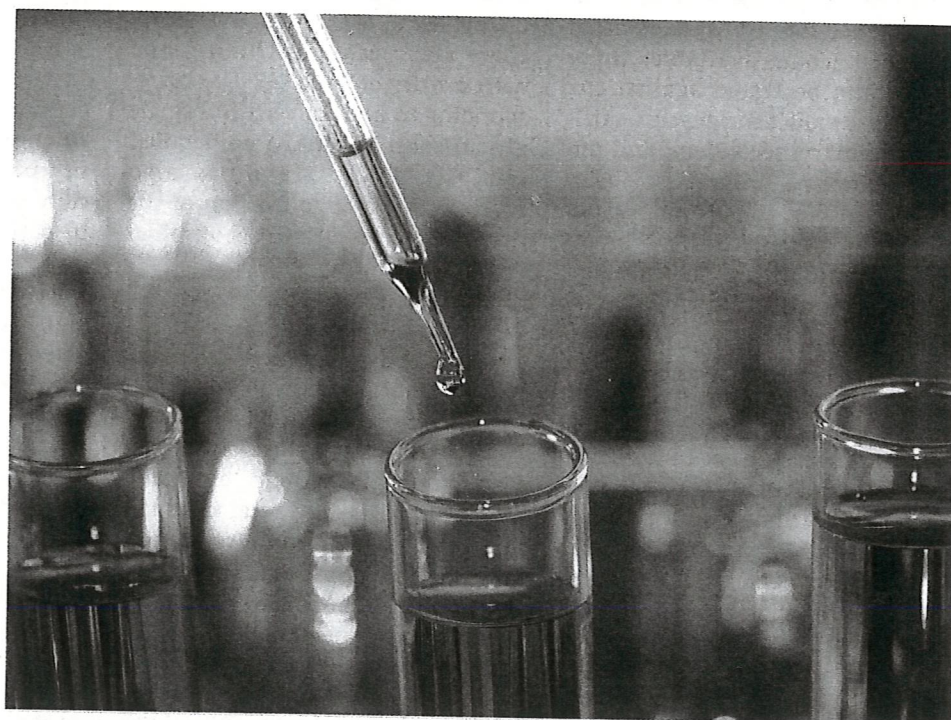


# Analytical Strategy in the Development of Biosimilars

Mario DiPaola

The author outlines an analytical strategy for establishing similarity in biosimilar development and approval.



SUWIT NGAOKAEW/Shutterstock.com

**I**n 2016, FDA announced its first three approvals of biosimilar products: Zorxio (filgrastim) and Erelzi (etanercept), made by Sandoz; and Inflectra (infliximab), made by Celltrion/Hospira (Pfizer) (1). While three approvals may not sound impressive, these approvals establish a clear path by which biosimilars can reach the United States market.

Since May 2016, applications for approval of biosimilars have been filed by Samsung Bioepis/Merck for infliximab; Coherus Biosciences for pegfilgrastim; Merck/Samsung Bioepis for insulin glargine; and Mylan/Biocon for trastuzumab. Four other filings have either not received FDA action yet or not received FDA complete response letters (1).

According to one publication (2), there are more than 1200 biosimilars in development against 16 major targets for potential US and global commercialization. The foundation for establishing similarity between a reference product and a biosimilar is a robust and sound analytical strategy demonstrating that the two products have similar primary, secondary, and tertiary structures, and similar functional and biological properties. This paper outlines some current methodologies and technologies available to develop a solid analytical strategy for use toward biosimilar development and approval.

## ESTABLISHING BIOSIMILARITY

In the very early stages of biosimilar development, it is important to carefully analyze a series of lots of the orig-

Mario DiPaola is senior scientific director at Charles River Laboratories in Woburn, MA.



inator product to determine the protein sequence, identify and quantify enzymatic and non-enzymatic post-translational modification (PTM), analyze biological functionality, and establish variability in product quality attributes against which the biosimilar product in development will be measured. Once lots of the biosimilar product become available, these lots should be tested against the reference product for physical attributes; primary, secondary, and tertiary structural properties; purity and presence of impurities, including those related to the product and its manufacture; and biological activity, all using orthogonal analytical methods that have sufficient sensitivity. It is important to recognize that product quality attributes should be ranked in criticality from critical to less critical and noncritical. For the critical quality attributes, it is expected that statistical equivalence be demonstrated between originator and biosimilar products. For less critical attributes, comparability should be based on reference product variability, typically based on the average  $\pm 2$  to 3 standard deviations; for noncritical product attributes, a graphical comparison will suffice for comparability. A comprehensive analytical characterization package that should be considered for comparison of biosimilar and originator products is provided in **Table 1**.

From the series of analytical methods listed in **Table 1**, one of the least complex tests to perform is intact mass analysis using a high-resolution mass spectrometer, preferably with electrospray ionization. Results from this method can reveal key product attributes, including the integrity of the molecule, post-translational modifications, and other types of protein modifications that may be present.

An example of the output from this type of analysis, often referred

to as top-down mass spectrometry, is provided in **Figure 1**. Results shown in this figure were obtained from the analysis of adalimumab. From the observed molecular weights in comparison to the theoretical one, one can easily conclude that major glycan species associated with this antibody are G0F (148083.0 Da) and G1F (148243.6 Da). Furthermore, there are species, including one with a mass of 146637.7 Da, corresponding to antibody with a single glycan substitution. More detailed analysis of the data has also revealed that a large fraction of the detected species lack the C-terminal lysine on the heavy chain.

### PEPTIDE MAPPING

While intact molecular mass data can be informative, more detailed chemical information is typically extracted from a well-resolved peptide map of the protein, especially when analyzed by liquid chromatography tandem-mass spectrometry (LC-MS/MS). The mirror-image peptide maps, shown in **Figure 2**, are derived from the tryptic digest of adalimumab, followed by analysis by reverse-phase high-performance liquid chromatography (RP-HPLC) with detection at 210 nm and tandem mass spectrometry (MS/MS).

Not only can the analysis by peptide mapping with LC-MS/MS confirm the protein sequence, it can also identify and locate within the primary structure of the molecule any post-translational modifications, including enzymatic ones such as N- and O-linked glycosylation, phosphorylation, and other common modifications; and non-enzymatic modifications including oxidation, deamidation, N-terminal cyclization, etc. In addition, by running maps with and without reduction, in most cases, it is possible to assign cysteines involved in the formation of disulfide linkages. Fragmentation by electron-transfer dissociation

(ETD) rather than collision-induced dissociation (CID) in MS/MS can further facilitate the localization of post-translational modification and disulfide bonds.

Visual comparison of the two maps shown in **Figure 2** indicates that the two molecules are quite similar overall in primary structure. Upon closer inspection, however, there is evidence of differences in height for some of the peptide peaks and the presence or absence of some of the low-intensity peaks. These differences typically should be closely scrutinized to better understand the types of molecular changes that would trigger changes in peptide maps, their significance, and if they could affect the efficacy and safety of the product.

### GLYCAN CHARACTERIZATION

Glycosylation can potentially impact not only the biological activity of the glycoprotein, but also the product circulation half-life and can render the protein immunogenic. Beyond establishing the location of glycosylation onto the glycoprotein, it is important to fully characterize the glycans. To do this, N-linked glycans can be simply released enzymatically and then derivatized for analysis by chromatography with fluorescence detection and LC-MS/MS for structural and linkage confirmation. O-linked glycans are released chemically by beta elimination and then permethylated before analysis by LC-MS. The glycosylation comparability between an originator antibody and a biosimilar is shown in **Figure 3**, with fluorescence profiles displayed as mirror images. The profiles are quite comparable in that the major species are G0F and G1F, but there are minor differences with respect to the observed lower-abundance glycan species. Interestingly, in this case, the originator molecule

**Table I.** Suggested analytical approaches for establishing biosimilarity.

Quality attribute	Methods
Primary structure	<ul style="list-style-type: none"> <li>• Intact mass by electrospray ionization–mass spectrometry (ESI–MS)</li> <li>• Peptide mapping by liquid chromatography–mass spectrometry (LC–MS) and MS/MS</li> <li>• N-terminal sequence by automated Edman degradation</li> <li>• Post-translational modifications               <ul style="list-style-type: none"> <li>o Glycosylation</li> <li>o Disulfides</li> <li>o Phosphorylation</li> <li>o Lipidation</li> </ul> </li> </ul>
Bioactivity (potency)	<ul style="list-style-type: none"> <li>• Cell uptake</li> <li>• Cell proliferation</li> <li>• Cytotoxicity               <ul style="list-style-type: none"> <li>o Antibody-dependent cell-mediated cytotoxicity (ADCC)</li> <li>o Antibody-dependent cellular phagocytosis (ADCP)</li> <li>o Complement-dependent cytotoxicity (CDC)</li> </ul> </li> </ul>
Target binding	<ul style="list-style-type: none"> <li>• Radioligand binding</li> <li>• Binding analysis by surface plasmon resonance</li> <li>• Binding analysis by fluorescence anisotropy</li> <li>• Immunoassays</li> <li>• Radioimmunoassay (RIA)</li> </ul>
Strength (protein concentration/content)	<ul style="list-style-type: none"> <li>• Absorbance at 280nm</li> <li>• Colorimetric assays</li> <li>• Chromatography (reverse-phase high-performance liquid chromatography [RP–HPLC], size-exclusion chromatography [SEC] etc...)</li> </ul>
High molecular weight species/aggregates	<ul style="list-style-type: none"> <li>• SEC–HPLC</li> <li>• Size-exclusion chromatography–multi-angle light scattering (SEC–MALS)</li> <li>• Native MS</li> <li>• Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE), (reduced and non-reduced)</li> <li>• Capillary electrophoresis (CE)–SDS (reduced and non-reduced)</li> <li>• Analytical ultracentrifugation</li> <li>• Field-flow fractionation</li> <li>• Dynamic light scattering</li> </ul>
High-order structures	<ul style="list-style-type: none"> <li>• Far and near ultraviolet (UV) circular dichroism</li> <li>• Fourier-transform infrared spectroscopy</li> <li>• Antibody array</li> <li>• Intrinsic/extrinsic fluorescence</li> <li>• Differential scanning calorimetry</li> <li>• Hydrogen/deuterium exchange MS</li> <li>• Ion-mobility MS</li> <li>• 1D and 2D nuclear magnetic resonance (NMR)</li> <li>• X-ray crystallography</li> </ul>
Charge distribution	<ul style="list-style-type: none"> <li>• IEF</li> <li>• cIEF</li> <li>• Ion exchange chromatography</li> </ul>
Non-enzymatic post-translational modification (PTM): oxidation	<ul style="list-style-type: none"> <li>• RP–HPLC</li> <li>• Peptide mapping with LC–MS</li> </ul>
Non-enzymatic PTM: deamidation	<ul style="list-style-type: none"> <li>• IsoQuant analysis</li> <li>• Peptide mapping with LC–MS</li> </ul>
N- or C-terminal truncation	<ul style="list-style-type: none"> <li>• MS</li> <li>• N-terminal sequencing</li> </ul>
Glycosylation	<ul style="list-style-type: none"> <li>• Oligosaccharide profiling               <ul style="list-style-type: none"> <li>o HPLC (normal phase, hydrophilic interaction) with fluorescence detection</li> <li>o HPLC with MS</li> </ul> </li> <li>• Sialic acid analysis</li> <li>• Monosaccharide analysis</li> </ul>
Process-related impurities	<ul style="list-style-type: none"> <li>• Host cell proteins by enzyme-linked immunosorbent assay (ELISA) or LC–MS</li> <li>• Residual deoxyribonucleic acid (DNA)</li> <li>• Expression inducers by chromatography</li> <li>• Anti-foam agents by chromatography</li> <li>• Leachates by LC or gas chromatography (GC)</li> </ul>



contains a higher abundance of high mannose species, which have been shown to lead to faster product-clearance from circulation (3). Of course, significantly faster clearance of the originator product will translate into higher relative potency for the biosimilar if the two are dosed similarly.

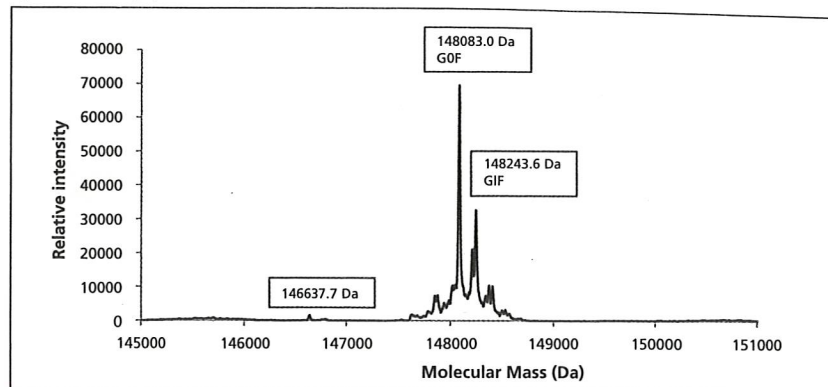
A variety of the methods listed in **Table I** are available to address charge variants, impurities, strength, and potency through binding and cell-based assay; for brevity, no further discussion on these attributes will be provided in this article.

### SPECTROSCOPIC METHODS

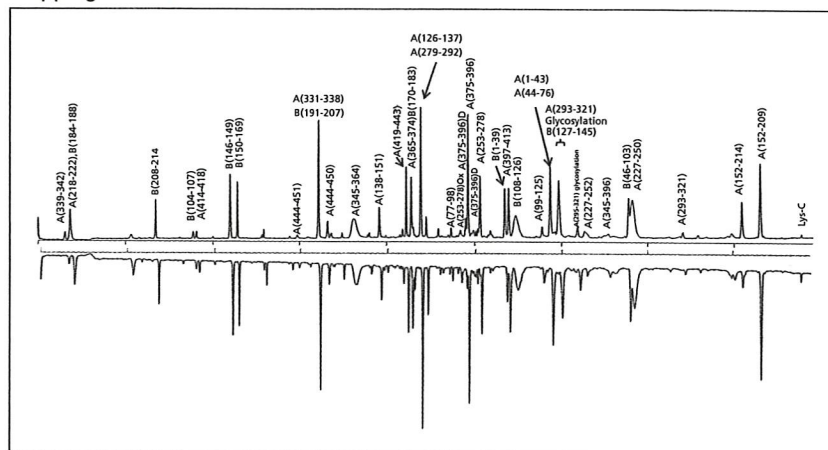
When applied correctly, tools for addressing primary structural comparability and potency (**Table I**) will produce high-quality data that can be used to establish a detailed picture of the primary structure and post-translational modifications of the protein. Unfortunately, not nearly as many tools are available to address protein/glycoprotein secondary, tertiary, and as applicable, quaternary structures. While there are some high-resolution methods such as multi-dimensional nuclear magnetic resonance (NMR) and X-ray crystallography, the cost, complexity, and time investment to carry out these methods precludes their widespread and routine use. Instead, there is reliance on lower-resolution methods that do not provide information at an atomic coordinate level, but provide structural information that is spatially averaged over the protein.

These lower-resolution spectroscopic methods, including circular dichroism (CD), fluorescence, and infrared spectroscopy, require some form of reporter moiety that relates to a structural characteristic of the protein, and ultimately can be correlated to the non-covalent bonding pattern of

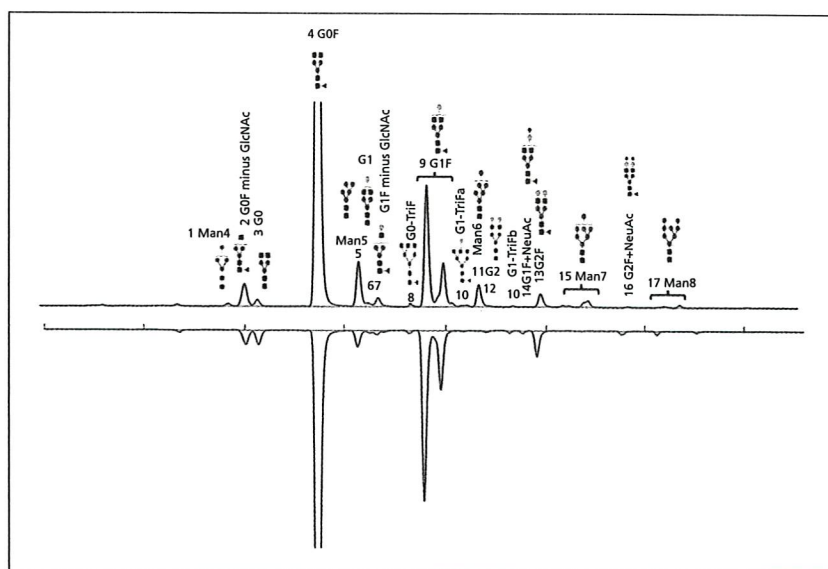
**Figure 1.** Molecular mass profile of adalimumab by electrospray ionization–mass spectrometry (ESI–MS).



**Figure 2.** Comparison of originator and biosimilar adalimumab lots by peptide mapping with detection at 210 nm.



**Figure 3.** Comparison of N-linked glycan profiles between originator and biosimilar antibodies.



**Table II.** Comparison of analytical strategies used in the development of recently FDA-approved biosimilars. LC is liquid chromatography; MS is mass spectrometry; RP is reverse phase; HPLC is high-performance liquid chromatography; UV is ultraviolet; CD is circular dichroism; SEC is size exclusion chromatography; FTIR is Fourier transform infrared spectroscopy; DSC is differential scanning calorimetry; ITS is intrinsic tryptophan fluorescence; NMR is nuclear magnetic resonance; SDS-PAGE is sodium dodecyl sulfate polyacrylamide gel electrophoresis; CE is capillary electrophoresis; IEF is isoelectric focusing; cIEF is capillary isoelectric focusing; FFF is field flow fractionation; IEX is ion exchange chromatography; CZE is capillary zone electrophoresis; AUC is analytical ultracentrifugation.

Product attribute		Method	Product		
			Filgrastim	Infliximab	Etanercept
Primary structure	Sequence confirmation	N-terminal sequencing	Yes	No	No
		LC-MS/MS	Yes	Yes	Yes
	Intact mass	MALDI-ToF MS	Yes	No	Yes
		ESI-ToF MS	Yes	Yes	No
	Post-translational modifications	Peptide mapping with LC-MS/MS	Yes	Yes	Yes
		DNA sequencing	Yes	No	No
		Amino acid analysis	No	Yes	Yes
	Glycosylation	N-linked	NA	Yes	Yes
		O-linked	NA	NA	Yes
		Sialic acid	NA	Yes	Yes
		Monosaccharide	NA	Yes	No
	Oxidation	RP-HPLC	Yes	No	No
LC-MS/MS		Yes	Yes	Yes	
Disulfide mapping	LC-MS/MS	Yes	Yes	Yes	
Potency	Receptor binding		Yes	Yes	Yes
	Cell based bioassay		Yes	Yes	Yes
Protein content (strength)		A280	No	Yes	Yes
		RP-HPLC	Yes	No	No
High-order structure		Far and near UV CD	Yes	Yes	Yes
		FTIR	No	Yes	Yes
		DSC	No	Yes	Yes
		ITF	No	No	No
		ETF	No	No	No
		Antibody array	No	Yes	No
		H/D exchange MS	No	No	Yes
		1D NMR	Yes	No	Yes
		2D NMR	Yes	No	No
	X-ray crystallography	No	No	Yes	
Impurities	Aggregates/high molecular weight species	SEC	Yes	Yes	Yes
		AUC	No	Yes	Yes
		FFF	No	No	Yes
		SDS-PAGE	Yes	No	No
	Truncation/fragmentation/partially reduced species	CE-SDS	No	Yes	No
		CE-SDS	No	Yes	Yes
		SDS-PAGE	Yes	No	No
		RP-HPLC	No	No	No
		LC-MS	Yes	Yes	Yes
	LC-MS	No	No	No	
Charge variants		IEF	Yes	Yes	No
		cIEF	No	No	No
		CZE	No	No	Yes
		IEX	Yes	Yes	No



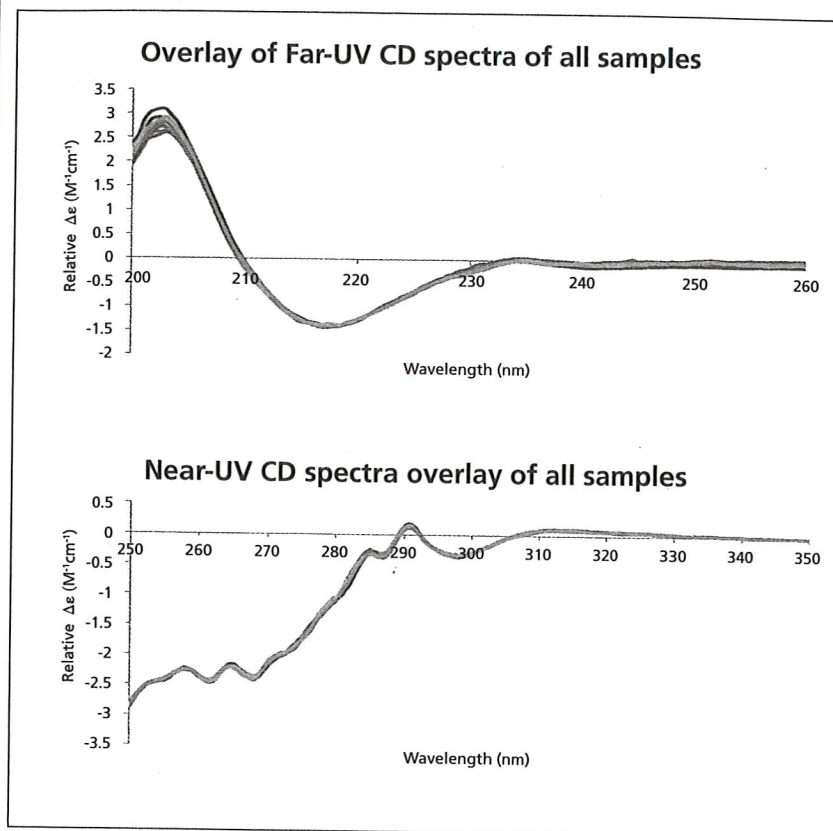
the folded polypeptide chains. As the reporter moieties are not often associated with specific locations in the protein structure, the obtained information is a spatial average. By combining methods that rely on different reporters and different modes of measurement, it is possible to build an integrated picture of the protein structural motifs and establish biosimilarity. An example of CD results from side-by-side comparison of seven biosimilar lots with seven originator lots is provided in **Figure 4**.

CD spectroscopy of proteins is quite sensitive to the three-dimensional orientation of the peptide-bond group (far-UV CD; 190 to 250 nm), disulfide bonds (near-UV CD; 250 to 320 nm), and aromatic side chains (near-UV CD; 250 to 290 nm). Deconvolution of CD spectra in the far-UV can be used to quantitatively estimate the types of secondary structure, whereas near-UV CD spectra can provide useful information on the local folding environment surrounding aromatic residues. Separately, Fourier transform infrared spectroscopy (FTIR) should also be included in the study of protein secondary structure. The absorption bands from stretching vibrations of the C=O (amide I) and C-N (amide II) groups of the protein backbone are useful for quantifying different types of secondary structure; notably, FTIR is more sensitive than CD for beta-derived secondary structural conformations, which is useful in the characterization of beta-rich proteins like monoclonal antibodies.

#### CALORIMETRIC TECHNIQUES

Besides spectroscopic methods, calorimetric techniques such as differential scanning calorimetry (DSC) should also be applied for the characterization of biosimi-

**Figure 4.** Comparison of N-linked glycan profiles between originator and biosimilar antibodies.



lars and comparability studies. DSC provides information on the structural stability of the folded polypeptide, given that the temperature ( $T_m$ ) at which denaturation occurs is characteristic of the protein stability. Because the denaturation transitions should be the same for a protein drug product and its biosimilar, one can use the DSC thermograms to demonstrate that two products derived from different manufacturing processes are structurally comparable.

#### ALTERNATIVE TECHNIQUES

Other techniques that would complement optical spectroscopic methods include hydrogen/deuterium exchange mass spectrometry, antibody array mapping, and ion-mobility mass spectrometry. Unlike the spectroscopic methods, these alternatives require greater

expertise, are likely to be more time consuming, and require significantly costlier instrumentation. However, they are expected to be much more informative than optical spectroscopy alone.

Protein aggregation is an undesired type of impurity that can result in enhanced product immunogenicity (4). Structural assays that have sufficient accuracy and precision to quantify such aggregates in protein drug products represent an important component in biosimilar comparability. Determination of the soluble aggregate levels in protein pharmaceuticals has historically relied on size-exclusion chromatography (SEC).

*Contin. on page 46*

Contin. from page 43

In recent years, however, there has been an increased awareness that SEC can yield erroneous aggregation results (5), due to several factors: possible adsorption of the aggregates to the SEC stationary phase; analysis under non-native conditions due to specific mobile-phase requirements; physical filtration of large aggregates; and dilution effects resulting in disruption of weak aggregates (6, 7). To avoid these SEC shortfalls, two alternate orthogonal methods—analytical ultracentrifugation (AUC) and field-flow fractionation (FFF)—have been more widely used. Both methods employ instrumentation, and more importantly, separation mechanisms that differ from SEC. Indeed, in many cases, AUC can provide evidence of large soluble aggregates present in a drug product that can go undetected by SEC, for reasons described in the previous paragraphs. Furthermore, AUC, by contrast to SEC, does not result in analyte dilution or capture on a stationary phase, and can often be done directly in the formulation matrix. With recent improvements in AUC and FFF resulting in better precision and accuracy, either of these methods can be used to accurately quantify the aggregate content in support of biosimilar comparability studies.

### COMPARABILITY STRATEGIES

To validate the analytical approaches discussed, a review of the comparability strategies related to FDA-approved biosimilar products has been undertaken. A summary of the comparability approaches used by Sandoz for filgrastim and etanercept and Celltrion for infliximab is provided in **Table II**.

Briefly, the developers placed significant emphasis on the confirmation by LC-MS/MS of the primary structure and post-translational modifications. Additionally, extensive effort has gone into the demonstration of biosimilarity by a combination of binding and cell-based assays. For aggregation and other high molecular weight species, the developers have relied on SEC but also with confirmation by either AUC or FFF. Interestingly, a less uniform strategy is evident among these three products for determining higher-order structures. The similarity in secondary/tertiary structures for filgrastim was conducted mainly through 2D NMR analysis, while crystallography and hydrogen/deuterium (H/D) exchange MS were used for etanercept, and a combination of CD, FTIR, and antibody array mapping was used for infliximab.

### CONCLUSION

To date, three biosimilar products have been approved by FDA. Many other products are currently in review

and are expected to receive approval in 2017. The regulatory path for getting to market in the US is becoming better defined; indeed, it is anticipated that there will be an acceleration in biosimilar approvals by FDA in the next five years.

As described in this paper, a variety of analytical tools are available to support the development of these biosimilars, especially with respect to chemical and biological/functional confirmation and comparability. While the three-dimensional structure of these molecules does present a significant challenge, there are several low-resolution methods that can be readily applied. However, more powerful methods are becoming available, including H/D exchange MS and ion-mobility MS, that can help with elucidating the spatial features of these molecules, and thus establishing comparability.

### REFERENCES

1. Pipeline Report—Biosimilar Drugs, US Specialty Care (July 2016).
2. R.A. Rader and E.S. Langer, "Future Manufacturing Strategies for Biosimilars," *BioProcess Int.*, May 17, 2016.
3. A.M. Goetze et al., *Glycobiology*. 21 (7) 949-959 (2011).
4. M.P. Baker et al., *Self Nonself*. 1 (4) 314-322 (2010).
5. J.F. Carpenter et al., *J. Pharm. Sci.* 99 (5) 2200-2208 (2010)
6. T. Arakawa et al., *J. Pharm. Sci.* 99 (4) 1674-1692 (2010).
7. K. Tsumoto et al., *J. Pharm. Sci.* 96 (7) 1677-1690 (2007). ♦

## Ad Index

Company	Page
AAPS	31
BIOMERIEUX INC	2
CHARLES RIVER LABS	19
CPHI	35
DISTEK INC	13
EUROFINS LANCASTER LABORATORIES	11
GE HEALTHCARE BIO-SCIENCES AB	39
HAMILTON CO	47
KERRY	7
NOVA BIOMEDICAL	27
NOVO NORDISK PHARMATECH A/S	23
THERMO FISHER SCIENTIFIC	45
TOSOH BIOSCIENCE	5
WUXI APP TEC	48