Peptide structure

Beyond transcriptional events

Anthony M. Treston and James L. Mulshine

PROTEIN chemistry, the demise of which was predicted in a News and Views article¹ 10 years ago, is still an essential technique for determining protein structure. Robert Anderegg and colleagues in the December 1988 issue² of the *Journal of Biological Chemistry*, for example, report the structure of a yeast mating pheromone using conventional protein-chemistry techniques. Their results demonstrate that, given the present lack of understanding of the mechanisms and control of post-translational processing, the prediction of processing events solely on the basis of nucleic-acid sequence data is speculative in the absence of supporting protein chemistry data.

The traditional approach for determining the structure of peptides and proteins has been through what could be called classical analytical methodology, including purification, derivation, hydrolysis and amino-acid characterization. Recently, however, these approaches have lost popularity because of the speed and precision of sequencing nucleic acids, enabling the prediction of the order of amino acids directly from DNA and RNA sequences. This 'molecular biology' approach to protein chemistry, although accurate, is confined to deriving the linear order of common amino acids in polypeptides.

Many peptides and proteins, however, contain amino acids which have been modified in various ways after translation from their messenger RNA, and these modifications cannot be derived directly from the primary structure of the gene. These post-translational modifications are essential for the biological function of the mature polypeptide, and both the types of modifications and the mechanisms by which they occur are apparently highly conserved across species ranging from yeast to man. The conservative nature of these modifications has allowed the determination of precursor protein consensus sequences which become the substrates for the post-translational processing reactions producing the mature molecule. The ability to predict polypeptide sequence from the primary nucleic-acid structure, coupled with knowledge of characteristic processing mechanisms, can in theory lead to the prediction of the mature processed form of polypeptides directly from a nucleic-acid sequence. But Anderegg *et al.*² demonstrate that a consensus sequence alone is not necessarily a reliable predictor of the structure of a mature peptide.

The nucleic-acid sequence encoding the mating hormone called a-factor from the yeast *Saccharomyces cerevisiae* has been known since 1985. The predicted precursor prohormone has at its carboxy terminus the sequence . . . Cys-Val-Ile-Ala, consistent with the consensus sequence for the acylation of the carboxy terminus of *ras* genes (a in the figure; sequences summarized in ref. 3). The identity of these sequences suggests that yeast a-factor should undergo the same post-translational processing as the yeast *ras*-related proteins. Furthermore, to

exert their activities as regulators of cell proliferation, *ras* proteins must be inserted into the cellular membrane, and a similar mechanism has been postulated to be necessary for a-factor activity³. Membrane insertion occurs only for *ras* proteins that have been modified at the cysteine residue near the carboxy terminus by conversion into a thioester of palmitic acid⁴. These considerations all point towards identical post-translational modifications for *ras* proteins and a-factor, that is, conversion to a palmitate cysteine-thioester (b in the figure).

In fact, a yeast gene has been identified on the basis of being involved with the processing of both *ras* and a-factor. *RAM* (for *ras* and a-factor maturation function) was described⁵ in 1986, and is essential for sexual activity of a-type haploid yeast and for the palmitate esterification and membrane localization of *ras* proteins. It was originally proposed that the *RAM* gene encoded an acyltransferase, with the *de facto* corollary that both a-factor and *ras*

have the same cysteine-thioester modifications. A gene called *DRP1* has now been shown⁶ to be responsible for a post-translational processing step occurring before fatty-acid alkylation of *ras*⁵. *DRP1* turns out to be identical to *RAM*.

The data of Anderegg *et al.*² now unequivocally show that the cysteine modifications of *ras* and a-factor are not identical, or even similar. These authors purified a-factor from *S. cerevisiae* and established by proton nuclear magnetic resonance and mass spectrometry that the biological activity of a-factor results from a pair of closely related post-translationally modified peptides with a carboxy-terminal α-carboxymethylcysteine-S-farnesyl ether (c in the figure). Farnesol, a branched-chain, polyunsaturated hydrocarbon alcohol, is best known as an intermediate in the biosynthesis of cholesterol and related molecules. The presence of a farnesyl moiety as a cysteine-S-thioether *per se* is not unusual, as it is known to be a component of peptide pheromones from other fungi. Not only is the alkyl group of a-factor an unsaturated branched-chain alcohol, whereas the *ras* protein has a saturated linear fatty acid moiety, but the chemical linkage is different in each case; a thioether for a-factor and a thioester for *ras*. These findings are totally incompatible with the predicted identical modifications postulated on the basis of identical carboxy-terminal precursor sequences.

In retrospect, the incorrect prediction of the nature of the a-factor modification and the identity of the protein encoded by *RAM* were caused by reliance on the 'molecular biology' approach to predicting peptide structure. The *RAM/DRP1* gene described by Powers *et al.*³ is obviously not an acyltransferase acting on both *ras* and a-factor. It seems more likely that *RAM* encodes a carboxypeptidase or endopeptidase responsible for removal of the tailing A-A-X amino acids of both precursor molecules to provide a free cysteine for modification. Alternatively, *RAM* could encode an enzyme responsible for the methylesterification of the α-carboxy terminus of both peptides. The assignment of a function for *RAM* is an illustration of the need for direct determination of peptide structures to resolve the nature of proposed post-translational modifications. □

1. Malcolm A.D.B. *Nature* **275**, 90-91 (1978).
2. Anderegg R.J., Betz, R., Carr, S.A., Crabb, J.W. & Duntzo, W. *J. Biol. Chem.* **263**, 18236-18240 (1988).
3. Powers S. *et al.* *Cell* **47**, 413-422 (1986).
4. Willumsen B.M., Christensen, A., Hubbert, N.L., Papageorge, A.G. & Lowry, D.R. *Nature* **310**, 583-586 (1984).
5. Tamanoi, F. *et al.* *J. cell. Biochem.* **36**, 261-273 (1988).

Anthony M. Treston and James L. Mulshine are at the Navy Medical Oncology Branch of the National Cancer Institute, Naval Medical Center, Bethesda, Maryland 20814, USA.