

Understanding Oligonucleotide Synthetic Yields

By Richard Hogrefe, Ph.D.; TriLink BioTechnologies

Most oligonucleotides are ordered by the scale instead of the quantity, particularly the smaller scales of 15 μmole and lower, with no guarantee of final yield. This is because every oligonucleotide is a unique molecule, with unique synthetic properties based on sequence and modifications. The quality of a synthesis can be affected greatly by the sequence, such as G rich sequences. It is not easy to predict how much material will result from any particular synthesis. The price of the oligonucleotide is based on the amount of material and labor needed to prepare that synthesis based on scale. We can either guarantee only a small amount from any synthesis to ensure full delivery, or send the customer the entire resulting yield (which is what TriLink and most of our competitors do).

There are many factors other than sequence and the attached modifications that can affect yield as well, such as reagent quality and even weather. These topics will be discussed shortly, but first it may help to have a brief overview of how yield is measured and reported and how even that process can affect the final yield of a product.

Measuring Up

The yield on your certificate of analysis (COA) is reported in one of two ways. It will be reported either in units of mass or in optical density units measured at 260 nm (OD_{260}). The yields from all larger scale syntheses (100 μmole and above) are reported in units of mass, milligrams or grams. The yields of the small scale syntheses are conventionally reported in OD_{260} units.

All oligonucleotides, even those made at the multi-gram scale, are measured for yield using absorbance. Since most orders are placed at scales yielding 0.05-10 milligram of product, obtaining an accurate weight on routine syntheses is not possible. UV spectroscopy is a much more accurate way to measure the yield of oligonucleotides since it ignores salt, water and other residuals of the synthesis that would affect the weight. The oligonucleotide is dissolved in water or buffer and the absorbance at 260 nm determined. This absorbance number is reported as the yield on the COA as it is the most accurate representation of the yield. The conversion to mass uses an approximation, as will be discussed in the next section, and therefore is not an accurate representation.

Larger scales are generally ordered by mass, not yield, simply because it is the convention. The yields on the COA are likewise reported in units of mass for similar reasons.

Conversions

Although the absorbance is the most accurate representation of the yield, it is by itself not very useful. Almost all reactions are expressed in terms of moles of reagents required. The yield reported as OD_{260} units on the COA can be converted to μmoles using Beer's Law that relates absorbance to concentration using another figure found on TriLink's COA, the extinction coefficient (ϵ), which is a constant unique to every substance.

Beer's Law: Absorbance = [concentration] ϵ
which can be derived to: [concentration] = Absorbance / ϵ

The units of ϵ are (OD_{260} units)(mL) (μmole)⁻¹, the units of concentration are (μmole)(mL)⁻¹, and absorbance is expressed in OD_{260} units. Since the absorbance readings we report are already normalized to 1 mL, the final calculation yields the number of μmoles of product directly, with no further math required.

For example, a reported yield of 100 OD_{260} units for a product with an ϵ of 175 (OD_{260} units)(mL) (μmole)⁻¹ calculates to a yield of 0.571 μmole (100/175).

Unfortunately, this very convenient calculation is premised on a number that is itself obtained by calculation, and sometimes with some guesswork added in. The ϵ is calculated for each and every oligonucleotide sold by TriLink. Different formulations exist for this calculation, the accuracy of each increases as more of the complexities of the system are taken into account.

The most accurate method is to generate the ϵ experimentally, but this is a very long and difficult process. TriLink does the next best thing in using what is called the nearest neighbor model. We take into account the affect on the absorbance of each nucleoside by the neighboring nucleoside, which can be considerable. However, even this fairly complicated calculation is only accurate to within 10%.

The guesswork occurs when some modified nucleosides and other compounds are added to an oligonucleotide. We can only use an educated guess as to what will be the affect of the modification on absorbance. We have determined the ϵ of some dyes we sell experimentally, but those were on unique oligonucleotides that may or may not extrapolate well across the spectrum of sequences.

Therefore, despite using a more complicated formula than many vendors, the ϵ offered on our COA is still only a reasonable approximation of the true figure. Be prepared to vary the concentration of oligonucleotide reagents by as much as 10% in order to obtain perfect results.

If you wish to determine the mass of the oligonucleotide, multiply the number of moles by the molecular weight (found on the COA in grams per mole, or perhaps more usefully, μg per μmole). This number is calculated for the free acid. Be sure to account for the salt form if you intend to actually weigh the samples. This is not reasonable for scales of 1 μmole and less, as the amount delivered makes it impractical.

What affects the yield of your oligonucleotide?

Synthesis

The amount of product theoretically possible from any particular synthesis is determined by the quality of the synthesis itself, generally done on an automated synthesizer. The coupling efficiency of the synthesis is very important. At TriLink we do our best to optimize this efficiency using modified programming and paying special attention to reagent preparation. The results of a drop in coupling efficiency of just a few percentage points on the yield can be devastating. This is easily demonstrated by calculating the theoretical yield with the following formula:

$$x^y$$

where x is the average coupling efficiency and y the number of couplings.

For example, a synthesis of a 30mer (which requires 29 couplings) with an average coupling efficiency of 99% theoretically yields 75% of product (0.99^{29}). That same synthesis at 98% efficiency will have a maximum yield of only 55%. That one percent costs nearly half the material. Consider the difficulty of making a 70mer. Even at 99%, the best one could hope for is 50% yield. At 98% it becomes an abysmal 25% yield.

It is very difficult to maintain every machine every day at an operational efficiency of 99% or greater, despite claims to the contrary. Not only does the instrument have to be finely tuned to operate at 99%+, but other factors have to be perfect, such as the moisture content in the acetonitrile and phosphoramidite quality. The weather plays a role, as well. Extremely humid days will adversely affect the quality of synthesis by making near complete water removal almost impossible, despite using rigorous anhydrous chemistry techniques.

Modified reagents often have poor coupling efficiencies for a variety of reasons. It is not unusual for a reagent to have a coupling efficiency as low as 90%.

Some modified oligonucleotides are prepared by conjugation to an amine, thiol or carboxyl functionality. This reaction is affected by many factors as well, including the quality and age of reagent, the sequence of oligonucleotide and the quality of the buffers. Reactions can range from 10% to nearly quantitative. A mediocre conjugation can have a considerable affect on final product yield. At TriLink we do our best to control the quality of all reagents through our ISO/GMP program, as well as operate with well established SOPs and well trained personnel. Despite our best efforts, some reactions will still go poorly time to time due to sequence effects.

Deprotection

After synthesis, the oligonucleotide is deprotected with base. With simple oligonucleotides this presents little difficulty, as long as the deprotecting reagents are fresh. However particular modifications require deprotecting conditions so mild that the bases are not fully deprotected. In most cases, some damage occurs to the bases themselves, but particularly to dye labeled compounds. These side reactions all result in loss of yield.

Purification

Depending on the quality of the synthesis, purification can be the step where the most yield is lost. A high quality synthesis will have only a moderate amount of impurities to remove, allowing a larger cut of the product peak. Moderate and poor syntheses will have more contaminating fragments that will crowd into the product peak, requiring a tighter cut to obtain an acceptable purity. Regardless of the quality of the synthesis, the overall process of purification is costly to yield. Upwards of 50% of the theoretical yield will be consumed in many preps for a wide variety of reasons.

Oligonucleotides that are partially deprotected and those containing degraded or damaged dyes often share enough properties with the product to make purification difficult. In some cases, the modifications on an oligonucleotide will cause the product to co-elute with shorter fragments, which is the case with unprotected primary amine modified oligonucleotides.

The requested purity makes a considerable impact on yield. The difference in delivered product between a final purity of 90% and 95% can be several fold.

What should I expect from my 1 µmole scale modified oligonucleotide order?

Here is a hypothetical synthesis to illustrate how yield is affected. We will return to the example of the 30mer described above, but add the following: it has a 5' dye and a 3' quencher. This hypothetical dye is only available as a succinimidyl ester, therefore requiring post-synthesis conjugation to a 5' amino labeled oligonucleotide. The 3' quencher in this case is a support bound reagent from which the synthesis begins, requiring no further chemistry. The oligonucleotide is further modified with three modified bases, each of which couple at 93% in this example. The amino linker couples with an efficiency of 95%.

The theoretical maximum synthetic yield based on an average of 99% for the other 26 couplings is calculated thusly:

$$(0.99^{26})(0.93^3)(0.95) = 0.59$$

The deprotection of the oligonucleotide always results in a small amount of undeprotected bases, undesired modifications, and in this case, some degradation of the quencher. It also involves transfers and filtrations that invariably result in most of the loss. Overall a loss of another 25% of potential product due to chemical modifications, incomplete deprotections and manual manipulation is not unusual. Our yield is now at 44%, or 0.44 µmole.

The protecting group on the amino group is left on to improve the purification of the amino labeled intermediate prior to conjugation. Without it, the amino labeled oligonucleotide will elute with the failure sequences on HPLC. It is common to lose 10% of this group prior to purification during deprotection and work-up, lowering our yield to 0.40 µmole.

This product is purified, resulting in another loss of at least 25% of the product due to contamination with the aforementioned undesired modified materials. We are now at 0.30 µmole of intermediate.

The conjugation of a dye generally goes at 70-90% unless there is a sequence or reagent issue. Let us assume an efficiency of 80%. The yield is now at 0.24 µmole.

This material is now re-purified to remove unconjugated material, free dye, and any side products due to damaged dyes, resulting in a further loss of 20% of the product, reducing the product to 0.19 µmole.

This is followed by a series of manipulations to desalt the sample, convert it to the proper salt and precipitate it, and then remove aliquots for final analysis prior to delivery. Unfortunately, this series of steps is more costly than appreciated, resulting in an overall loss of another 20% of the product, reducing the final yield to 0.15 µmoles.

Given a 30mer with an ϵ of 345 (OD₂₆₀ units)(mL) (µmole)⁻¹ and a molecular weight of 10,500 g/m, or 10.5 mg/µmole, this final yield would correspond to a delivery of 52 OD₂₆₀ units, 1.6 mg, of product.

We would consider this an excellent yield for a highly modified compound like this. Yields of 20-40 OD₂₆₀ units for compounds of this sort are much more common. The losses described can easily be compounded by many different factors, some completely out of our control (mostly construct/sequence issues) and some within our control (level of training, routine maintenance). A compound with high guanosine content can cause the loss of at least half of the material if it has secondary structure that makes purification difficult. Also, recall that a day with higher than usual humidity can result in the loss of a third of the product right from the onset of the synthesis by dropping coupling efficiency just one percentage point.

Why is my yield different this time although the order is exactly the same?

We understand how annoying it must be for a customer to obtain different quantities of material for subsequent batches of the exact same oligonucleotide. Much of the answer is in the preceding paragraphs. Subtle changes in one aspect of the synthesis can have grave consequences on the yield. Also, many of our reagents have a single source, upon whom we are dependent for high quality material. If that falters we must make do with what we can obtain.

When should I request a set quantity?

Almost all mid-scale syntheses (50 mg and larger) are ordered by set quantity, usually in units of mass. Some customers order 10 mg samples, depending on us to decide on the proper starting scale.

There are cases when ordering a set quantity for smaller amounts may make sense. Many diagnostic firms require guaranteed amounts of material delivered so that they can produce the allotted number of kits, even at small quantities. This will increase your cost considerably in some cases since we may have to prepare a larger scale to ensure the yield you requested. It is best to order by the scale and trust us to do the best possible under the current circumstances.

A Table of General Expectations

For a general guideline, see tables of expected yields on pages 17 and 18.