Chromatographic Alignment of ESI-LC-MS Proteomics Data Sets by Ordered Bijective Interpolated Warping

ARTICLE *in* ANALYTICAL CHEMISTRY · OCTOBER 2006

Impact Factor: 5.64 · DOI: 10.1021/ac0605344 · Source: PubMed

2 AUTHORS:

John T Prince
Brigham Young University - Provo Main Cam…
36 PUBLICATIONS **539** CITATIONS

SEE PROFILE

Edward Marcotte
University of Texas at Austin
191 PUBLICATIONS **16,244** CITATIONS

SEE PROFILE

Anal. Chem. **2006,** *78,* 6140−6152

Chromatographic Alignment of ESI-LC-MS Proteomics Data Sets by Ordered Bijective Interpolated Warping

John T. Prince and Edward M. Marcotte*

Center for Systems and Synthetic Biology, Institute for Cellular and Molecular Biology, and Department of Chemistry and Biochemistry, University of Texas at Austin, Austin, Texas 78712

Mass spectrometry proteomics typically relies upon analyzing outcomes of single analyses; however, comparing raw data across multiple experiments should enhance both peptide/protein identification and quantitation. In the absence of convincing tandem MS identifications, comparing peptide quantities between experiments (or fractions) requires the chromatographic alignment of MS signals. An extension of dynamic time warping (DTW), termed ordered bijective interpolated warping (OBI-Warp), is presented and used to align a variety of electrospray ionization liquid chromatography mass spectrometry (ESI-LC-MS) proteomics data sets. An algorithm to produce a bijective (one-to-one) function from DTW output is coupled with piecewise cubic hermite interpolation to produce a smooth warping function. Data sets were chosen to represent a broad selection of ESI-LC-MS alignment cases. High confidence, overlapping tandem mass spectra are used as standards to optimize and compare alignment parameters. We determine that Pearson's correlation coefficient as a measure of spectra similarity outperforms covariance, dot product, and Euclidean distance in its ability to produce correct alignments with optimal and suboptimal alignment parameters. We demonstrate the importance of penalizing gaps for best alignments. Using optimized parameters, we show that OBI-Warp produces alignments consistent with time standards across these data sets. The source and executables are released under MIT style license at http://obi-warp.sourceforge.net/.

One major goal of proteomics, the comprehensive study of cellular proteins across a variety of conditions, has been intensely pursued through chromatographic separation and mass spectrometry analysis. Top-down approaches focus on intact proteins while bottom-up approaches study peptide fragments, usually created enzymatically. In "shotgun proteomics", protein mixtures are digested en masse. The complexity of the resulting peptide mixture has prompted the advent of the MudPit style experiment where proteins are subjected to multiple chromatographic dimensions of separation and peptide identities are determined by fragmentation spectra in addition to their mass.[1] Quantification of peptides (reviewed in ref 2) is generally achieved through integration of peptide peaks in chromatographic traces acquired by MS. Notable exceptions involve the use of iTRAQ (Applied Biosystems) reagents (where quantification is coupled with MS/MS acquisition) and peptide sampling methods (quantification using statistics generated during peak picking).[3] Isotope labeling strategies (e.g., ICAT or SILAC) allow peak ratios to be calculated between two samples without regard to ion suppression, although some have managed direct comparisons between successive samples without isotopic labels.[4,5]

A formidable challenge in proteomics remains the integration of quantitation information across multiple mass spectrometry runs. When possible, data sets may be integrated based on peptide identities discovered through tandem MS. Unfortunately, the result of tackling high-complexity samples—tryptic digests may contain vast numbers of peptides (there are ~300 000 yeast tryptic peptides before considering posttranslational modifications (PTMs)), fluctuating quantities, and orders of magnitude differences in abundance—with the semistochastic, time-, and abundance-dependent nature of MS/MS sampling and variable confidence identifications, is that peptide identities in one run are never fully duplicated in another. Paradoxically, as more samples are run on a subject, so diminishes the overlapping set of peptide quantities that can be compared. Also, proteins or peptides most important to the experimental variable(s) in question are those most likely to go unidentified in some of the analyses since their concentrations are most likely to be in the greatest flux.

If LC−MS runs can be aligned chromatographically, then identities acquired in one run can be leveraged across all the others. For a given experimental subject, the union of peptide identities discovered in this way would approach completeness over time, and quantities could be compared for peptides without an ID in each run. More complete peptide coverage and quantitation has implications for low-abundance peptides. Peptides with

* To whom correspondence should be addressed. E-mail: marcotte@icmb.utexas.edu. Fax: (512) 232-3432.

(1) Washburn, M. P.; Wolters, D.; Yates, J. R. *Nat. Biotechnol.* **2001**, *19*, 242−7.
(2) Ong, S. E.; Mann, M. *Nat. Chem. Biol*. **2005**, *1*, 252−62.
(3) Rappsilber, J.; Ryder, U.; Lamond, A. I.; Mann, M. *Genome Res.* **2002**, *12*, 1231−45.
(4) Wang, W.; Zhou, H.; Lin, H.; Roy, S.; Shaler, T. A.; Hill, L. R.; Norton, S.; Kumar, P.; Anderle, M.; Becker, C. H. *Anal. Chem*. **2003**, *75*, 4818−26.
(5) Wiener, M. C.; Sachs, J. R.; Deyanova, E. G.; Yates, N. A. *Anal. Chem*. **2004**, *76*, 6085−96.

10.1021/ac0605344 CCC: $33.50
Published on Web 07/28/2006

low stoichiometry compared to sibling peptides are candidates for PTMs, and additional searches may uncover these modified peptides, which are often missed. Chromatographic alignment may also benefit the development of better tandem MS search algorithms by providing peptide identities for low confidence or previously unidentified MS/MS spectra on a large scale.

The alignment and subsequent extraction of quantities from MS proteomics data sets is a difficult problem. Fractions may contain thousands to hundreds of thousands of eluants, many of which will coelute. Significant percentages of the peaks may vary or be absent in tests of biological variation. In the extreme but common case, prefractionated samples to be aligned may only contain a small fraction of overlapping signal, depending on the quality and execution of the separation technique. Chromatographic reproducibility will likely degrade across multiple samples, especially over a long period of time. Nanoliter flow rates are becoming common to reduce sample consumption and enhance sensitivity but can be difficult to accurately reproduce.[6] Finally, spectra with low signal-to-noise ratios are often encountered as MS sensitivities are pushed to their limits. Thus, alignment algorithms for MS proteomics data must be especially robust.

Chromatographic Alignment of Mass Spectrometry Data. We briefly review methods that have been applied to the chromatographic alignment of complex data or have some potential in this regard, noting possible shortcomings for application in this domain.

Methods relying on the introduction of chemical standards have been proposed. Internal standards have been used for alignments in metabolic studies.[7] A limited number of internal standards may not be able to accurately capture nonlinear chromatographic variability with a high degree of accuracy, but too many standards may compete for valuable analytical signal and introduce unwanted signal suppression. The addition of mobile-phase tracer molecules can provide a continuous chromatographic reference (as a ratio of intensities) by which alignment may occur and even enhance MS/MS identifications in complex proteomics samples.[8] However, the precision of alignment appears to be constrained by the gradient program, and although the ratio of intensities shows less deviation than peak retention times, this ratio is indirect to the actual peaks in question.

A class of alignment algorithms builds on MS/MS-derived peak identities.[9−12] While the reliance on MS/MS data appears to be effective in these cases, the use of tandem spectra for alignment presents some drawbacks: (1) It requires that there be a significant percentage of overlapping MS/MS identifications. In particular, isotope labeling experiments where tandem spectra are only collected for peaks with high concentration differences may suffer a shortage of shared tandem spectra. Also, depending on prefractionation quality, consecutive fractions to be aligned may contain few common identifications. (2) Misidentified MS/MS spectra may be problematic to the alignment. The inclusion of more tandem spectra may result in better alignments, but the larger the group used, the more false positives will likely be included in the set. (3) Samples to be aligned may suffer from biases if they use different data-dependent sampling parameters (the point along a chromatographic peak where a peptide is sampled is subject to change). (4) Techniques relying on MS/MS data for chromatographic alignment are incompatible with pure MS approaches, which may offer significant advantages[13,14] such as reduced analysis time and greater quantification accuracy.

Several methods are capable of producing alignments independent of tandem spectra through peak picking and peak matching. Some derive from metabolic studies where tandem spectra are generally absent or rare.[15,16] "XCMS", for example, identifies and iteratively groups similar peaks from MS data and fits with a loess function.[16] Choosing similar peaks may be more challenging in proteomics with noisier data sets and 1 or 2 orders of magnitude more peaks. Scenarios with few overlapping peaks (low overlapping signal) may be especially troublesome for these methods. In addition, it has been noted recently that all existing peak matching methodologies require that "the deviation in retention time from sample to sample be no greater than the time between two adjacent peaks",[16] a restraint not necessarily relevant to other methods. Some methods require high-precision mass spectrometers for the generation of accurate mass and retention time pairs.[13,17] These typically rely on a user-defined window for alignment. While they seem promising, the accurate mass/time approaches are incompatible with lower resolution mass spectrometers and may be difficult to apply to multistage prefractionation experiments.[2]

There are multiple methods from chemometrics for aligning signals (e.g., using Bessel's inequality[18] or the Fraga-Synovec 2D alignment method[19]). Many require calibrant samples, and since these are typically applied to fairly well defined systems, it is unclear how they might respond to more complex alignment scenarios. Polynomial fitting is an alternative approach that has been proven difficult to apply to complex data.[20] Since the desired result of an alignment is usually quantitation information, we highlight the parallel factor analysis (PARAFAC2) algorithm. It corrects for misalignment while simultaneously applying the PARAFAC algorithm for determining component concentrations in multiway analysis.[21] The coupling of alignment and multiway analysis[22] may be a drawback in some applications by proscribing

(6) Gershon, D. *Nat. Methods* **2005**, *2*, 466.
(7) Frenzel, T.; Miller, A.; Engel, K.-H. *Eur. Food Res. Technol.* **2003**, *216*, 335−42.
(8) Chen, S. S.; Aebersold, R. *J. Chromatogr., B* **2005**, *829*, 107−14.
(9) Idborg, H.; Edlund, P. O.; Jacobsson, S. P. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 944−54.
(10) Idborg-Björkman, H.; Edlund, P. O.; Kvalheim, O. M.; Schuppe-Koistinen, I.; Jacobsson, S. P. *Anal. Chem.* **2003**, *75*, 4784−92.
(11) Higgs, R. E.; Knierman, M. D.; Gelfanova, V.; Butler, J. P.; Hale, J. E. *J. Proteome Res.* **2005**, *4*, 1442−50.
(12) Andersen, J. S.; Wilkinson, C. J.; Mayor, T.; Mortensen, P.; Nigg, E. A.; Mann, M. *Nature* **2003**, *426*, 570−4.
(13) Silva, J. C.; Denny, R.; Dorschel, C. A.; Gorenstein, M.; Kass, I. J.; Li, G. Z.; McKenna, T.; Nold, M. J.; Richardson, K.; Young, P.; Geromanos, S. *Anal. Chem.* **2005**, *77*, 2187−200.
(14) Aebersold, R. *Nature* **2003**, *422*, 115−6.
(15) America, A. H.; Cordewener, J. H.; van Geffen, M. H.; Lommen, A.; Vissers, J. P.; Bino, R. J.; Hall, R. D. *Proteomics* **2006**, *6*, 641−53.
(16) Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. *Anal. Chem.* **2006**, *78*, 779−87.
(17) Anderle, M.; Roy, S.; Lin, H.; Becker, C.; Joho, K. *Bioinformatics* **2004**, *20*, 3575−82.
(18) Grung, B.; Kvalheim, O. M. *Anal. Chim. Acta.* **1995**, *304*, 57−66.
(19) Fraga, C. G.; Prazen, B. J.; Synovec, R. E. *Anal. Chem.* **2001**, *73*, 5833−40.
(20) Eilers, P. H. *Anal. Chem.* **2004**, *76*, 404−11.
(21) Kiers, H. A.; TenBerge, J. M. F.; Bro, R. *J. Chemom.* **1999**, *13*, 275−94.
(22) Pravdova, V.; Walczak, B.; Massart, D. L. *Anal. Chim. Acta.* **2002**, *456*, 77−92.

the use of other quantification methods (e.g., higher-order singular value decomposition) on the aligned data.

Correlation Optimized Warping (COW),[23,24] recently described as a constrained variant of dynamic time warping (DTW),[25] was developed with the alignment of chromatographic samples in mind and has seen much use toward that end.[24,26,27] However, COW depends on proper selection of node positions, which may be difficult to achieve in complex samples, and it may overlook areas of significant nonlinear change.

Dynamic Time Warping. DTW, first used in speech processing,[28] is an approach to align single or multivariate signals across time while preserving the internal ordering of the signals. It has been applied to the alignment of chromatographic signals for some time.[22,29,30]

The dynamic time warping algorithm is as follows (adapted from ref 31; see also ref 22): For time series X and Y having lengths $|X|$ and $|Y|$,

$$X = x_1, x_2, \ldots x_i, \ldots x_{|X|}$$

$$Y = y_1, y_2, \ldots y_i, \ldots y_{|Y|}$$

determine a warp path W (where $W = w_1, w_2, \ldots w_k$) with a length K, where the kth element of the path is

$$w_k = (i, j)$$

(i and j indices into X and Y respectively) such that

$$\max(|X|, |Y|) \leq K < |X| + |Y|$$

Formally, the beginning of each time series should be used at the start of the warp path, $w_1 = (1,1)$, and the end of each series finishing at $w_k = (|X|, |Y|)$, though this requirement is often relaxed in practice. In addition, every index in both time series must be used, and i and j must be monotonically increasing:

$$w_k = (i,j),\ w_{k+1} = (i', j')\ i \leq i' \leq i + 1, j \leq j' \leq j + 1$$

A path's total score (T) is computed by summing the similarity (S) between the data points X_i and Y_j in the kth element of the warp path (w_{ki}, w_{kj}):

$$T(W) = \sum_{k=1}^{k=K} S\ (w_{ki}, w_{kj})$$

If a warp path is drawn in two dimensions (with X and Y indices on their respective axes) then steps along the diagonal represent a perfect correspondence between the data. In this representation, horizontal or vertical steps, equivalent to the advance of one index without the other, are called gaps or transitions.

In speech processing, the similarity function, S, and total score, T, are generally interpreted geometrically—similar data points have a small Euclidean distance between them and the best score is the shortest path in Euclidean space. For spectra, a benefit function has been used as the similarity function, with the maximum sum defining the optimal score.[32] If the full spectrum is used, each spectral comparison is at best $O(n)$ (where n is the number of peaks) in time.

The solution to DTW lends itself to dynamic programming (DP) approaches where a globally optimal solution may be found by recursively solving subproblems. For dynamic time warping, the DP solution is $O(n^2)$ (where n is the number of time points per sample) in space and time, a relatively efficient means to a guaranteed optimal solution. Techniques are available to increase the speed of DTW, such as warp path radii,[33,34] or FastDTW.[31]

DTW Suited to Complex Proteomics Data. DTW has been used to align complex proteomics samples in two cases: Wang et al. use a small subset (~200) of chromatograms and apparently a straightforward application of dynamic time warping for their alignments.[4] Prakash et al. apply several novel methodologies to dynamic time warping, including a score function that minimizes the noise distribution in spectra and the use of adjacent spectra in calculating spectral similarity scores.[32] They make available an online web service (ChAMS) demonstrating their alignment method and present a measure of alignment accuracy for at least one alignment.

The potential absence or change in peak height for a large fraction of peaks, and irregularities in the precise elution of individual peaks introduces a strong element of noise to the alignment of complex proteomics data. DTW as applied to spectra is democratic—if different alignment interpretations exist, it is the alignment with the greatest number of peaks in agreement that wins (under ordering and gap penalty constraints). Random fluctuations and variation in the elution time of individual analytes may offer alternative alignment paths, but since these should be self-canceling as they will be evenly distributed, the combined voice of legitimate shared signal should be ample to overcome significant amounts of spectral and biological noise. It should be emphasized that DTW, like the alignment approaches discussed earlier, produces a global alignment—individual peaks deviating from the general ordering will be misaligned in proportion to their deviation.

In this work, we optimize dynamic time warping parameters for complex proteomics data sets, including several candidate

(23) Nielsen, N. V.; Carstensen, J. M.; Smedsgaard, J. *J. Chromatogr., A* **1998**, *805*, 17−35.
(24) Bylund, D.; Danielsson, R.; Malmquist, G.; Markides, K. E. *J. Chromatogr., A* **2002**, *961*, 237−44.
(25) Tomasi, G.; Berg, F.; Andersson, C. *J. Chemom.* **2004**, *18*, 231−41.
(26) Debeljak, Z.; Srecnik, G.; Madic, T.; Petrovic, M.; Knezevic, N.; Medic-Saric, M. *J. Chromatogr., A* **2005**, *1062*, 79−86.
(27) Christensen, J. H.; Tomasi, G.; Hansen, A. B. *Environ. Sci. Technol.* **2005**, *39*, 255−60.
(28) Silverman, H. F.; Morgan, D. P. *IEEE ASSP* **1990**, *7*, 6−25.
(29) Riner, E.; Abbey, L. E.; Moran, T. F.; Papamichalis, P.; Shafer, R. W. *Biomed. Mass Spectrom.* **1987**, *6*, 491−8.
(30) Wang, C. P.; Isenhour, T. L. *Anal. Chem.* **1987**, *59*, 649−54.
(31) Salvador, S.; Chan, P. *KDD Workshop MTSD*, 2004.
(32) Prakash, A.; Mallick, P.; Whiteaker, J.; Zhang, H.; Paulovich, A.; Flory, M.; Lee, H.; Aebersold, R.; Schwikowski, B. *Mol. Cell. Proteomics* **2006**, *5*, 423−432.
(33) Sakoe, H.; Chiba, S. *IEEE ASSP* **1978**, *ASSP-26*, 43−9.
(34) Myers, C.; Rabiner, L. R.; Rosenberg, A. E. *IEEE ASSP* **1980**, *ASSP-28*, 623−35.

spectral similarity functions and a gap penalty function (such as is used in biological sequence alignment) using MS/MS-derived peak identities as externally derived validation. We also demonstrate a method to transform the discrete, many-to-many (nonbijective) results obtained with dynamic time warping into a smooth, bijective warping function. Finally, we compare our alignments to those found by using ChAMS.

METHODS

Data Set Preparation. Raw MS runs and corresponding SEQUEST search data were downloaded from the Open Proteomics Database (OPD)[35] and the Peptide Atlas Repository.[36] ThermoFinnigan RAW files were converted to mzXML[37] with the linux binary ReAdW (v. 2.5). No preprocessing was performed but could easily be incorporated into the analysis if desired. MS spectra were transformed into a uniform matrix by rounding to the nearest m/z (summing in the event of multiple values) and applying monotone piecewise cubic hermite interpolation (PCHIP)[38] along the time dimension to fill in missing values and create regularly spaced time samples of varying frequency (3-, 6-, and 12-s increments).

Chromatographic Standards and Measures of Accuracy. Protein/Peptide Prophet data were downloaded from the data repositories, if available. Data sets without these data were further analyzed by Peptide/Protein Prophet (Trans-Proteomic Pipeline v. 1.2.3). These analyses have been posted to OPD. MS/MS identifications whose protein probability, initial peptide probability, and NSP adjusted probability were ≥ 0.99 were chosen as candidate time standards, and the time of the MS scan containing the precursor ion was recorded. Peptides at a given charge sampled >2 times were discarded, as it was reasoned that these represent especially broad peaks that might offer less chromatographic precision than those with fewer (albeit still highly confident) IDs. A single elution time for peptides with two identifications was estimated by taking a weighted average based on the intensity of the precursor ion peak. Peak maximums could serve as more precise time standards; however, we avoided this since the peak finding process for complex, noisy data sets may introduce additional noise. After selecting the overlapping peptides for each alignment, outliers (likely caused by a mistaken identification) were thrown out by iteratively removing time points appearing >5 standard deviations from the regression line. Although this is only a rough estimate for outliers, the results were generally consistent with the 0.99 probability threshold chosen, thereby providing additional verification of Protein/Peptide Prophet accuracy.

Bijective Synchronization. To create a bijective warp path, we move sequentially along the DTW warp path and include or discard points according to certain criteria: (1) The first and last points of a warp path are always included. (2) Points along the diagonal not belonging to a transition are included. (3) The point with the best similarity score in a transition is included, unless (4) no more than one point in a transition can be included.

Occasionally, a horizontal and vertical transition may meet and so include the same point at their vertex. In the case where the vertex is the point of greatest similarity in the first transition, the other points in the second transition would be ignored. In the event that a nonvertex point in the second transition had a similarity score higher than the vertex, a minor bias toward earlier points would result.

Normalization of Similarity Functions. To standardize the influence of a global gap penalty to different data sets using different score comparisons, we normalize the similarity scores by the mean and standard deviation of the similarity score distribution. For consistency in executing the score functions, we negated the normalized Euclidean distance distribution and maximized the warp path (the exact equivalent of minimization). We verified that the functions used here produce normal distributions of similarity scores using data sets of randomly generated spectra (1000 spectra $\times$ 1000 time points; 0−1 000 000 in intensity), justifying the normalization. Large sections of chromatographic lead-in or trail-out time may influence the overall distribution of similarity scores; we do not address that in this work.

Interpolation. Through interpolation, a bijective warp map is used as the basis for constructing a smooth warping function that may be applied to either run. The choice of interpolant and how it is applied may influence the final outcome of an alignment. We chose the monotone PCHIP[38] method for all alignments here.

The responsiveness of the bijective interpolation can be adjusted by altering the number of included bijective anchors—fewer anchors will give a smoother interpolant (smaller derivatives as viewed from the diagonal). Ordered bijective interpolated warping (OBI-Warp) determines which anchors to include (if all are not selected) by subdividing the bijective warp map based on the number of selected anchors and selecting the point of highest similarity in each section to become an anchor for interpolation. Choosing fewer anchors effectively allows a user to disregard subsections with low (i.e., potentially spurious) alignment signal and still achieve a globally satisfying alignment using points of strongest similarity spread throughout. Given a desired number of anchors, A, the total set of bijective anchors T is divided into segments with N anchors in each ($N \sim T/A$) where rounding is distributed across the segments. A bijective anchor B is included in A if it has the highest similarity score S among the N anchors in the segment G:

$$B = \{x \in G : S(x) = \max\{S(x) : x \in G\}\}$$

For this paper, all possible bijective anchors were used so that the warp path would be highly responsive to variation in the alignment path during optimizations.

Interpolation in this work is applied to the bijective anchors as points in traditional x, y coordinate space giving a diagonal warp function. The bijective anchors may also be transformed into $x, y - x$ space to give an alternative interpolation, but this variant is not explored here.

After choosing bijective anchors, OBI-Warp interpolation may be applied in different ways with implications for peak integration

(35) Prince, J. T.; Carlson, M. W.; Wang, R.; Lu, P.; Marcotte, E. M. *Nat. Biotechnol.* **2004**, *22*, 471−2.

(36) Desiere, F.; Deutsch, E. W.; King, N. L.; Nesvizhskii, A. I.; Mallick, P.; Eng, J.; Chen, S.; Eddes, J.; Loevenich, S. N.; Aebersold, R. *Nucleic Acids Res.* **2006**, *34*, D655−8.

(37) Pedrioli, P. G.; Eng, J. K.; Hubley, R.; Vogelzang, M.; Deutsch, E. W.; Raught, B.; Pratt, B.; Nilsson, E.; Angeletti, R. H.; Apweiler, R.; Cheung, K.; Costello, C. E.; Hermjakob, H.; Huang, S.; Julian, R. K.; Kapp, E.; McComb, M. E.; Oliver, S. G.; Omenn, G.; Paton, N. W.; Simpson, R.; Smith, R.; Taylor, C. F.; Zhu, W.; Aebersold, R. *Nat. Biotechnol.* **2004**, *22*, 1459−66.

(38) Fitsch, F.; Carlson, R. *SIAM J. Num. Anal.* **1980**, *17*, 238−46.

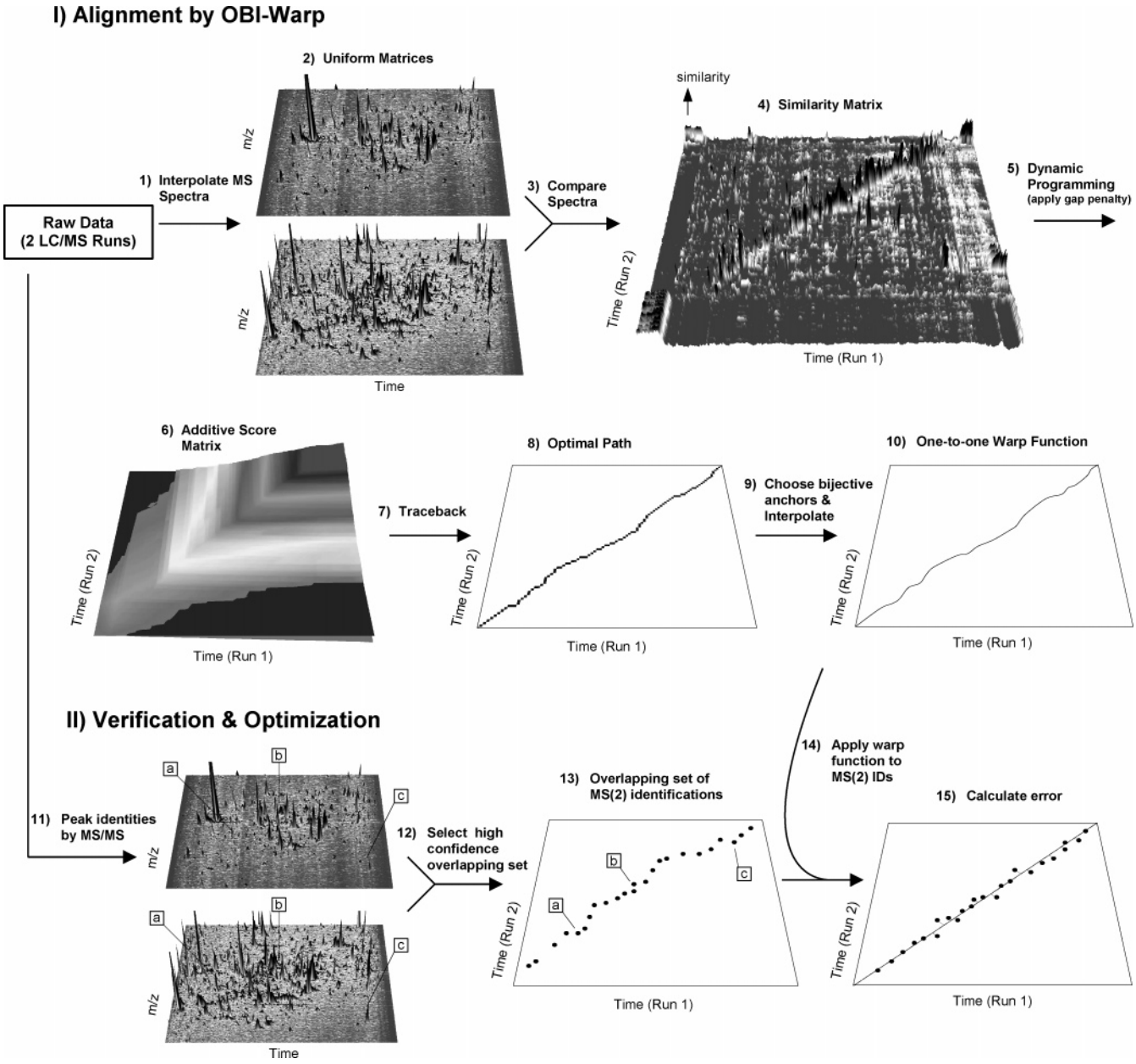

Figure 1. Flow diagram showing the chromatographic alignment of mass spectra by OBI-Warp and verification/optimization by MS/MS peak identities. (I) Alignment with bijective interpolated dynamic time warping. Raw data from two LC−MS runs, whether successive fractions or across different biological conditions, (1) is interpolated into a (2) uniform matrix (or rectilinear matrix). (3) An all vs all similarity matrix of the spectra is constructed. (4) The similarity matrix distribution is mean centered and normalized by the standard deviation. (5) Dynamic programming is performed by adding similarity scores along a recursively generated optimal path while off-diagonal transitions are penalized by either a local or global gap penalty to give (6) an additive score matrix. (7) Pointers are kept in a traceback matrix used to deliver (8) the optimal alignment path. (9) High scoring points in the optimal path are selected to create a bijective (one-to-one) mapping, which is used as anchors for PCHIP interpolation to generate a smooth warp function. (II) Verification and optimization. (11) MS/MS spectra from the raw MS runs are searched via SEQUEST and Peptide/Protein Prophet to determine peak identities. (12) High-confidence identifications are selected and (13) the overlapping set of peptide identifications (after filtering outliers) is used as the alignment standard. (14) The warp function produced through the comparison of MS data is applied to the standards. (15) The ideal alignment would shift all standards to the diagonal. The accuracy of an alignment is calculated as the sum of the square residuals from the diagonal.

algorithms. The underlying data may be warped (changing the underlying intensities through interpolation, stretching, and shrinking), or, as applied in this work, the time points labeling the intensity matrix axis may be altered, preserving the original intensities of the data.

Measures of Alignment Accuracy. To measure alignment accuracy we warp MS/MS-derived time standards with the bijective warping obtained by aligning MS signals. A perfect alignment would position (perfectly derived) standards precisely along the diagonal. We calculate error in one of two ways: the sum of the square residuals from the perfect alignment (SSR) and the average absolute time difference between time standards (AAD).

Other Alignment Experiments. Alignments for Supporting Information Figure 1 were performed on the 6-s interpolated alignment of F5 and F6 of the scx data set. For Figure A, random

Table 1. Overview of the LC−MS Data Sets Used in Optimization

Label	Emphasis	Run Label	Database - Acc# / Name	Date MS acquired	Organism / Description	Comparisons
ecoli	Same sample with different injection volume	e10	OPD - opd00005_ECOLI 021016.jp32A.10ul.3	10/16/2002	*Escherichia coli* 10uL injection	(e10 :: e15) 000_ld_020 :: 000_ld_020, 020_ld_040 :: 020_ld_040, 040_ld_060 :: 040_ld_060, 060_ld_080 :: 060_ld_080, 080_ld_100 :: 080_ld_100, 100_ld_150 :: 100_ld_150, 150_ld_200 :: 150_ld_200, 200_ld_300 :: 200_ld_300
		e15	OPD - opd00006_ECOLI 021010.jp32A.15ul.1	10/10/2002	*Escherichia coli* 15uL injection	
scx	Adjacent chromatographic fractions	scx	PeptideAtlas - (none given) Comp12vs12standSCX	NA	*Saccharomyces cerevisiae* standard strong cation exchange fractions	F2 :: F3, F2 :: F4, F3 :: F4, F4 :: F5, F5 :: F6, F6 :: F7, F7 :: F8, F8 :: F9, F9 :: F10, F10 :: F11, F11 :: F12, F12::F13
size	Chromatographic fractions with little overlap	size	PeptideAtlas - (none given) Comp12vs12sizefrac	NA	*Saccharomyces cerevisiae* size fractions	A1 :: A2, A2 :: A3, B1 :: B2, B2 :: B3, C1 :: C2, C2 :: C3, A1 :: B1, A2 :: B2, A3 :: B3, B1 :: C1, B2 :: C2, B3 :: C3
msmeg	Different biological state	early	OPD - opd00014_MYCSM 7-17-03	07/17/2003	*Mycobacterium smegmatis* early exponential growth	(early :: middle :: stat) [compared all 3 ways] 000 :: 000 :: 000 020 :: 020 :: 020 040 :: 040 :: 040 060 :: 060 :: 060 080 :: 080 :: 080 100 :: 100 :: 100
		middle	OPD - opd00009_MYCSM 6-17-03	06/17/2003	*Mycobacterium smegmatis* middle exponential growth	
		stat	OPD - opd00028_MYCSM 6-06-03	06/06/2003	*Mycobacterium smegmatis* stationary phase	
yeast	Different biological state (very different analysis conditions from msmeg)	gly	OPD - opd00044_YEAST 6-22-04-YM_N14N15_DAYGly040704_14_45_v0.4	06/22/2004	*Saccharomyces cerevisiae* one carbon metabolism with glycine	(gly :: ser) 005a :: 005_1, 005b :: 005_2, 005c :: 005_3, 020a :: 020_1, 020b :: 020_2, 020c :: 020_3, 060a :: 060_1, 060b :: 060_2, 060c :: 060_3, 900a :: 900_1, 900b :: 900_2, 900c :: 900_3
		ser	OPD - opd00043_YEAST 6-21-04-YM_N14N15_DAYSer040704_14_45_v0.4	06/21/2004	*Saccharomyces cerevisiae* one carbon metabolism with serine	

noise at a desired fraction was generated by providing a random value between 0 and 2× the fractional part of the signal selected. The propagative multiple alignment test was performed on the 020 mM salt fraction of OPD accession numbers 8−21 (opd00008_MYCSM−opd00021_MYCSM). Chromatographic time values were taken directly from the data (no interpolation to create a uniform matrix was performed).

Implementation. The software OBI-Warp is written in C++, compiles under Linux and Windows (with MinGW), and should compile under any system with gcc. All C++ objects are also known to compile with Microsoft Visual C++ (6.0). Vector and matrix classes were modeled after the Template Numerical Toolkit (http://math.nist.gov/tnt/). Most scripts for the generation of MS/MS time standards were written in Ruby. The PCHIM and PCHFE routines (and dependencies) from the public domain engineering/mathematical suite SLATEC were translated and rewritten with modifications and additions into C++ code. In particular, a subroutine for the interpolation of a sorted array of evaluation points (as occurring in both instances used herein) was written that only requires a single traversal of the input arrays, an algorithmic improvement on the SLATEC routines of $O(N+M)$ compared to $O(N*M)$. The OBI-Warp package is released under an unrestrictive MIT style license and can be downloaded from http://obi-warp.sourceforge.net/. A plugin, obiwarp, was written in C++ for zlab (http://zlab.sourceforge.net/), a custom library of low-level gui-tools allowing direct access to openGL for scientific data visualization. The plugin OBI-Warp allows the manipulation of alignment parameters (e.g., gap penalty) and the real-time visualization of the similarity matrix, additive score matrix, traceback matrix, optimal path, bijective anchors, smooth warp function, time standards, and sum of square residuals calculation. The plugin is also released under the MIT style license and is available with the OBI-Warp package.

RESULTS AND DISCUSSION

Figure 1 outlines the overall methodology followed in this paper, illustrating the major steps in alignment by OBI-Warp and also depicting the validation process used for verification and optimization. We interpolate MS spectra into matrices, compare spectra for their similarities, and perform dynamic time warping to create a warp path. We transform the warp path into an optimal one-to-one (bijective) mapping and interpolate to create a smooth warp function. Overlapping, high-confidence MS/MS identifications are used as time standards by which to judge the validity of alignments and optimize alignment parameters.

Data Set Selection. The data sets and comparisons used in this study are shown in Table 1. All samples are of high complexity, consisting of analyses of crude, trypsinized protein fractions containing >1000 peptides each. They were selected to provide a cross section of electrospray ionization liquid chroma-

tography mass spectrometry (ESI-LC-MS) alignment scenarios, ranging from adjacent chromatographic fractions of the same sample to parallel analyses of biologically varying samples. We use the scx and size multifraction experiments to compare alignments between successive fractions derived from these two different prefractionation techniques. While strong cation exchange (SCX) fractions generally have significant overlap between successive chromatography fractions, it is expected that there is little overlap between different size fractions, thus testing perhaps the most extreme alignment scenario. The ecoli data set represents an easier case for alignment where only the injection quantity was altered between the two LC/LC/MS/MS runs. The msmeg and yeast data sets represent tests of biological variation, with each set employing different chromatography and MS technique. The msmeg data sets were collected over nearly a dozen SCX fractions. The yeast data sets contain only four salt fractions and use a mass fractionation technique, dividing the length of full MS scans for each analysis into thirds. These data sets also have a 1:5 MS to MS/MS scan ratio, so the number of MS scans per unit time is less than in other runs. Because all alignments between different samples (as opposed to successive fractions) were performed over multiple SCX fractions, we expect that prefractionation variability will make some fractions more difficult to align with their counterpart fraction.

Spectra Similarity Function. Although initial applications of DTW to spectral data used spectra reduced to total ion chromatograms (TIC),[30] the alignment of complex data sets undoubtedly requires greater dimensionality than the TIC or base peak. Wang et al. selected 200 representative m/z values and scored the differences in intensities across these values.[4] Prakash et al. used a fuzzy dot product on unbinned spectra and subtracted the value of a shuffled comparison to minimize noise-related signal.[32] We compare Euclidean distance, the dot product, covariance, and Pearson's product-moment correlation coefficient (abbreviated in figures "euc", "prod", "cov", and "corr", respectively) for their merit as spectral similarity functions for DTW by using these functions. Besides being the similarity function most frequently used in DTW, Euclidean distance has been shown to be interpolatable between sampling points, a desirable property for infrequently sampled data.[39]

The covariance is the mean corrected dot product and correlation coefficient the standard deviation corrected covariance. OBI-Warp caches factorable values in the calculation of covariance and correlation coefficient (i.e., the mean and standard deviations) so that they are only computed once per spectra. With this implementation, $X \times Y$ spectra comparisons for these two measures of similarity take on the algorithmic equivalence of calculating the dot product. Figure 2 provides empirical verification of the similarity in performance of these score functions and highlights the difference between the cached and noncached score functions. It demonstrates that the correlation coefficient and covariance take essentially the same time to calculate as the dot product. We note that the use of very small warp path radii may decrease the benefit of the caching somewhat.

To test these similarity functions for merit in correctly aligning MS spectra, we apply each of them to the optimization of a local weighting scheme and a global gap penalty function across

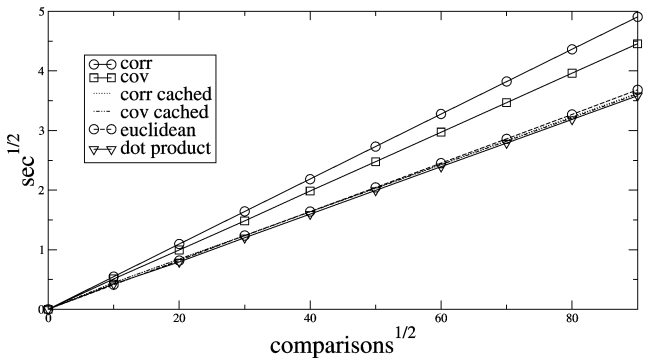

Figure 2. Comparison of spectra similarity function speed. Plot demonstrates the equivalence of the cached correlation coefficient (corr) and covariance (cov) compared to the dot product and Euclidean distance. Each data point represents 10 computed $n \times n$ comparisons on scans of length 10 000.

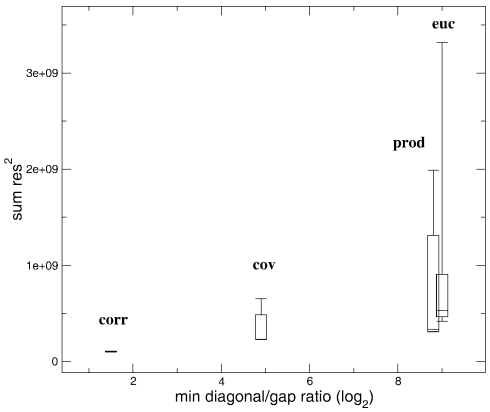

Figure 3. Local weighting distributions by minima. Diagonal/gap (D/G) ratios ($\log_2$) from 0.0 to 9.9 by 0.1 increments were summed across all data sets and all time increments to create a distribution of SSR scores. Candlesticks represent the min, first quartile, median, third quartile, and max of these distributions. Each distribution is centered on the diagonal/gap ratio of the minimum of the distribution. Based on its minima and distribution, correlation coefficient (corr) clearly outperforms the other similarity functions. The optimal D/G ratio (position on the *x*-axis) correlates with the SSR distributions.

variably incremented (3, 6, 12 s) LC−MS runs, in all performing 1 934 400 alignments. Two measures of performance are of particular interest: error with optimal parameters and error across a broad range of suboptimal conditions. By comparing similarity functions across a range of DTW constraints we allow each function to be compared on its own best terms (i.e., at optimal parameters). At the same time, by inspecting performance across different parameters and time increments, we measure a score function's robustness—a good similarity function will amplify the true signal over the noise to give correct alignments not only under ideal parameters but also under less than perfect conditions. The similiarity score distribution for each alignment was normalized to allow unbiased comparison of the different similarity functions using the same range of optimization parameters.

Local weighting optimization was performed without the application of a global gap penalty. In all, 186 comparisons (62 comparisons at 3 time increments each (3, 6, and 12 s)) were tested against 100 local weight ratios. The $\log_2$ diagonal/gap (D/G) ratio was varied from 0 to 9.9 by 0.1 increments. Figure 3 plots the distribution of sum of the SSR scores obtained for the 100 ratios for each similarity function. Correlation coefficient performs

(39) Aach, J.; Church, G. M. *Bioinformatics* **2001**, *17*, 495−508.

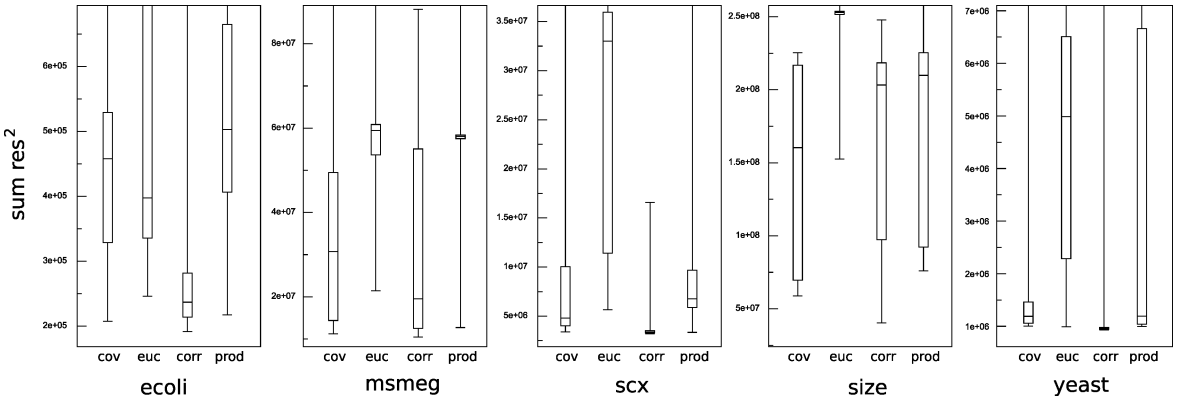

Figure 4. Gap penalty distributions. Candlesticks depict the min, first quartile, median, third quartile, and max of the distributions of SSR scores summed for each gap penalty combination (0.0−14.7 by 0.3 increments for init and elongation penalties). Distributions include all time increments and are plotted by score function for each data set. The corr shows the optimal performance (indicated by the minima) across these data sets and is generally the most robust across the range of gap penalties tested (evidenced by low SSR score distributions).

best under these conditions, with good performance across the range of D/G ratios. The minimum D/G ratio (distribution location on the *x*-axis) is also indicative of the power of the similarity function: a high D/G ratio suggests that the score function needs "prodding" to keep from wandering off the diagonal because amplification of the true signal is insufficient. Thus, the minimum D/G ratios correlate with the SSR distributions.

We also aligned the set of 186 comparisons while varying global gap penalty parameters. We tested all combinations of initiation and elongation penalties from 0 to 14.7 by 0.3 increments and then summed the resulting optimization landscapes to find the global optimum for each score function. Figure 4 shows a quartile summary of these distributions. Following the trend in the local weighting optimization, the correlation coefficient stands out as having the best minima and distributions, followed by covariance and dot product, with Euclidean distance performing poorly. Score distributions for individual alignments confirm this trend, although some exceptions to this ordering do occur. Based on these results, we select correlation coefficient as the default spectral similarity function for OBI-Warp.

Various preprocessing techniques (e.g., baseline correction) may alter these results somewhat, but the ordering of score function performance should, in general, remain unaltered. Future studies could examine the correlation coefficient in combination with the benefit function of Prakash et al.[32] and the quadratic variants used by Stein and Scott.[40]

DTW Constraints. Various constraints may be applied to DTW, including local weighting, slope constraints,[34] and a global gap penalty. Without constraints, DTW has been shown to be "too flexible" for some univariate chromatographic data sets.[25]

The main purpose for a local weighting scheme is to correct for the double score bias of transitions. Assuming identical similarity scores, paths incorporating transitions will receive twice the score of diagonal paths. Without correction of some kind, this phenomenon will produce alignments biased toward the diagonal when minimizing a cost function (e.g., for Euclidean distance) and tending toward transitions when maximizing a benefit function (e.g., dot product). A simple weighting scheme has been used to increase the benefit of a diagonal step compared with that of a transition.[30,33] To compare local alignments (i.e., where $w_1 \neq (1,1)$ or $w_K \neq (|X|,|Y|)$), the total alignment score is normalized by the sum of the weighting coefficients used.[33]

To examine the influence of different score functions across a broad range of conditions, we also modulated the local weighting scheme (see Figure 3).

Although DTW has traditionally relied upon slope constraints (or rules) to generate optimal alignments,[34] here we use a global gap penalty function such as is used in biological sequence alignment.[41] To relate DTW with chromatography, any change in chromatographic rate will introduce gaps in the DTW alignment. A gap penalty function is used to penalize transitions relative to the gap length. It achieves the same end as traditional slope constraints—discouraging high rates of change in the alignment path—but it may offer more flexibility. A global gap penalty may be preferable to slope constraints for a number of reasons: a global gap penalty (1) encourages the warp path to disregard instances of local noise, preferring transitions only when they are significantly better than the diagonal (as determined by the function), (2) allows very disjoint alignment segments (i.e., large transitions) if there is enough evidence to support them, and (3) can be finely adjusted for the alignment task at hand. The optimal global gap penalty is influenced by deviations from a perfect alignment and noise in the alignment signal (see Supporting Information Figure 2). To effectively test the gap penalty, we used relaxed slope constraints—allowing vertical or horizontal transitions (even following one after the other). Athough OBI-Warp allows for any arbitrary gap function (requiring slightly more memory to keep track of gap sizes), here we tested a linear gap penalty with a separate parameter for initiation.

The use of a global gap penalty (at a set local weighting D/G ratio of 2) was found to be comparable or superior to the use of a local weighting scheme alone in almost all cases (see Table 2). Optimal initiation and elongation penalties determined across all 186 comparisons are shown in Figure 5. From Figure 5A it is evident that the initiation and elongation penalties are anticorrelated. The data suggest that better score functions prefer less stringent gap penalties (with Euclidean distance being the pronounced exception). The optimal gap penalty shows moderate

(40) Stein, S. E.; Scott, D. R. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 859−66.

(41) Gotoh, O. *J. Mol. Biol.* **1982**, *162*, 705−8.

Table 2. Percent Difference between Minima from Local Weighting and Global Gap Penalty Optimizations[a]

	ecoli (% diff)		msmeg (% diff)		scx (% diff)		size (% diff)		yeast (% diff)	
	AAD	avg res^2	AAD	avg res^2	AAD	avg res^2	AAD	avg res^2	AAD	avg res^2
corr	51.1	117.5	0	3.6	36	124.6	40.3	31.3	-0.7	-0.4
cov	108	733.6	-12.8	-8.8	366.7	2670.3	20.6	98.7	16.3	29.6
prod	219.7	6926.1	-6.7	3.9	399.5	2883.1	28.5	108.4	25.4	43.7
euc	18.5	21.4	73.7	320.5	159.7	684.2	35.5	84.8	21.3	23.2

[a] Calculated as 100(LW − GP)/GP. In all cases except three, the global gap penalty outperforms the local weighting alone, as indicated by positive AAD values and the average squared residual (avg res^2) values.

levels of clustering across the four similarity functions. Since optimal gap penalties often had different initiation and elongation penalties, future work on nonlinear gap functions may be useful. Figure 5B shows that optimal gap penalties vary somewhat between different run types across the four similarity functions. Figure 5C compares the optimal gap penalties found using the two different measures of alignment accuracy, SSR and the average absolute time difference (AAD), for each similarity function. SSR penalizes big mistakes more than smaller ones, while AAD weights all time differences equally and reports the accuracy in units easy to understand (seconds between eluting peaks). We excluded the size data set from this calculation since it represents an exceptional case and its distributions and minima were extreme compared to the other data sets and significantly skewed the final results. In all cases, AAD favored heavier gap penalties. The default gap penalty parameters in OBI-Warp are set to the optima discovered here using SSR as the accuracy measure ([init, elong] corr 0.3, 2.4; cov 0.0, 11.7; prod 0.0, 7.8; euc 0.9, 1.8).

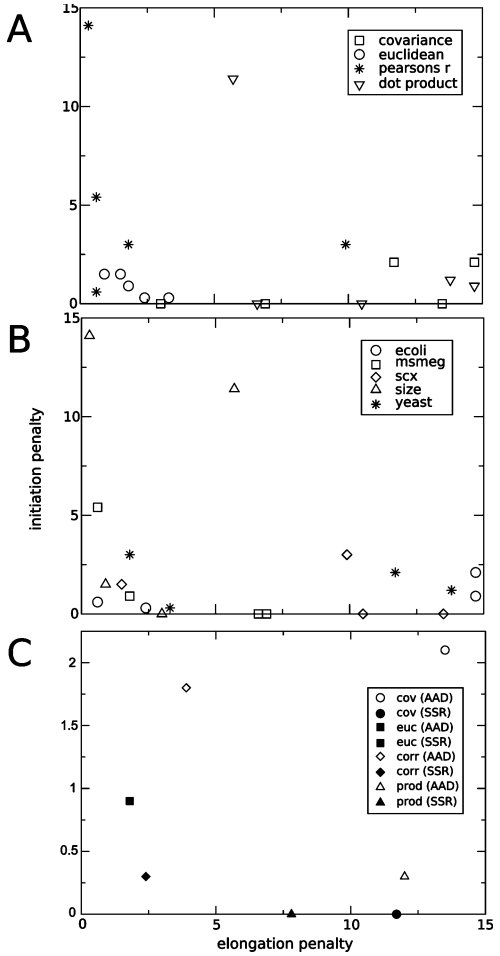

Figure 5. Global gap penalty optima. Each symbol represents the minimum elongation/initiation parameter obtained. SSR scores are summed within each data set by similarity function and (A) shown by score function and (B) shown by data set. (C) depicts the minima obtained by summing across all data sets with both SSR and AAD per similarity function. Anticorrelation between initiation and elongation penalties is evident. Score function minima tend to cluster somewhat. There is less, but still evident, clustering by data sets across the different score types. Optimization using the AAD measure of error selects larger gap penalties than the SSR.

The choice of an optimal gap penalty is related to signal/noise of the true alignment path, the frequency of chromatographic variation, and the extent of the variations. Algorithms to estimate these parameters from a given similarity matrix could produce gap penalties tailored to a given alignment scenario.

Time Increment Comparison. Analyte in an ESI experiment is infused continuously but sampled discretely by the mass spectrometer in time; however, the sampling frequency may be varied by interpolation of the discrete signal. Since the unit upon which DTW and associated constraints act is a single spectrum, the sampling frequency (i.e., the time increment of interpolation) may influence the accuracy of an alignment for a given gap penalty. We examined the influence of sampling frequency on the accuracy of alignments. Figure 6A suggests a slight bias toward larger increments when viewed across all data sets and global gap penalties tested above. However, Figure 6B shows that, for the correlation coefficient, it is the 3-s increment that is favored over longer increment times. This may be attributed to the range tested for gap penalties: a significant subset of the initiation and elongation parameters are high (to include parameter space for all the score functions), and smaller increments would allow the warp path to transition to correct paths more easily than large increments under stringent conditions (e.g., the warp path may make lots of single step transitions even with a large elongation penalty).

Bijective Interpolation. In some applications of DTW including chromatography, a process termed "synchronization" is used to apply the DTW mapping to one or both data sets in order to bring them into register one with another.[22,25] In symmetric synchronization, transitions are dealt with by duplicating the response of the series whose index remains unchanged. With this approach, however, the units describing a time series become inapplicable to the warped run, and the length of the warped signals is increased to the length of the warp path. In asymmetric

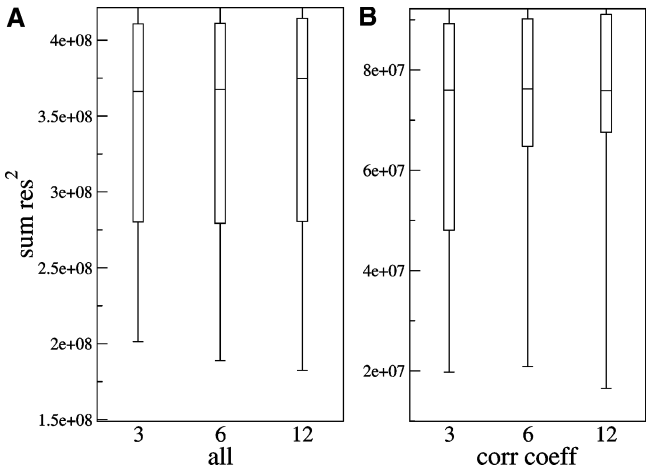

Figure 6. Effect of sampling frequency (*x* axis (in seconds)) on alignment accuracy. (Left) SSR across all score types and data sets. (Right) response to varying time increments across all data sets when using correlation coefficient as the spectral similarity function. Across all score types, a small bias toward runs incremented with larger steps is evident while the opposite seems to be the case for the correlation coefficient.

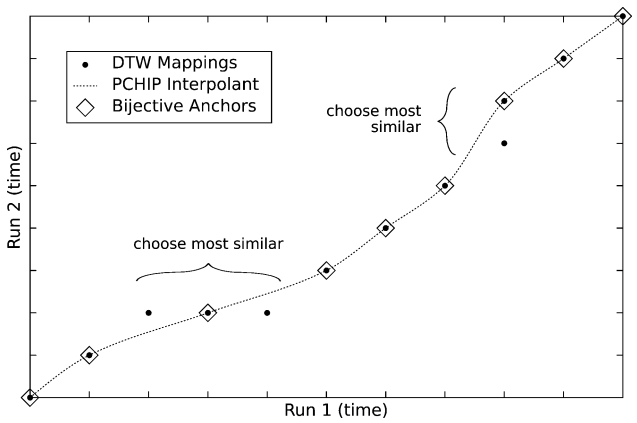

Figure 7. Bijective interpolation. The dynamic time warp path is discrete and not one-to-one. To create a bijective (one-to-one) mapping, all diagonal points and the point of highest similarity in each transition is included. The resulting map is PCHIP interpolated into a continuous warping function.

synchronization, one run is considered the reference and the other run is warped to fit the reference. Multiple points in the reference corresponding to a single point in the other are treated as above, duplicating the points in the nonreference sample. However, when multiple points in the nonreference correspond to a single point in the reference, the average of these points is taken. The run warped by asymmetric synchronization takes on the same length as the reference. A synchronization method where segments with more points are interpolated to have the same number of data points as the reference has also been suggested.[25]

In initial investigations, we found that the use of asymmetric synchronization as the basis for combining a series of runs gave unnatural emphasis to transition areas (data not shown). A desire to achieve more natural synchronizations and produce mappings that could be applied symmetrically (e.g., for use in the alignment of multiple SCX fractions) led to the development of an algorithm to produce a one-to-one (bijective) mapping from the many-to-many warp map produced by DTW and a smooth interpolant for warping either run. The one-to-one mapping includes all bijective points from the traditional DTW mapping and creates bijectivity by including a single point of greatest similarity per nonbijective region (dropping other, less similar points in a one-to-many transition area). In areas with adjacent horizontal and vertical transitions, this gives the effect of rounding off the corners.

We use monotone PCHIP to create a smooth warp function from the bijective warp path. PCHIP ensures that the interpolant will not exceed the extrema at each change in monotonic direction, avoiding some of the "wiggles" or extremes seen in other interpolation methods. It was designed to give a reasonable interpolant for data with both steep and flat sections,[38] conditions frequently observed in chromatography. In practice, we observe this approach to give natural, conservative interpolations consistent with actual chromatographic variation. In this study, we included all bijective mappings in the interpolant. The final bijective interpolant is a warping function that conservatively strings together the bijective mappings. Figure 7 shows a DTW warp path, the bijective warp anchors, and the smooth PCHIP interpolant through these points. The interpolant in this example uses all possible bijective anchors just as in the current study.

Interpolated bijective synchronization offers several benefits: (1) Either run may be warped to fit the other. This may be especially useful in situations involving the alignment of multiple runs where no single run can be considered the reference. (2) The warping preserves the essence of the nonlinear changes suggested by DTW without its drawbacks. (3) Points of greatest similarity become interpolation anchors in transition areas, whereas in traditional DTW, all points in a transition are considered equivalent (though they likely are not). (4) Warped time series are "natural" in appearance and are likely better approximations of most chromatographic variation. Potential shortcomings of the method would include instances where chromatographic variation took on a fully discrete form; these will be modeled as rapid, but smooth transitions, where the slope will be proportional to the length of the transition—nearly, but not completely, capturing the discrete form.

When normalized by the path length, the DTW traceback score provides a means for comparing the similarity of runs. This score is based on similarity in analyte signal less differences in chromatography as modeled by the gap penalty. The smooth warp function allows the deconvolution of these factors. Warped runs may be judged similar in composition (i.e., in the substance of the spectral signal) by any applicable metric (e.g., covariance) apart from chromatographic differences. Chromatographic differences can be measured directly from the warp function itself (as viewed from the diagonal). For instance, the integral of the warp function gives a measure of total chromatographic difference while the derivative gives a measure of chromatographic variability.

Comparison With ChAMS. The recent study by Prakash et al. used dynamic time warping to align spectra from complex proteomics samples. Their method uses a benefit function designed for use with mass spectrometry proteomics: a fuzzy dot product based on mass spectra resolution and a noise estimating parameter. They attempted to reduce the influence of noise on alignments by using adjacent scans to influence the similarity score of the scan in question, reasoning that true signals are preserved over time. We ran each comparison used here through the ChAMS server. To compare the results to ours in a quantitative

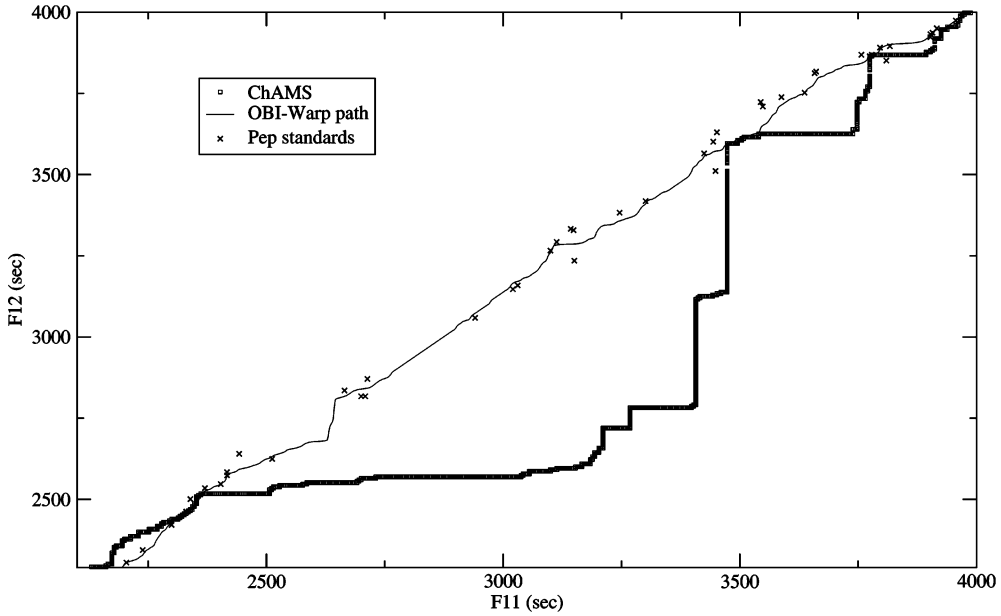

Figure 8. ChAMS misalignment. In stretches of low similarity—but with confident MS/MS identifications—ChAMS often transitions. OBI-Warp successfully aligns these areas, due in part to local weighting and a global gap penalty.

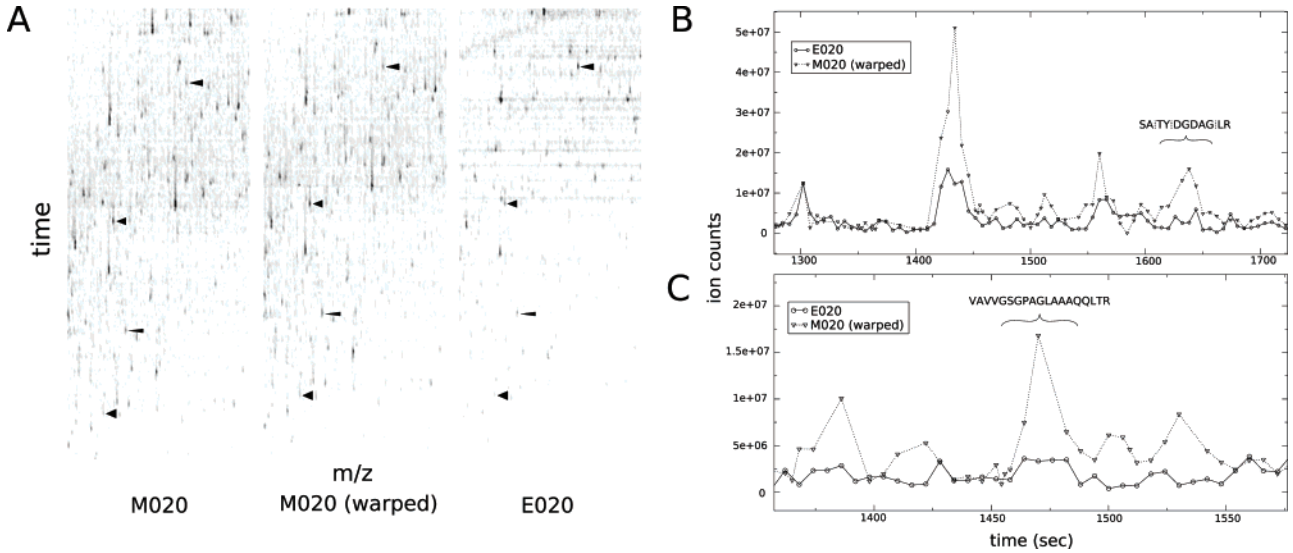

Figure 9. Example alignment. msmeg early and msmeg middle fractions (E020 and M020, respectively) were aligned using optimized parameters. (A) Section of the global alignment. All three plots share the same axis dimensions: *m*/*z* 829−1153 and time ~860−2500 s (of ~5100 s MS runs). Triangular markers to shared peaks are provided as points of reference; markers between M020 (warped) and E020 are on the same horizontal line. (B) A citrate synthase peptide not identified in E020. XIC is from *m*/*z* 733 (+2 peptide charge). (C) A peptide from an NADH-dependent glutamate synthase. No peptides from this protein were identified at any confidence level in the msmeg early runs. XIC is from *m*/*z* 884 (+2 peptide charge).

fashion, we applied our bijective interpolation algorithm to their warp path and calculated SSR from the MS/MS time standards used here. We alternatively chose anchors based on the points of most and least similarity and interpolated linearly and using PCHIP to give four possible interpolants. We note that Prakash et al. used the ChAMS-generated warp path to successfully pick related peaks and that using their results for bijective warping goes beyond the scope of their work. In addition, since four of these data sets were collectively used to determine OBI-Warp gap penalty parameters, these represent a biased test set. However, the size data set was held back from determining gap penalty parameters and represents a fair test (for difficult alignments). In these comparisons, we use the default parameters determined above (correlation coefficient, the optimized gap penalty parameters (init 0.3, elong 2.4)) and runs interpolated to 6-s intervals. The actual warp paths, interpolated paths, and time standards may be compared for all alignments (see Supporting Information Figure 1). Supporting Information Table 1 gives the SSR and AAD scores for each alignment and indicates that the OBI-Warp alignments are comparable to or better than those derived from the ChAMS warp paths.

The ChAMS alignments are nearly identical to OBI-Warp's in runs and sections within runs with medium to high similarity. However, as demonstrated in Figure 8 on late SCX fractions, in runs or sections with weaker signal—but still having legitimate MS/MS identifications—ChAMS often wanders from the true path.

Several factors may be responsible for this outcome and are discussed in likely order of influence: (1) *Gap Penalty.* ChAMS makes no correction for the double benefit gained from transitions. Thus, in instances where the true signal is weaker, the benefit from transitioning often outweighs the benefit of following the correct signal. Furthermore, ChAMS implements no global gap penalty. Our study shows that a modest gap penalty, in addition to doubling the benefit of stepping diagonal, is beneficial to recovering correct alignments, probably by persuading the warp path to ignore small amounts of distracting noise. (2) *Minimal Similarity Cutoff.* ChAMS sets a 0.2 score threshold before including a point in the warp path. As it turns out, some of the runs used here had few data points passing this criterion. To be fair, comparisons were only made on time standards falling within the upper and lower bounds of the ChAMS warp path. This still left some legitimate time standards (high-confidence MS/MS IDs) without mappings in ChAMS, and so many of the runs contained interpolated sections that had little to no guidance from the actual warp path. In some instances, these sections made disproportional contribution to poor scores. (3) *Similarity Function.* It may be that the correction for mean and standard deviation offers some advantage over the ChAMS benefit function. A similarity function such as the one ChAMS uses, but with added mean and standard deviation correction, might be superior to either individually. Comparison with ChAMS highlights the importance of constraints in DTW, especially between runs with low alignment signal.

Multiple Alignment and Example Application To Identify Differentially Expressed Proteins. Chromatographic alignment can be useful, or even critical, across a variety of mass spectrometry proteomics experiments. For single chromatographic dimension runs (as is typical of current biomarker studies), OBI-Warp may be used to align each successive analysis of a particular specimen type to some base run or some meta-run formed by summing the mass spectrometry signals across the aligned experiments. Then, identities may be extended to peaks across the data set and quantities extracted from each run. Further, OBI-Warp may allow multidimensional chromatographic separation for these experiments, since multiple dimensions may be aligned, signal summed, and compared with other analyses. OBI-Warp can be used in a similar fashion to extend the reach of isotope labeling experiments. Successive runs may be aligned, peak identities extended to unknown isotopic doublets, and these quantities extracted and compared with those from other runs.

To demonstrate OBI-Warp's suitability for multiple alignment scenarios, 14 runs were aligned in succession with each run being aligned to the warped form of the previous (and then warped itself). Such a test likely exceeds the requirements of any typical multiple alignment (e.g., each run aligned to the same template). ChAMS alignments, postprocessed to create warp functions, were applied to the same set for comparison. The time differences for OBI-Warp shown in Table 3 compare favorably with the direct alignment of two runs and reveal little propagative error.

Finally, to illustrate the differential identification of peptides using OBI-Warp, we present an alignment of biologically varying samples. Figure 9A depicts a segment of the alignment of the 020 fractions of msmeg early (referred to as E020) and middle (called M020) using default parameters and a 6-s time increment. These

Table 3. Transitive Error Measured as the Average AAD in Seconds of Warped Peptide Standards from Each Run Back to the Original Template Run (4-03-03)[a]

	AAD (sec) between 4-04-03 and last aligned		
Alignment	None	OBI-Warp	ChAMS
4-03-03 :: 7-25-03 (Direct)	149.87	11.56	164.19
4-03-03 :: 6-17-03	242.72	19.25	43.35
6-17-03(warp) :: 6-18-03	306.58	27.94	49.06
6-18-03(warp) :: 6-28-03	258.04	22.47	43.78
6-28-03(warp) :: 7-11-03	216.67	21.6	35.81
7-11-03(warp) :: 7-13-03	206.52	21.59	44.7
7-13-03(warp) :: 7-17-03	179.59	19.07	38.88
7-17-03(warp) :: 7-19-03	176.46	17.87	36.59
7-19-03(warp) :: 7-20-03	185.82	18.11	30.23
7-20-03(warp) :: 7-21-03	212.69	23.1	48.86
7-21-03(warp) :: 7-22-03	171.45	21.2	40.81
7-22-03(warp) :: 7-23-03	433.96	104.53	92.32
7-23-03(warp) :: 7-24-03	152.15	52.45	58.46
7-24-03(warp) :: 7-25-03	149.87	**38.81**	**72.14**

[a] The transitive error accumulated through 13 alignments and warping compares favorably with a direct alignment for OBI-Warp. ChAMS alignments were postprocessed to create bijective interpolants for warping so its results could be used for comparison.

samples are derived from *Mycobacterium smegmatis* cells harvested early in log-phase growth and from mid-log phase cells. The 100 high-confidence peptides (as defined above (including the criteria for duplicate peptides)) found in M020 were not found at any confidence in any of the msmeg early data sets. We highlight two peptides from this set that belong to proteins directly involved in metabolism (protein identities were assigned by sequence similarity to *Mycobacterium tuberculosis*). Figure 9B shows the extracted ion chromatogram (XIC) of a peptide from citrate synthase, the enzyme responsible for the first step in the citric acid cycle. The surrounding XIC appears to be correctly aligned for neighboring peaks. There are 27 peptides in this high-confidence set where the entire protein (at any confidence) is missing from all msmeg early fractions. Figure 9C shows the alignment of one of these peptides, which is derived from a probable NADH-dependent glutamate synthase (small subunit). This example demonstrates the potential for OBI-Warp to provide identities and quantities for differentially expressed peptides and proteins.

CONCLUSIONS

In this study, we have verified the ability of dynamic time warping to correctly align ESI-LC-MS runs of varying similarity, both between samples differing due to biological variation and between those differing due to prefractionation. We show that Pearson's correlation coefficient, followed by covariance, dot product, and Euclidean distance, achieved the most correct alignments under the widest set of gap penalty parameters. We optimize gap penalty parameters and show that a global gap penalty function generally outperforms a local weighting scheme alone. We demonstrate the utility of bijective interpolated synchronization for delivering smooth, natural warpings based on the discrete DTW warp path. We compare our results with those of a recent, independent implementation of DTW and, because of the use of local weighting, a global gap penalty, and other factors, find our implementation to be comparable or better across a series

of alignments. Finally, we demonstrate OBI-Warp's utility for multiple alignments and present a case of using OBI-Warp to identify differentially expressed proteins between bacteria harvested from two different growth conditions, illustrating the potential of LC−MS alignment for differential proteomics.

ACKNOWLEDGMENT

We thank Jason Davis for suggesting the caching of score functions. We thank Parag Mallick for helpful discussion on various alignment methods. We thank Benno Schwikowski for helpful discussion on the application of dynamic time warping to MS samples. We thank Daniel Miranker for helpful discussions on score functions, mass spectrometry, and dynamic time warping. This work was supported by grants from the National Science Foundation, the National Institutes of Health, the Welch Foundation (F1515), and a Packard Fellowship (E.M.M.).

SUPPORTING INFORMATION AVAILABLE

(1) Figure 1: visual comparison of OBI-Warp and ChAMS alignments. (2) Figure 2: factors influencing the choice of an optimal gap penalty function. (3) Table 1: ChAMS comparison. This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review March 23, 2006. Accepted June 28, 2006.

AC0605344