

Welcome to the TechTalk feature of US HUPO News. TechTalk will be an ongoing feature in this newsletter and on www.USHUPO.org. TechTalk is intended to provide insight into a specific technique and then to generate online discussion via a WIKI at www.USHUPO.org. Share your ideas, questions, and experiences with iTRAQ and suggest future TechTalk topics. We hope to enliven and to create a resource for the Proteomics community with this feature.

Go to [US HUPO WIKI](#) to comment on this TechTalk.

Special thanks to Darryl Pappin for his contribution to our inaugural TechTalk.

Enjoy - Josh LaBaer and the US HUPO Newsletter Committee

An iTRAQ Primer

Darryl Pappin, Cold Spring Harbor Laboratory

This is not a review. It is also deliberately brief on background. What follows is a collection of tips and tricks to help get the best from using iTRAQ in quantitative proteomics experiments. These have been assembled over time and with experience using these reagents, by ourselves and many others. Seasoned iTRAQ users will be fully aware of many of the issues discussed below (and will have applied their own workarounds). New or prospective users will hopefully find the following useful. In places, I have added my own personal thoughts and biases as to which steps are important, why they are important, and which are of lesser consequence. Quite a lot of what follows is not in the manual!

A brief retrospective:

The modern proteomics researcher is faced with an ever increasing (and somewhat confusing) set of tools for quantitative proteomics. These include ICAT, SILAC, iTRAQ and label-free methods, to name a few. For the labeling methods, I think it more convenient to classify them all in only two groups: mass-difference or isobaric. It is all about the introduction of heavy isotope labels into your analytes (in this case peptides or proteins). The fact that this is achieved metabolically (SILAC) or by chemical reaction (iTRAQ) is less important, and often down to personal choice or particular experimental constraints. Most conveniently, analysis of quantitative information from these two different classes is essentially the same for any member of the group. In software, isobaric reporter-ion methods such as iTRAQ and TMT are handled identically, as are mass-difference approaches using SILAC, ICAT and ICPL.

iTRAQ was conceived as one possible way to solve the problem of sample multiplexing. Put simply, how might we analyze multiple samples with less analytical load. If your experimental question can be effectively answered by a simple A versus B binary comparison (e.g. control versus treated, or control versus a single time-point) then any of the quantitative methods will probably give useful data in the hands of expert practitioners. Which method you choose comes down to other factors, such as expertise, time and cost. Reagents in the iTRAQ family really start to show distinct advantages if you can use

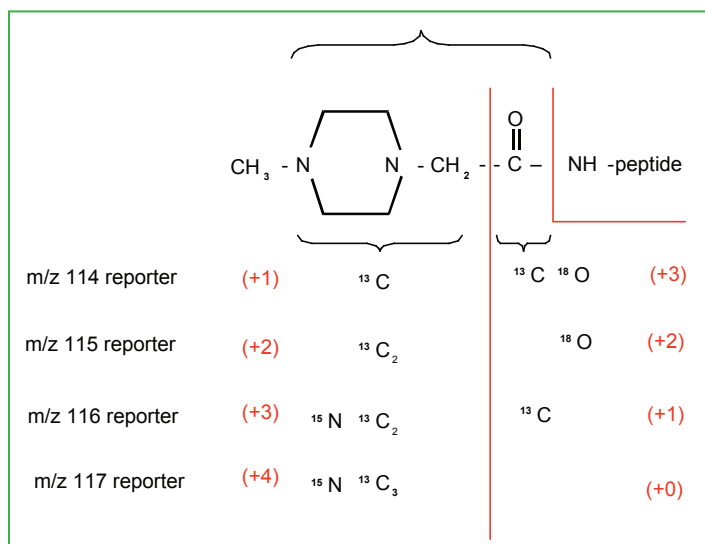


Figure 1. Structure of the iTRAQ 4-plex reagent used for multiplexed measurement of protein expression levels. This isobaric peptide tag is designed to fragment in such a way that mass information from the reporter section of the molecule can be resolved and used to measure relative protein levels in up to four samples at once. The isobaric reporter/balance character of the molecule is achieved by inserting isotopically enriched atoms of carbon, nitrogen and oxygen at various points in the structure. Reporter ions are generated by sequential cleavage of the tag-peptide amide bond by CID, followed by neutral loss of the carbonyl (C=O) balance moiety.

them in a multiplexed manner. This may be a time-course or comparison of multiple sample states or samples (e.g. patients). For my part, iTRAQ was also conceived and executed primarily as an interesting academic exercise. It is about making bonds break in a predetermined order, and making charge go where you want so that you get to see the simplest set of fragments [Fig 1].

How to get the best from iTRAQ:

As with any analytical method, iTRAQ has strengths to use and weaknesses to mitigate. A principal strength is the ability to achieve efficient sample multiplexing without increasing MS or MS/MS complexity. A disadvantage is that all peptides are labeled with the same chemical tags, so overlap of precursors with similar or identical m/z will lead to merging of iTRAQ reporter ions from different peptides and incorrect calculation of intensity ratios. This

was understood in principal from conception, but only more recently discussed in a more thorough analytical framework [1]. In very complex sample mixtures, where precursor overlap or interference is often high, this will lead to fold-change underestimation (for both up- and down-regulation), although the direction of change is correct. There are a number of interesting and ingenious ways that have been proposed to correct for precursor overlap and background [For a review, see 2]. *My personal bias is to try to reduce potential problems in the first place.* Steps can be taken in three areas:

Sample preparation and the iTRAQ reaction:

More separation is always better. For very complex samples, such as a profiling shotgun approach on whole cell lysates, a 2D separation is an absolute requirement (at the least). Anyone attempting to use iTRAQ with a 1D separation on such samples is wasting their time and/or not realizing what they are missing. In-source suppression that reduces the signal intensity of mixed components is often massively underestimated, and the effects of precursor interference or overlap are significantly increased. For the first dimension of separation, increasing the number of SCX fractions is the simplest and most practical approach. Using the OFFGEL device, 24 fractions will always yield more proteins than 12. In classical MudPIT, 12-16 salt-step gradients are needed. For the second dimension, gradient profiles can be optimized and lengthened (2 hours or more). Longer capillary analytical columns (20cm or greater) with finer particle size (<5um) will improve chromatographic separation and reduce sample complexity being sprayed into the MS, albeit at the cost of increasing column backpressure. Of course, in a busy core laboratory one has to balance time available for any given sample against overall sample throughput, and compromises must often be made.

Get the protein concentrations right, and only use the right buffers. The iTRAQ chemistry is pretty immune to abuse (salts, detergents etc.), but is quenched very rapidly by primary (Tris) or secondary amine buffers and to some extent by free thiols (e.g. DTT). Amazingly, the iTRAQ manual is stuffed full of very useful tables that detail reducing agent, buffer, detergent and salt (in)compatibilities. *Please read!* Getting your total protein measurements right will not only make sure you use appropriate concentrations with the standard iTRAQ protocol, but also means that any later global bias or normalization corrections will be small.

Keep the alcohol concentration up! The iTRAQ reaction is typically performed in ~70% v/v alcohol for a reason: to control aqueous hydrolysis. As with all active esters, and especially in the case of basic compounds like iTRAQ, you

are balancing rapid hydrolysis with water against reaction with amines (aminolysis). In pure water at ~pH 8, the half-life of iTRAQ is around 15 seconds (i.e. no good to anybody). In 70% v/v alcohol, the half life increases to 15 minutes, which is eminently useable.

Don't think you can swap to acetonitrile. At 70% v/v this can severely suppress the reaction rate, which this solvent can do for many nucleophilic reactions. Alcohols are very good, as aminolysis is often increased from that in water, and they are good peptide solvents. If you really want to change things, then longer-chain alcohols such as isopropanol will be fine (e.g. to improve solubility of hydrophobic peptides or proteins), and even aprotic water-miscible solvents such as DMSO and DMF may be useful, if more difficult to remove. Seasoned chemists might happily use 70% aq. pyridine which, apart from being a great peptide solvent, is also its own buffer.

Buffers and more. There is also a good reason why the concentration of reaction buffer is kept at 120-150mM in the reaction mixture. As the iTRAQ reagents hydrolyse, the end products are N-methyl piperazine acetic acid and N-hydroxysuccinimide. Without sufficient buffering capacity, the solution pH will rapidly fall below ~pH 8 and the reaction with peptide amines will stop as they protonate fully. Triethylammonium bicarbonate (TEAB) was specifically chosen as a good tertiary amine buffer which is also very volatile and can therefore be easily removed if required. *Everything is there for a reason.*

Solubility and chromatography. iTRAQ labeled peptides are larger and can be more hydrophobic than their underivatized, native peptide counterparts (especially for the 8-plex labels). Thus, one should always be aware of possible solubility issues. Simple steps, such as increasing the ACN content in LC loading solutions to 5% /0.1% formic acid or greater can reduce the loss of poorly soluble peptides, particularly when some samples may sit for 12 hours or more in a cooled autosampler vial before injection! Chromatographic resolution of labeled peptides can be improved by steepening LC acetonitrile gradients (by ~10% at the final %B concentration) to counteract increased hydrophobicity.

Salt, salt and more salt. Often poorly understood is the effect of residual iTRAQ reagents on SCX or OFFGEL fractionation. For both methods, the total salt concentration must be kept below ~10mM at sample loading, or peptides will not bind (SCX) or focus (OFFGEL). Following the iTRAQ reaction, residual buffer (TEAB: ~120mM) and ethanol are removed *in vacuo* (as designed). The remaining iTRAQ reagents (principally N-methylpiperazine acetic acid as the hydrolysis product) are not volatile, and are present

at around 40mM concentration in the standard reaction volume. These must be removed, for example by C_{18} cartridge or stage-tip cleanup, before SCX or OFFGEL loading.

Data Acquisition:

Fragment harder. iTRAQ labeled peptides are more stable to CID than their native counterparts. For a labeled peptide of any given mass, it is advisable to increase collision energy slightly to give equivalent fragmentation and increase the intensity of reporter ions. This is beneficial, as it reduces the effects of weak ion statistics that increase the scatter of iTRAQ ratio measurements. As a general rule, increasing the collision energy or voltages by 10-15 % relative to the native peptides is sufficient, although more energy may be applied to give stronger reporter signals.

Use tighter precursor selection tolerances for MS/MS to minimize precursor overlap. This can be optimized for any given instrument, and is a compromise between using a narrow isolation window without losing too much signal sensitivity.

2+ is better than 3+. Fragmenting the $[M+2H]^{2+}$ ions of any given peptide will almost always give higher intensity iTRAQ reporter ions relative to backbone fragments than the $[M+3H]^{3+}$ (or higher) ion of equivalent intensity. If your acquisition software allows you to set precedence to the lower charge-state ions, then do so (unless the 3+ ion is massively more intense).

Data processing:

Use the best. Use your better peptides for calculating protein-level iTRAQ ratios (e.g. only use peptides with expect values >0.05). Ratio calculations may also be applied only to unique peptides – those peptides which are *not* shared by members of closely related protein families (even though this can significantly reduce the number of measurements available for ratio calculations).

Apply intensity weighting to calculation of average iTRAQ ratios and statistics. In this way, more intense spectra contribute more to the final calculated ratio, and background contributions are lessened. A very convenient way to do this is using the summed absolute intensities of iTRAQ reporters in any give channel (114,115 etc.) from peptides that map to any given protein. This is a natural match to the way a mass spectrometer collects data by averaging or summing together a series of scans or spectra.

Apply appropriate normalization or bias corrections. Only apply global normalizations with large data sets (e.g. >1000 proteins) where statistics are on your side. Again, this can be very conveniently calculated by first summing all the absolute reporter ion intensities in each iTRAQ channel. Simple corrections are then applied such that the global ratios are corrected to be unity. If you got your protein measurements right in the first place, these corrections should be small.

Correct for isotope overlap. Use the supplied correction factors to compensate for overlapping isotope contri-

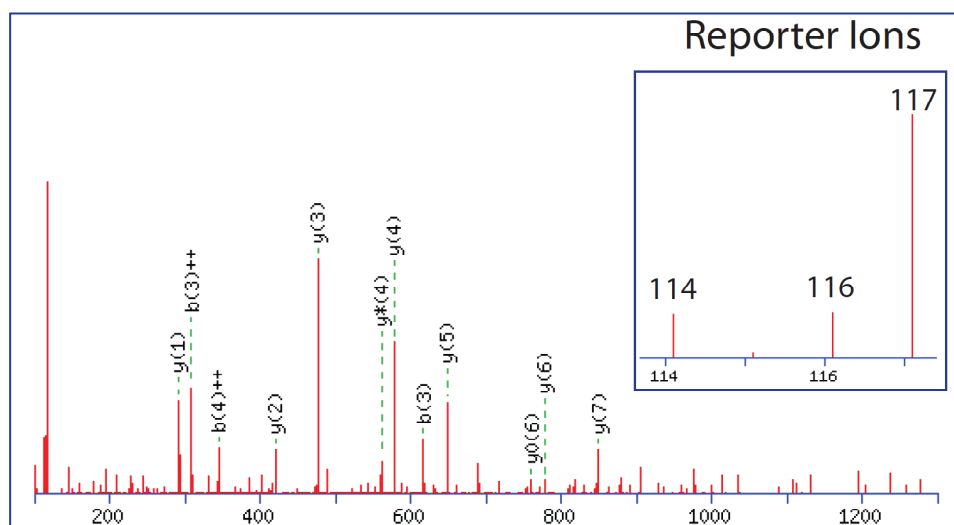


Figure 2. An example of a protein showing significant up-regulation is shown. The peptide was identified with high confidence as evident by an almost complete y ion series ($p=0.001$). The 117 reporter ion signal is the most intense ion in the MS/MS spectrum. The average fold-change of upregulation (from 5 peptide measurements) was 7.7.

butions, both from natural isotope abundance (+1, +2 Da) and those that result from incomplete enrichment at any given carbon or nitrogen atom (-1, -2 Da).

How to (quickly) judge significance. A simple approach is to use the global statistics derived from any given iTRAQ channel. For example, one can take all ratio measurements in any given iTRAQ channel for all peptides of $p > 0.05$, and calculate mean and standard deviations. If normalized, the mean will be at or close to unity, and standard deviations can be in the range of 0.15-0.5. As a quick filter, one can judge the significance of fold-changes very broadly using this measured value, and fold changes can be ascribed simply in terms of intervals of SD above and below unity (say $>2SD$ up or down). This quick-and-dirty approach to judging significance cannot replace more detailed statistical analyses [2] but is often extremely useful in giving a qualitative assessment of how well an experiment went.

Things that people worry about but which can (mostly) be ignored. In this last category I include things such as interfering fragment ions that happen to have similar m/z to iTRAQ reporter ions, and correction of some overlapping isotope contributions. Interfering fragment ions are generally sequence specific (thus relatively rare) and in large data sets the effects are negated by averaging

of ratios from multiple peptide measurements or removal by outlier rejection methods. Isotope impurity overlaps are generally of the order of $\sim 3\%$ for the 4-plex reagents, so the corrections would be only be one tenth of typical global standard deviations of $\sim 25\%$ seen in real biological samples. For the 8-plex reagents, the impurity corrections are even less ($\sim 1\%$ on average) and can be ignored.

Figure 2 shows one example from an iTRAQ experiment where all the above tips were applied. This example is from a global proteomics screen of whole-cell lysate from cells treated with short DNA oligos targeted to let-7 microRNA. Cells were lysed, digested with trypsin, labeled with iTRAQ in a 3-plex experiment and fractionated by high-resolution (24 fraction) OFFGELL followed by capillary LCMS. More than 200,000 MS/MS spectra were collected, with some 2,500 proteins identified and quantified. Treatment with let-7 antagomirs caused significant increase in the expression of on-target proteins, in this example by almost 8-fold.

References:

- [1] Ow, S. et al. (2009) *J. Proteome Res.* **8**, 5347-5355
- [2] Karp, NA. et al. (2010) *Mol. Cellular Proteomics*, epub ahead of press, April 10th 2010