

## Instructions for Providing your Sequence

*Before we can perform any analysis of your “known” protein, the following information must be sent to us by email. Incorrect or incomplete information can seriously jeopardize your results, increase the costs and delay the analysis.*

- **Email to us the exact sequence of the mature protein or construct you expect your sample to have.** A *single* amino acid mistake in the sequence you email us will adversely affect the results you obtain. The most common mistake that researchers make is to send us the public database entry when in fact their sample has known differences such as cleaved signal sequences, amino acid substitutions or modifications, tags, isoforms, etc. Research the molecule you are sending and indicate the true amino- and carboxy- termini and any known or predicted modifications.
- **Send the sequence as single letter amino acid code in the text of an email.** Please do *not* send as an attachment. The sequence should be in the body text of the email, no numbers or characters other than the sequence itself.
- **Email the sequence to:** [hmf@harvard.edu](mailto:hmf@harvard.edu)
- **In the same email, indicate project specifics, such as species, construct, vector, or expression system.** We must know all species that are involved. For example, you may have a human gene being expressed in e. coli as a GST fusion. That is three species: human, e.coli and schistosoma.
- **For phosphorylation or other modification site experiments, in the email:**
  1. **Make a clear distinction between (a)** requests to find *any* site of the modification versus **(b)** requests to prove/disprove specific (known or putative) sites of modification. *Our design and analysis strategy are different for each.*
  2. **If specific residues are of interest (b - above),** indicate them in the email.
  3. **If you are surveying for any site (a - above),** please research the literature and indicate what is known or predicted by motif.
  4. Indicate the conditions that may cause your modification, e.g. expected kinase(s), exposure to UV, etc.
  5. If there are tags or relevant domains, indicate the residues at which they begin and end.
  6. Finally, use a residue numbering system consistent with the sequence you are emailing. For instance, if specifying that there is a known modification at residue 200 in the native protein, but you are sending a recombinant with an eight amino acid Nterm tag, be sure to indicate residue 208, not 200.
- **Above all, we need to know what you know.** Anything that is not indicated above but is germane to the project should also be in the email. This is also a good place to indicate items above that you *do not* know, e.g. “we do not know where the exact amino-terminus begins.” If you have received this form without you or your PI having a prep and strategy discussion by phone, *do not send a sample.*

*Thank you.*