Costabilization of Peptide and RNA Structure in an HIV Rev Peptide−RRE Complex†

Ruoying Tan and Alan D. Frankel*

Department of Biochemistry and Biophysics and Gladstone Institute of Virology and Immunology, University of California, San Francisco, P.O. Box 419100, San Francisco, California 94141

Received July 29, 1994; Revised Manuscript Received September 30, 1994®

ABSTRACT: An arginine-rich peptide corresponding to amino acids 34−50 of the human immunodeficiency virus Rev protein has been shown to bind specifically to its RNA-binding site (RRE) when the peptide is in an α-helical conformation. Mutation of any one of six amino acids (Thr34, Arg35, Arg38, Arg39, Asn40, or Arg44) was shown to strongly decrease specific RNA-binding affinity in vitro, suggesting that these residues may contact specific bases or distinct structural features of the RNA. We now show that the four arginine side chains, and not just their charge, are important for specific binding in vivo, and present evidence that three additional arginines (Arg46, Arg48, and Arg50) may make electrostatic contacts to the RRE. RNA-binding specificity of the Rev peptide is temperature-dependent in vitro, correlating with α-helix unfolding. Circular dichroism experiments indicate that the peptide helical structure is stabilized when bound specifically to the RRE and that the RNA undergoes a conformational change upon binding. Because the structures of the peptide and RNA in this model system appear to be mutually stabilized upon binding, it is suggested that the entire complex may be viewed as a single folding unit.

From a biological perspective, the study of RNA−protein recognition is interesting because many cellular functions, including transcription, RNA splicing, and translation, depend on the specific interaction of proteins and RNA. From a macromolecular perspective, the problem is interesting because RNAs can fold into a wide range of tertiary structures, although so far only a few structures have been solved. The three-dimensional structure of tRNA has been known for quite some time, and the cocrystal structures of three tRNA synthetase−tRNA complexes are now known (Rould et al., 1989, 1991; Ruff et al., 1991; Cavarelli et al., 1993; Biou et al., 1994). While the basic shapes of the tRNAs in the three protein−RNA complexes are similar, important recognition features are quite different. For example, protein contacts occur largely in the RNA minor groove in the glutaminyl-tRNA synthetase complex (Rould et al., 1989, 1991), largely in the major groove in the aspartyl complex (Ruff et al., 1991; Cavarelli et al., 1993), and largely to the backbone of a long variable tRNA stem in the seryl complex (Biou et al., 1994). The tertiary folds of the synthetase proteins are very different, further illustrating how structurally diverse RNA−protein recognition is likely to be.

Fortunately, several common amino acid motifs have recently been identified in RNA-binding proteins, allowing them to be grouped into families (Mattaj, 1993; Burd & Dreyfuss, 1994). It was anticipated that related family members would use similar protein frameworks and amino acid contacts to recognize RNA, as observed, for example, with certain members of the helix−turn−helix DNA-binding family (Pabo et al., 1990). However, the wide diversity of RNA structures recognized by closely related RNA-binding proteins suggests that the "rules" for RNA recognition may be less obvious. For example, binding studies with RNP domain-containing proteins show that the same protein architecture, or even the same protein, can be used to recognize rather different RNA structures (Mattaj, 1993). Conversely, the diversity of tRNA synthetase structures indicates that it is possible for very different proteins to recognize the same basic RNA fold, although some small domains may be common to several synthetases (Schimmel et al., 1993). To further complicate the picture, RNA-binding domains from proteins even within the same family may adopt different conformations. For example, an arginine-rich domain from the human immunodeficiency virus (HIV) Rev protein recognizes its RNA site only when in an α-helical conformation (Tan et al., 1993) whereas related domains from the HIV and bovine immunodeficiency virus (BIV) Tat proteins appear to recognize their sites as unstructured or nonhelical peptides (Calnan et al., 1991a,b; Chen & Frankel, 1994; L. Chen and A.D.F., submitted; R.T. and A.D.F., in preparation). In this study, we further examine RNA recognition by the α-helical Rev peptide.

Rev is an essential regulatory protein encoded by HIV. Although its mechanism of action is not yet fully determined, Rev is believed to activate the nucleocytoplasmic transport of unspliced and incompletely spliced HIV mRNAs either directly, by facilitating their access to a nuclear RNA transport pathway (Emerman et al., 1989; Malim et al., 1989), or indirectly, by inhibiting their interaction with cellular splicing factors (Chang & Sharp, 1989; Kjems et al., 1991b). The unspliced and incompletely spliced mRNAs encode the viral structural proteins needed to assemble infectious virus particles.

To function, Rev must interact with a cis-acting RNA element, the Rev response element (RRE), located within the *env* gene. A single hairpin within the relatively large

† Supported by NIH Grants GM39589 and AI29135. R.T. is a Scholar of the American Foundation for AIDS Research (70405-15-RF).

* Address correspondence to this author at Gladstone Institute/ Biochemistry and Biophysics, UCSF, P.O. Box 419100, San Francisco, CA 94141. Phone: 415-695-3816. Fax: 415-826-8449.

® Abstract published in *Advance ACS Abstracts,* November 1, 1994.

0006-2960/94/0433-14579$04.50/0