 1793.7, 1809.6, 1825.7 and 1841.7, suggesting that different ratios of monomers contribute to the composition of crofelemer.

For a more detailed mass spectrometry analysis of crofelemer and its fractions at various time and temperature conditions, please refer to the companion MS paper in the series.¹⁸

***In vitro* Bioactivity Evaluation of Fractionated and Stability Samples**—All of the crofelemer stability samples tested were able to inhibit Cl^- ion efflux from T84 cells, regardless of fractionation, time, and temperature (Figure 12). Though there was visual evidence of variability between crofelemer time zero samples and their 7 and 30 day degraded forms, Cl^- efflux levels persisted below passive Cl^- efflux in every case. Fluorescence values for crofelemer samples when challenged by either ionomycin or forskolin also remained below blank controls in all cases. Challenge with forskolin resulted in higher Cl^- efflux compared to crofelemer samples challenged with ionomycin and crofelemer alone. Overall Cl^- efflux behavior when cells were stimulated with ionomycin varied across samples; however, fluorescent values remained at or slightly above Ft-F0 values for crofelemer alone. No significant differences, given the precision of the bioassay, between samples incubated at 25 or 40°C could be detected within the same fractions or between different fractions tested (unfractionated, 10 kDa Top, 3 kDa Top). Fluorescence values for crofelemer sample groups were consistently above values seen for quercetin after 25 minutes, indicating that saturation of Cl^- channels did not occur.

Discussion

Spectroscopic Changes to the Samples During Stability

Crofelemer is a highly heterogeneous drug product due to its botanical origins and complex polymeric composition. It consists of procyanidins and prodelphinidins with varying degrees of polymerization and oxidation. Multiple methods are thus necessary to capture the wide variety of structural features presented in this biopolymer mixture. For example, due to variations of the composition of crofelemer, and potential changes of its averaged extinction coefficient, UV-Visible absorbance alone may be insufficient to determine mass quantities of samples across batches.

Ideally, samples from several batches should be used to analytically characterize the heterogeneous nature of complex pharmaceutical molecules such as crofelemer. However, since only limited commercial batches of crofelemer were available, we used product fractionations and accelerated stability conditions to simulate the heterogeneity that may occur across batches due to differences in raw material source, handling, and storage. Several gross structural and chemical changes were observed in the materials upon thermal treatment. The intensity of the 215 nm CD minimum decreased with increasing incubation time and temperature, suggesting the oxidation state or oligomerization level may affect the overall structures of crofelemer. The CD spectra presented different peak maxima between the crofelemer extracts and the published SP-303 material, suggesting that subtle structural differences may occur across batches or because of product purity. This result may also be derived from the interactions among the aromatic rings (π - π stacking) that produce

increased optical activity. In contrast to CD, the FTIR spectra of crofelemer agreed with SP-303 results in the literature.²⁰

Size and Compositional Changes to the Stability Samples During Stability

Fractionation of crofelemer remains challenging although a number of previous studies^{25, 29, 30} successfully applied normal phase and reversed phase HPLC for identifying monomer/oligomer species. In this study, both SEC and HILIC HPLC were applied to analyze the composition of crofelemer. SEC with standards was used to estimate the MW and degree of polymerization, but the results were different from MS and qCNMR by overestimating the size of the polymer. Since there are no commercially available standards of crofelemer oligomers, a set of similarly structured compounds was used to develop a regression curve. This could be a potential cause of the overestimation. Moreover, HILIC HPLC showed a very broad main peak instead of well separated individual species. Previous HPLC methods described in the literature^{31, 32} commonly focused on homogeneous oligomers. In contrast, crofelemer is composed of heterogeneous oligomers of different repeating units, which greatly increases the potential total number of possible species to 4^{DP} (degree of polymerization) and overpowered the separation capacity. Interestingly, the overall peak area for both SEC and HILIC significantly increased over the period of incubation. Since the samples were stored in sealed vials, diluted and injected in the same manner, it is reasonable to attribute the peak area changes to changes in extinction coefficients, which may be due to sample oxidation as well as aggregation.

The variability of chemical composition in crofelemer was evident in negative-ion ESI-MS analysis. Crofelemer is a mixture of proanthocyanidins with different degrees of polymerization and different ratios of monomers; in addition, the linkage type between monomers and the stereochemistry of subunits contribute to the complexity of the mixture. It is known that the bioavailability and optical absorption of proanthocyanidins tend to decrease with increases in the degree of polymerization.³³ The highest MW species observed in negative-ion ESI-MS was a hexamer with MW ~1800 Da. The upper mass limit of QToF in this analysis was set to m/z 2000 and multiple charged species were not observed. Hexamers, pentamers, tetramers, trimers, dimers and monomers are well resolved and separated by a mass of a single monomer, either (epi)C or (epi)GC, and appeared as clusters of peaks in the mass spectrum. Within a cluster, oligomer chains with different monomeric compositions could be visualized by the addition of (epi)GC or (epi)C units. Between the major ion clusters, a relatively low intensity series of peaks corresponded to ions derived by retro-Diels-Alder (RDA) fragmentation. The peaks for m/z 305 and 303 were much more intense in comparison to the peaks for m/z 289 and 287, suggesting that the most abundant monomer is (epi)GC. Our quantification also revealed that almost all crofelemer polymers had terminal units of (epi)GC. Thiolytic combined with UV detection at 280 nm was used as an orthogonal method to analyze the terminal composition of crofelemer polymers. The (epi)GC dominated as the terminal units in comparison to (epi)C. In addition, epi forms in which the pyrogallol ring at the second carbon and the OH group at the third carbon are cis to each other, were slightly predominant in the terminal units.

The peaks in the ^1H NMR spectra were relatively broad due to the heterogeneity of monomer composition and polydispersity of the crofelemer. Although distinct peaks were not resolved through ^1H NMR, some noticeable changes in the signature features of the peaks were observed over the course of incubation. The δ 2.0–3.1 ppm peak arises from the proton(s) located on the C4 sp^3 hybridized carbon, the δ 3.1–5.5 ppm peak consists of the signals from the C2 and C3 sp^3 hybridized carbons in their cis and trans conformations, and the δ 5.5–7.3 ppm peak consists of protons from sp^2 hybridized carbons of the A and B-rings of the repeating units. Changes observed in the δ 5.5–7.3 ppm peak, which occurred under forced degradation, may be indicative of oxidation and/or chemical changes in the B-ring, such as quinone formation or PD to PC conversions in crofelemer's substituents.

The qCNMR spectra revealed more distinctive chemical and physical changes that took place in crofelemer samples over time. Peak broadening occurred over the course of sample incubations, indicative of variation in the microenvironments of constitutive monomers. This broadening can be attributed to changes in resonance frequencies of adjacent monomers, either by introduction of heterogeneous species via chemical degradations of proximal monomers (e.g. oxidation), or through changes in intra- and interchain associations (also evidenced in changes to CD spectra over the course of incubation), or a combination of both. IU/TU, trans/cis ratios, and PC and PD percentages could be determined in most samples except for the 40°C, 30 Day unfractionated sample. The trans/cis, PC% and PD% values in all time zero samples were in agreement with those observed in the MS analysis. Samples were comprised primarily of GC (PC%) units in the cis configuration (trans/cis ratios < 1). An unexpected decrease in IU/TU ratios were observed in the time zero sample compared to the size fractionated samples. Association with regenerated cellulose, or self association of crofelemer through hydrogen bonding and/or π - π stacking could potentially form larger aggregates which could also result in the unexpected IU/TU ratios observed for the fractionated samples. Increases in the IU/TU values of the 30 Day, 40°C fractionated samples compared to their time zero counterparts indicate potential chemical additions at the C4 position of terminal units.

Increases in the presence of the procyanidin moieties, with consequent decreases in prodelphinidin moieties were observed across all fractions over time (excluding the 10 kDa Top, 30 Day, 25°C sample). The decrease in the 132–134 ppm signal range relative to total signal of the spectra indicates a loss of the triphenol structure of the gallo catechin units at the C5', C4' and C3' carbons. Interestingly, an increase in the 146.17 ppm signal relative to total signal was observed. A simple conversion from gallo catechin to catechin units via loss of the C5' hydroxyl should not result in an increase of 146.17 signal, as catechin and gallo catechin moieties contribute equally to this peak.

Effect of Structural Changes on *in vitro* Bioassay

A fluorescent microplate assay was utilized for assessing total chloride efflux from T84 monolayers to assess variations of activity between crofelemer and its adulterated forms. Ideally, degradation of the starting material should result in observable changes in a drug's ability to initiate a measurable response. Although the exact mechanism of action for crofelemer in the GI tract is not known, it has been shown to selectively inhibit the cystic

fibrosis transmembrane conductance regulator (CFTR) and calcium-activated chloride channels (CaCC).^{34–36} T84 cells express both CaCC and CFTR channels.²⁶ By comparing total passive Cl[−] efflux (blank) to efflux measured in the presence of crofelemer samples, the overall ability of crofelemer samples to inhibit Cl[−] efflux could be visualized. Challenge of the crofelemer samples was performed using selective initiators of either CaCC (ionomycin) or CFTR (forskolin) in an effort to distinguish potential loss of inhibitive activity toward single channels after crofelemer degradation had occurred. Quercetin, a highly similar monomeric form of the catechin and gallo catechin subunits which comprise crofelemer, was used at higher concentrations (100 μM) in an effort to saturate Cl[−] channels as a positive control. The subtle changes from the initial lot of crofelemer during the degradation process did not have any substantial effect on its ability to inhibit Cl[−] efflux. Degradation of crofelemer did not substantially impact the polymers ability to inhibit individual channels when challenged with forskolin or ionomycin. It is entirely possible that the extent of degradation investigated in this study was not sufficient to create detectable changes beyond the precision of the chloride channel assay. As a heterogeneous polymeric species, slight or even substantial degradation in the repeating units of the overall chains which comprise the polymer may not cause observable changes in activity with the cell-based assay used. If only a single repeating unit and not a coherent chain of catechin and/or gallo catechin of crofelemer is necessary for chloride channel inhibition, then minor degradations in a sample of crofelemer – though observable through analytical analysis – may not translate into distinguishable changes in a bioassay. This hypothesis is further supported by quercetins ability to inhibit chloride efflux as a mimic of monomeric crofelemer. A more in-depth analysis of the biological data using principle component analysis is presented in the third paper of this series.

Characterizing physicochemical and biological properties is challenging due to the complex nature of the crofelemer. UV-Visible absorption spectroscopy, CD, FTIR analyses were applied to determine the spectroscopic characters while NMR, MS, SEC and HILIC HPLC were used to provide detailed information regarding the MW and composition of the polymer. Although potential oxidation and sample aggregation were observed within multiple techniques, no obvious changes were found in the bioactivity assay, even with a month of incubation at 40°C. However, with the help of fractionation and incubation at various conditions, a large set of data regarding the physicochemical properties of crofelemer was collected and there were subtle but distinguishable changes obtained from individual assays. In the companion paper,¹⁹ a mathematical modeling/machine learning approach will be established to compare different crofelemer samples.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding for this work was made possible by the Food and Drug Administration through grant 1U01FD005285-01. Views expressed in this publication do not necessarily reflect the official policies of the Department of Health and Human Services nor does any mention of trade names, commercial practices, or organization imply endorsement by the United States Government. Support for the NMR instrumentation was provided by NIH Shared Instrumentation Grant #s S10RR024664 and S10RR014767 and NSF Major Research Instrumentation Grant # 0320648. PK was

partially supported by NIH R01CA173292 during these studies. The authors wish to also acknowledge Justin Douglas and Sarah Neuenswander of the KU Nuclear Magnetic Resonance Laboratory.

References

1. Federici M, Lubiniecki A, Manikwar P, Volkin DB. Analytical lessons learned from selected therapeutic protein drug comparability studies. *Biologicals*. 2013; 41(3):131–147. [PubMed: 23146362]
2. Lubiniecki A, Volkin DB, Federici M, Bond MD, Nedved ML, Hendricks L, Mehndiratta P, Bruner M, Burman S, Dalmonte P, Kline J, Ni A, Panek ME, Pikounis B, Powers G, Vafa O, Siegel R. Comparability assessments of process and product changes made during development of two different monoclonal antibodies. *Biologicals*. 2011; 39(1):9–22. [PubMed: 20888784]
3. Crommelin DJ, Shah VP, Klebovich I, McNeil SE, Weinstein V, Fluhmann B, Muhlebach S, de Vlieger JS. The similarity question for biologicals and non-biological complex drugs. *Eur J Pharm Sci*. 2015; 76:10–17. [PubMed: 25912826]
4. Bui LA, Hurst S, Finch GL, Ingram B, Jacobs IA, Kirchhoff CF, Ng CK, Ryan AM. Key considerations in the preclinical development of biosimilars. *Drug Discov Today*. 2015; 20(Suppl 1):3–15.
5. Alt N, Zhang TY, Motchnik P, Taticek R, Quarmby V, Schlothauer T, Beck H, Emrich T, Harris RJ. Determination of critical quality attributes for monoclonal antibodies using quality by design principles. *Biologicals*. 2016; 44(5):291–305. [PubMed: 27461239]
6. FDA. Botanical Drug Development Guidance for Industry (Draft Guidance). 2015 ed.
7. FDA. Orange book: approved drug products with therapeutic equivalence evaluations. Silver Spring, MD; US FDA: 2010.
8. Anderson J, Bell C, Bishop J, Capila I, Ganguly T, Glajch J, Iyer M, Kaundinya G, Lansing J, Pradines J, Prescott J, Cohen BA, Kantor D, Sachleben R. Demonstration of equivalence of a generic glatiramer acetate (Glatopa (TM)). *J Neurol Sci*. 2015; 359(1–2):24–34. [PubMed: 26671082]
9. Lee S, Raw A, Yu L, Lionberger R, Ya N, Verthelyi D, Rosenberg A, Kozlowski S, Webber K, Woodcock J. Scientific considerations in the review and approval of generic enoxaparin in the United States. *Nat Biotechnol*. 2013; 31(3):220–226. [PubMed: 23471071]
10. Administration TUSFaD. , editor. FDA. Guidance for industry – quality considerations in demonstrating biosimilarity of a therapeutic protein product to a reference product. Silver Springs, MD: 2015.
11. Administration TUSFaD. , editor. FDA. Guidance for industry – scientific considerations in demonstrating biosimilarity to a reference product. Silver Springs, MD: 2015.
12. Chow SC, Song F, Bai H. Analytical Similarity Assessment in Biosimilar Studies. *AAPS J*. 2016; 18(3):670–677. [PubMed: 26873509]
13. Burdick R, Coffey T, Gutka H, Gratzl G, Conlon HD, Huang CT, Boyne M, Kuehne H. Statistical Approaches to Assess Biosimilarity from Analytical Data. *AAPS J*. 2016
14. Kim JH, Joshi SB, Tolbert TJ, Middaugh CR, Volkin DB, Smalter Hall A. Biosimilarity Assessments of Model IgG1-Fc Glycoforms Using a Machine Learning Approach. *J Pharm Sci*. 2016; 105(2):602–612. [PubMed: 26869422]
15. More AS, Toprani VM, Okbazghi SZ, Kim JH, Joshi SB, Middaugh CR, Tolbert TJ, Volkin DB. Correlating the Impact of Well-Defined Oligosaccharide Structures on Physical Stability Profiles of IgG1-Fc Glycoforms. *J Pharm Sci*. 2016; 105(2):588–601. [PubMed: 26869421]
16. Mozziconacci O, Okbazghi S, More AS, Volkin DB, Tolbert T, Schoneich C. Comparative Evaluation of the Chemical Stability of 4 Well-Defined Immunoglobulin G1-Fc Glycoforms. *J Pharm Sci*. 2016; 105(2):575–587. [PubMed: 26869420]
17. Okbazghi SZ, More AS, White DR, Duan S, Shah IS, Joshi SB, Middaugh CR, Volkin DB, Tolbert TJ. Production, Characterization, and Biological Evaluation of Well-Defined IgG1 Fc Glycoforms as a Model System for Biosimilarity Analysis. *J Pharm Sci*. 2016; 105(2):559–574. [PubMed: 26869419]

18. Hewarathna AN, Mozziconacci O, Nariya MK, Kleindl PA, Xiong J, Forrest ML, Middaugh CR, Joshi SB, Volkin DB, Deeds EJ, Schöneich C, TBD. Chemical stability of the botanical drug substance crofelemer: a model system for the analysis of biosimilars. *J Pharm Sci.* TBD(TBD):TBD.
19. Nariya MK, Kim JH, Xiong J, Kleindl PA, Hewarathna AN, Joshi SB, Schöneich C, Forrest ML, Middaugh CR, Volkin DB, Deeds EJ, TBD. Biosimilarity assessment of crofelemer samples using data mining and machine learning approaches. *J Pharm Sci.* TBD(TBD):TBD.
20. Ubillas R, Jolad SD, Bruening RC, Kernan MR, King SR, Sesin DF, Barrett M, Stoddart CA, Flaster T, Kuo J, Ayala F, Meza E, Castañel M, McMeekin D, Rozhon E, Tempesta MS, Barnard D, Huffman J, Smee D, Sidwell R, Soike K, Brazier A, Safrin S, Orlando R, Kenny PTM, Berova N, Nakanishi K. SP-303, an antiviral oligomeric proanthocyanidin from the latex of *Croton lechleri* (Sangre de Drago). *Phytomedicine.* 1994; 1:77–106. [PubMed: 23195881]
21. Newman RH, Porter LJ, Foo LY, Johns SR, Willing RI. High-resolution ¹³C NMR studies of proanthocyanidin polymers (condensed tannins). *Magn Reson Chem.* 1987; 25:118–124.
22. Davis AL, Cai Y, Davies AP, Lewis JR. ¹H and ¹³C NMR assignments of some green tea polyphenols. *Magn Reson Chem.* 1996; 34:887–890.
23. Stringano E, Gea A, Salminen JP, Mueller-Harvey I. Simple solution for a complex problem: proanthocyanidins, galloyl glucoses and ellagitannins fit on a single calibration curve in high performance-gel permeation chromatography. *J Chromatogr A.* 2011; 1218(43):7804–7812. [PubMed: 21930278]
24. Karonen M, Liimatainen J, Sinkkonen J. Birch inner bark procyanidins can be resolved with enhanced sensitivity by hydrophilic interaction HPLC-MS. *J Sep Sci.* 2011; 34(22):3158–3165. [PubMed: 21998029]
25. Gu L, Kelm M, Hammerstone JF, Beecher G, Cunningham D, Vannozzi S, Prior RL. Fractionation of polymeric procyanidins from lowbush blueberry and quantification of procyanidins in selected foods with an optimized normal-phase HPLC-MS fluorescent detection method. *J Agric Food Chem.* 2002; 50:4852–4860. [PubMed: 12166971]
26. West MR, Molloy CR. A microplate assay measuring chloride ion channel activity. *Anal Biochem.* 1996; 241:51–58. [PubMed: 8921165]
27. FDA. Center for Drug Evaluation and Research. 2012. Fulyzaq NDA 202292 chemistry review http://www.accessdata.fda.gov/drugsatfda_docs/nda/2012/202292Orig1s000ChemR.pdf ed
28. Seto R, Nakamura H, Nanjo F, Hara Y. Preparation of epimers of tea catechins by heat treatment. *Biosci Biotech Bioch.* 1997; 61(9):1434–1439.
29. Engstrom MT, Paljarvi M, Frygas C, Grabber JH, Mueller-Harvey I, Salminen JP. Rapid qualitative and quantitative analyses of proanthocyanidin oligomers and polymers by UPLC-MS/MS. *J Agric Food Chem.* 2014; 62(15):3390–3399. [PubMed: 24665824]
30. Wallace TC, Giusti MM. Extraction and normal-phase HPLC-fluorescence-electrospray MS characterization and quantification of procyanidins in cranberry extracts. *J Food Sci.* 2010; 75(8):C690–696. [PubMed: 21535487]
31. Lazarus SA, Adamson GE, Hammerstone JF, Schmitz HH. High-performance liquid chromatography/mass spectrometry analysis of proanthocyanidins in foods and beverages. *J Agric Food Chem.* 1999; 47:3693–3701. [PubMed: 10552707]
32. Kelm MA, Johnson JC, Robbins RJ, Hammerstone JF, Schmitz HH. High-performance liquid chromatography separation and purification of cacao (*Theobroma cacao* L.) procyanidins according to degree of polymerization using a diol stationary phase. *J Agric Food Chem.* 2006; 54:1571–1576. [PubMed: 16506802]
33. Neilson AP, O’Keefe SF, Bolling BW. High-Molecular-Weight Proanthocyanidins in Foods: Overcoming Analytical Challenges in Pursuit of Novel Dietary Bioactive Components. *Annu Rev Food Sci Technol.* 2016; 7:43–64. [PubMed: 26735794]
34. Gabriel SE, Davenport SE, Steagall RJ, Vimal V, Carlson T, Rozhon EJ. A novel plant-derived inhibitor of cAMP-mediated fluid and chloride secretion. *Am J Physiol.* 1999;G58–G63. 276(Gastrointest Liver Physiol. 39). [PubMed: 9886979]

35. Tradtrantip L, Namkung W, Verkman AS. Crofelemer, an antisecretory antidiarrheal proanthocyanidin oligomer extracted from *Croton lechleri*, targets two distinct intestinal chloride channels. *Mol Pharmacol*. 2010; 77(1):69–78. [PubMed: 19808995]
36. Fischer H, Machen TE, Widdicombe JH, Carlson TJ, King SR, Chow JW, Illek B. A novel extract SB-300 from the stem bark latex of *Croton lechleri* inhibits CFTR-mediated chloride secretion in human colonic epithelial cells. *J Ethnopharmacol*. 2004; 93(2–3):351–357. [PubMed: 15234776]
37. Isogai A. NMR analysis of cellulose dissolved in aqueous NaOH solutions. *Cellulose*. 1997; 4:99–107.
38. Gast JC, Atalla RH, McKelvey RD. The ^{13}C -N.M.R. spectra of the xylo- and cello-oligosaccharides. *Carbohydrate Research*. 1980; 84:137–146.

Figure 1.
Structure of crofelemer defined in its application summary review.

Figure 2.

Crofelemer tablet before and after the coating is removed (top). Unfractionated crofelemer (bottom, left) and membrane fractionation of crofelemer using a 10 kDa MWCO filter (bottom, right). Reddish-brown colored crofelemer was separated into a reddish-brown fraction, which is retained in the insert, and a colorless fraction as filtrate.

Figure 3.
Carbon chemical shift values observed for crofelemer's monomeric constituents.

Figure 4.

UV-Vis of unfractionated crofelemer and its fractions over a total incubation time of one month at 25 or 40°C. The absorbance maxima were observed at 275 nm and scattering correction was applied to estimate the concentration.

Figure 5.

Comparison of representative CD spectra of five different crofelemer samples (Unfractionated (A), 10 kDa Top (B), 10 kDa Bottom (C), 3 kDa Top (D) and 3 kDa Bottom (E)) over one month incubation at two temperatures (25 and 40°C). One minimum at 217 nm is observed, and there were weak positive features located at 275 nm.

Figure 6.

^1H NMR comparison of Day 0 and Day 30 25°C, 40°C crofelemer samples ($n=1$). Sharp peaks (δ 3.4 ppm to δ 4.0 ppm) present in the 10 kDa and 3 kDa MWCO filtered samples are characteristic of cellulose that contaminated the sample during centrifugal filtration.³⁷

Figure 7.

Quantitative ^{13}C -NMR comparison of Day 0 and Day 30, 25 and 40 °C crofelemer samples ($n=1$). No visible changes in peak positioning were observed. Precipitation of sample over the 30 day period contributed to signal loss in the 40 °C samples, making some calculations unreliable. Residual acetone from the NMR tube was present in the unfractionated Day 0 sample, as well as the Day 30, 25 °C 10 kDa Top and 3 kDa Top samples. The signal contribution of the δ 31.2 ppm contaminant peak has been excluded in analysis of proximal peaks. Contributions from regenerated cellulose to the carbon spectra could potentially overlap with the δ 76.6 ppm signal from cis moieties in crofelemer,³⁸ however no consistent increase in cis population within samples was observed over time.

Figure 8.

Crofelemer and its fractions by SEC. Five MW standards references were applied to estimate the MW of crofelemer. MW distributions can also be estimated by total peak area from crofelemer and its fractions.

Figure 9.

RP-HPLC analysis of thiolyzed crofelemer at 280 nm. Standard solutions of C, epiC, GC and epiGC were separately injected onto the column to identify the corresponding peaks in the thiolyzed crofelemer mixture.

Figure 10.

TIC chromatogram of thiolyzed C (A). C and epiC are eluted at 5.31 and 7.13 min, respectively. Peaks eluting at 5.89 and 6.36 min correspond to C-thioether adducts. Negative ESI mass spectrum of the peak eluting at 6.36 min (B) and 5.89 min (C).

Figure 11.

Negative ESI mass spectrum of crofelemer. The polymeric nature of crofelemer is reflected by the sequential occurrence of clusters of peaks which were separated by the mass of one monomer unit. The insert shows the zoomed region of hexamers illustrating the substitution pattern. Within a cluster, the 16 Da mass difference between peaks corresponded to the addition of one hydroxyl group into the aromatic ring, and these data indicate that different ratios of monomers contribute to the composition of crofelemer.

Figure 12.
Representative T84 assay results for unfractionated crofelemer (Day 0, Day 30) compared to respective 3 kDa Top samples.

Table 1

Summary of the analytical techniques applied to characterize crofelemer.

Type	Critical Quality Attributes	Assays
Physical and Composition	Mass Recovery	UV-Vis
	Compound ID and Purity	¹ H NMR
		¹³ C NMR
	Average Degree of Polymerization	¹³ C NMR
		SEC-DAD
	MW Distribution	SEC-DAD
	Composition	HILIC HPLC-DAD
		FTIR
		Q-ToF
Chemical Assays	Oxidation	Thiolysis-LC-MS
		Circular Dichroism
Biological Activity	Cl ⁻ Channel Inhibition	C18 HPLC-UV
		QToF
Biological Activity	Cl ⁻ Channel Inhibition	Fluorescent Assay in T84 Intestinal Cell Monolayer

Table 2

Differences in IU/TU, trans/cis ratios and PC, PD percentages for crofelemer samples (n=1) as determined by qCNMR.

Sample	Trans/Cis	PC%	PD%	IU/TU
<i>Unfractionated, T0</i>	0.69	31.16	68.84	6.76
<i>Unfractionated, 30 Day, 25°C</i>	0.82	34.51	65.49	4.93
<i>Unfractionated, 30 Day, 40°C</i>	IS *	50.54	49.46	IS *
<i>10 kDa Top, T0</i>	0.72	33.52	66.48	4.27
<i>10 kDa Top, 30 Day, 25°C</i>	0.72	24.88	75.12	4.87
<i>10 kDa Top, 30 Day, 40°C</i>	0.64	72.58	27.42	4.53
<i>3 kDa Top, T0</i>	0.58	23.6	76.4	3.79
<i>3 kDa Top, 30 Day, 25°C</i>	0.73	47.33	52.67	3.61
<i>3 kDa Top, 30 Day, 40°C</i>	0.77	43.02	56.98	4.27

*
insufficient signal

Composition of terminal monomers of crofelemer.

Monomer	(A) Composition (%) of terminal units				(B) Actual composition (%) of terminal units (considering epimerization)				
	Exp 1 n=4	Exp 2 n=5	Exp 3 n=5	Exp 4 n=4	Exp 1 n=4	Exp 2 n=5	Exp 3 n=5	Exp 4 n=4	Average
C	0.22 ± 0.03	0.26 ± 0.02	0.19 ± 0.06	0.22 ± 0.03	0.21 ± 0.03	0.24 ± 0.02	0.18 ± 0.06	0.22 ± 0.03	0.21±0.02
epiC	0.29 ± 0.02	0.32 ± 0.03	0.32 ± 0.03	0.33 ± 0.03	0.29 ± 0.02	0.33 ± 0.03	0.33 ± 0.03	0.35 ± 0.03	0.32±0.03
GC	30.49 ± 10.42	24.15 ± 6.49	17.49 ± 7.27	24.47 ± 13.06	24.6 ± 10.42	15.65 ± 6.49	29.25 ± 7.27	16.49 ± 13.06	21.50±6.56
epiGC	20.34 ± 1.69	22.55 ± 1.82	20.66 ± 2.65	21.8 ± 1.79	26.79 ± 1.69	31.43 ± 1.82	29.64 ± 2.65	30.18 ± 1.79	29.51±1.96

Table 4

Percentage conversion of monomers into their epimers.

Monomer	Epimerization factors (%)				
	Exp 1	Exp 2	Exp 3	Exp 4	Average
C	0.94±0.02	0.83±0.01	0.67±0.002	0.81±0.001	0.81±0.01
epiC	3.58±0.02	3.67±0.03	2.31±0.02	2.54±0.01	3.02±0.02
GC	18.42±0.60	7.40±0.42	7.37±0.14	7.93±0.23	10.28±0.35
epiGC	37.99±7.07	34.68±3.09	29.38±10.32	30.86±5.39	33.23±6.47