Probing the structure of the HIV-1 Rev trans-activator protein by functional analysis

Sarah L.Thomas, Joachim Hauber and Georg Casari

Introduction

The biology of Rev

The human retrovirus human immunodeficiency virus type 1 (HIV-1), the causative agent of acquired immunodeficiency syndrome (AIDS), encodes several regulatory proteins in addition to the viral structural proteins Gag,Pol and Env (for reviews, see Pavlakis and Felber, 1990; Cullen, 1991; Greene, 1991). Two such viral trans-acting regulators, Tat and Rev, are essential for virus replication. The Rev protein acts post-transcriptionally and primarily at the level of nucleo-cytoplasmic transport of incompletely spliced and unspliced viral mRNAs (reviewed in Cullen and Malim, 1991; Gerace, 1995). Rev recognizes and directly binds these viral mRNAs via its RNA target sequence, the highly structured cis-acting Rev response element (RRE), which is encoded by sequences residing in the env gene. The subsequent interaction of Rev with one or multiple cellular cofactor(s) finally results in the nucleo-cytoplasmic transport of the RRE-containing viral mRNAs. Indeed, various cellular proteins have recently been described which appear to be specific interaction partners of Rev, including the nucleoporin-like protein hRIP/Rab (Bogerd et al., 1995; Fritz et al., 1995) and nuclear eukaryotic initiation factor 5A (eIF-5A) (Ruhl et al., 1993; Bevec et al., 1996).

Organization of Rev and predicted structural models

CD measurements (Auer et al., 1994) combined with functional studies of mutated variants of the HIV-1 rev gene suggest that Rev essentially consists of two structural domains with distinct regions of defined activities. The 50–55 residues of the C-terminal region which contain the protein activation domain (Hope et al., 1991; Malim et al., 1991; Weichselbraun et al., 1992b), appear to be more variable and are structurally less well defined. This part contains a leucine-rich peptide core motif which binds to cellular cofactors, thereby mediating the transport of viral mRNA from the nucleus to the cytoplasm. Mutations within this region result in Rev mutant proteins with trans-dominant in vivo phenotypes (Malim et al., 1989, 1991; Mermer et al., 1990; Venkatesh et al., 1990).

The N-terminal half of the protein is formed by the first 61–66 aa residues. A stretch of eight arginine residues has been shown to be responsible for the nuclear/nucleolar localization of Rev and also mediates the direct and specific binding of Rev to the viral RRE RNA (Hope et al., 1990a; Böhnlein et al., 1991; Zapp et al., 1991; Kjems et al., 1992; Hammerschmid et al., 1994). Residues flanking the arginine-rich region on both sides are thought to be involved in the oligomerization of Rev molecules (Olsten et al., 1990; Malim and Cullen, 1991; Zapp et al., 1991) which is necessary for biological activity. The precise amino acid sequences involved have yet to be mapped and characterized.

CD-measurements have provided strong evidence that most of the N-terminal domain is in an α-helical conformation. A proline-rich stretch is incompatible with this conformation and is supposed to form a loop. A structural model of the N-terminal domain has been proposed (Auer et al., 1994) and is depicted schematically in Figure 1. It suggests two α-helices, one N-terminal and one C-terminal of the proline rich stretch, that contact via a series of hydrophobic residues. Helix 1 includes a number of conserved and important hydrophobic residues that can combine to provide a number of alternative helix contact points. As neither analysis of CD spectra nor other spectroscopic methods could discriminate where these contacts take place, the structural model is still poorly defined regarding the relative orientation of the helices and the precise points of contact. A set of alternative models has thus been proposed (Auer et al., 1994).

Despite intensive attempts to solve the three-dimensional structure of Rev, its physical and biological properties mean that X-ray crystallography and NMR studies have so far been unsuccessful. In this paper we present a new approach that uses recombinant techniques and selectively designed mutants to determine specifically points of contact within a protein. In the case of the N-terminal domain of Rev, we demonstrate that this strategy can provide the information needed to decide on a particular structural arrangement in a relatively straightforward manner.
Designing mutations to probe the structure of HIV-1 Rev

Materials and methods

Plasmid constructions

The pcREV, pBC12/CMV/IL-2, pBC12/CMV/βGal and pDM128/CMV expression plasmids have been described in detail previously (Cullen, 1986a; Malim et al., 1988, 1991; Ruhl et al., 1993). Oligonucleotide-directed mutagenesis with a bacteriophage M13 mutagenesis system (United States Biochemical, Cleveland, OH) was used to introduce targeted nucleotide substitutions encoding single and double amino acid alterations into the rev gene of pcREV. All mutations introduced were confirmed by DNA sequencing using Sequenase 2.0 (United States Biochemical).

Cell culture, transfections and assays

COS cells were maintained as described previously and were transiently transfected using DEAE-dextran and chloroquine (Cullen, 1986b). Rev trans-activation was investigated by cotransfection of COS cell cultures of 3.5×10^5 cells with 375 ng of pDM128/CMV (Rev-reporter) (Malim et al., 1991) together with 250 ng of either pcREV (positive control), mutant Rev expression plasmid (TC series) or pBC12/CMV/IL-2 (negative control) and 75 ng of pBC12/CMV/βGal (transfection efficiency control). pDM128/CMV contains the gene encoding chloramphenicol acetyltransferase (CAT) and the RRE target sequence of Rev (Hope et al., 1990a), both of which are positioned between HIV-1 splice sites and expressed transiently using the cytomegalovirus immediate early (CMV-IE) promoter. At ~60 h post-transfection, cell lysates were prepared and the levels of β-galactosidase activity (Fiering et al., 1991) were measured. In contrast, co-expression of functional Rev results in nucleo-cytoplasmic transport of unspliced (cat-containing) mRNAs and induces their cytoplasmic expression (Hope et al., 1990a; Ruhl et al., 1993). Protein expression of the rev mutant genes was confirmed as described in detail recently (Hammerschmid et al., 1994) by Rev-specific radioimmunoprecipitation analysis of COS cell cultures (3.5×10^5 cells) transfected transiently with 500 ng of expression plasmid DNA.

3D modelling

The new structural model of Rev was interactively constructed on a Silicon Graphics Indigo R4000 using the modelling software SYBYL (TRIPOS). The two helical regions were set in default α-helical conformation as provided by the modelling package with the two helices arranged in typical geometry of helix contacts forming the measured contact Ile55 with Ile19. As there were no other data available, the connecting proline-rich loop was set in a typical proline conformation. Conformations of side chains were only altered from standard geometry in cases where steric clashes had to be avoided. In this way, a level of detail compatible with the level of detail provided by the experimental data used was maintained.

Results and discussion

Designing mutations to probe the structure of HIV-1 Rev

The strategy we used to test structural hypotheses was designed to detect contacts between residues in the three-dimensional structure of the functional Rev protein by testing carefully chosen single and double mutations. Using the structural models of HIV-1 Rev that have already been proposed (Auer et al., 1994) as a starting point, we pinpointed residue pairs that are predicted to be in contact in the respective alternative structures. In order to refine the choice of mutations to be introduced, we used 3D structural models which take into account the chirality of the side chains and 3D geometry of α-helix packing.

To test for structural contacts, it was necessary to verify the structural importance of the hydrophobic residues concerned. We first confirmed that these residues were consistently conserved in all sequences available. The important role of these residues was then confirmed experimentally by functionally testing mutants that replaced hydrophobic amino acid residues by hydrophilic charged side chains (Asp) which are incompatible with forming buried hydrophobic contacts. Non-functional mutants therefore provide evidence for important hydrophobic contacts. Rev mutant proteins that affect the amino acid residues Val16, Leu18, Ile19 (Hope et al., 1990b) and Leu22 (Weichselbraun et al., 1992a) have previously been described as non-functional. In this study we replaced Ile52 and Ile55 by aspartate to test the importance of these residues with respect to trans-activation.

Second, it was necessary to discriminate residues that fulfill a direct functional role (like binding molecular partners) from those of structural importance. As functional requirements are usually much more stringent, a subtle mutation (here to Ala) will generally already abolish function, as has been described previously for the leucine motif of the activation domain (Malim et al., 1991). Positions making stabilizing hydrophobic contacts are tolerant to such exchanges and have been well studied in other proteins. A moderating effect comes from the inherent general flexibility of proteins which allows them partly in the cytoplasm. Thus, only background levels of CAT activity are measurable. In contrast, co-expression of functional Rev results in nucleo-cytoplasmic transport of unspliced (cat-containing) mRNAs and induces their cytoplasmic expression (Hope et al., 1990a; Ruhl et al., 1993).
Table 1. Description of the HIV-1 Rev mutants

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Position (aa)</th>
<th>Mutation (aa)</th>
<th>Contact in</th>
</tr>
</thead>
<tbody>
<tr>
<td>RevTC1</td>
<td>16</td>
<td>V → A</td>
<td>–</td>
</tr>
<tr>
<td>RevTC2</td>
<td>18</td>
<td>L → A</td>
<td>–</td>
</tr>
<tr>
<td>RevTC3</td>
<td>19</td>
<td>I → A</td>
<td>–</td>
</tr>
<tr>
<td>RevTC4</td>
<td>23</td>
<td>Y → P</td>
<td>–</td>
</tr>
<tr>
<td>RevTC5</td>
<td>52</td>
<td>I → D</td>
<td>–</td>
</tr>
<tr>
<td>RevTC6</td>
<td>52</td>
<td>I → A</td>
<td>–</td>
</tr>
<tr>
<td>RevTC7</td>
<td>55</td>
<td>I → D</td>
<td>–</td>
</tr>
<tr>
<td>RevTC8</td>
<td>55</td>
<td>I → A</td>
<td>–</td>
</tr>
<tr>
<td>RevTC9</td>
<td>16/52</td>
<td>V → A/I → A</td>
<td>Model 1</td>
</tr>
<tr>
<td>RevTC10</td>
<td>16/55</td>
<td>V → A/I → A</td>
<td>Model 1</td>
</tr>
<tr>
<td>RevTC11</td>
<td>18/55</td>
<td>L → A/I → A</td>
<td>Model 2</td>
</tr>
<tr>
<td>RevTC12</td>
<td>19/52</td>
<td>I → A/I → A</td>
<td>New model</td>
</tr>
<tr>
<td>RevTC13</td>
<td>19/55</td>
<td>I → A/I → A</td>
<td>New model</td>
</tr>
</tbody>
</table>

The protein name, position(s) mutated within the 116 amino acid HIV-1 Rev trans-activator protein and the amino acid changes introduced are listed in columns 1, 2, and 3, respectively. The structural models of Rev N-terminus (compare Figure 1), which are tested or suggested by the double mutants (TC9 to TC13), are listed in column 4.

to close the cavity created by making minor compensatory rearrangements of the surrounding amino acids (Eriksson et al., 1992; Buckle et al., 1993; Jackson et al., 1994). Such a mutation was introduced and tested for function for all hydrophobic positions of interest.

In order to test for a particular contact between two residues within the structure, it is necessary to compare the phenotypes of the individual single mutation Rev proteins with that of a double mutant. In general, the stabilizing and destabilizing effects of different mutations are independent and their combined impact on stability is simply the sum of the individual effects (Wells, 1990; Gregoret and Sauer, 1993). Therefore, a combination of two mild mutations will usually show little effect on stability and the double mutant would be expected to have virtually wild-type activity. If, however, the two mutated residues are in close contact in the functional structure, the double mutant would be expected to be severely destabilized and consequently have reduced functional activity. This can be explained by the fact that compensatory rearrangements for a single mutation that involve the other contacting residue can no longer take place in the double mutant. Furthermore, if the compensatory rearrangements for each single mutation involve partly moving the same individual residues in different ways it would be impossible for these movements to occur simultaneously in the double mutant. Thus, contacting residues restrict each other's range of possible functional mutations, a fact that has already been observed by comparing natural or wild-type Rev in transfected COS cells, as measured by Rev-specific radioimmunoprecipitation analysis (Figure 2).

The whole series of rev mutant vectors were tested by cotransfecting COS cell cultures with both the Rev mutant plasmid and the pDM128/CMV reporter and controlled to take into account varying transfection efficiencies by addition of the pBLC12/CMV/βGal (Ruhl et al., 1993) construct to the transfections as described in the Materials and methods section. Clearly, expression of Rev wt resulted in significant transactivation rates. The relative CAT levels obtained as a percentage of Rev wild-type (wt) activity (set to 100%) are shown in Figure 3.

Rev mutants TC5 and TC7 represent Ile to Asp mutations and were designed to prove the importance of residues Ile52 and Ile55. Clearly, TC5 has completely lost function showing crucial
importance for Ile52. The case is not as clear for TC7 (Ile55) but the activity of this mutant is reduced.

Mutants TC1, TC2, TC3, TC6 and TC8 represent replacements of the chosen hydrophobic residues by alanine, a mutation expected to do little harm at structurally important positions. TC1, TC3 and TC6 (replacing Val16, Ile19 and Ile52) undoubtedly show wild-type activities, in spite of being sensitive to replacement by aspartate, clearly indicating a structural and not a functional role for those residues. TC2 and TC8 (replacing Leu18 and Ile55) show slightly reduced activity. Although the results are not as clear as for TC1, TC3 or TC6, these mutants are functional and are at a higher level than the Asp-mutants which is consistent with them playing a structural role.

TC9, TC10, TC11, TC12 and TC13 represent double mutations involving two residues that make direct contact in either of the alternative structural models. Each of these residues’ single mutation to alanine had no or only a minor effect on Rev trans-activation activity. As generally expected for arbitrary combinations of harmless mutations, TC9, TC10 and TC11 have wild-type level activity, indicating that these residue pairs (Val16–Ile52, Val16–Ile55 and Leu18–Ile55) are not touching and making crucial stabilizing hydrophobic contacts. However, in spite of the single mutant phenotypes, TC12 and TC13 show drastically reduced trans-activation levels. This ‘unpredictable’ drastic reduction can be interpreted as a direct consequence of disruption of important structural contacts made by the mutated pairs (Ile19–Ile52 and Ile19–Ile55).

Besides those mutants which were specifically designed to indicate structural contacts, we also introduced a tyrosine to proline mutation at position 23 (TC4). This position is generally very tolerant to mutations [e.g. to Asp (Berger et al., 1991) or Ser (Hope et al., 1990b)]. Mutant TC4 shows strongly reduced functionality, providing strong evidence that Tyr23 is still part of α-helix 1 as a proline at this position is deleterious.

A new N-terminal model of HIV-1 Rev

The data presented above can be used to infer the structural organization of the Rev N-terminus when the protein is in its functional conformation. We can now test whether or not the central contacts of alternative structural Rev models are in agreement with the data obtained experimentally.

Model 1 (see Figure 1) proposed that Val16 makes important contacts to Ile52 and Ile55, stabilizing the helix 1–helix 2 contact by their hydrophobic interaction. Our findings contradict this model, as double mutants involving these residues (TC9 and TC10; Figure 3) show wild-type levels of activity.

Model 2 puts the very conserved Leu18 as a central residue in the helix 1–helix 2 interaction surface (see Figure 1), thereby proposing an intensive contact between Leu18 and Ile55. No reduction in the level of trans-activation was measured for the double mutant. (TC11; Figure 3), which contradicts this structural arrangement.

Our results reject both current alternative structural models. However, we have identified isoleucine at aa position 19 as a clear candidate for being directly involved in helix–helix contact, as both TC12 (Ile19–Ile52) and TC13 (Ile19–Ile55) show drastically decreased levels of activity. We used this new finding as an additional constraint for the construction of a structural model for the Rev N-terminal domain. It is possible to arrange helix 1 and helix 2 in such a way that Ile19 takes the central position in the helix contact (depicted schematically in Figure 1, ‘new model’). The new constraint fixes the orientation of helix 1 with respect to helix 2. This new model is distinct from the structural arrangements that have been proposed to date. In the process of searching for a consistent geometry, a full atom model has been constructed by interactive computer modelling and is shown in Figure 4. A full atom representation was chosen to keep the geometry and chirality of helices and amino acid residues realistic; it does not express the resolution of the model.

With the existence of this more refined structural model, it is now possible to identify those amino acid residues in the N-terminal sequence of HIV-Rev that co-operate to form potential interaction surfaces. Rev protein–RRE RNA complexes need to bind specifically to at least a second Rev molecule to mediate RNA export (Malim and Cullen, 1991). Residues Leu18, Phe21, Leu22, Ile55 and Ile59 create a hydrophobic area that is not buried through the helix–helix contact. An alternative smaller hydrophobic patch at the opposite side of the domain involves Leu12, Leu13, Val16 and Leu60 and may play a role in the interaction of the N-terminal domain with the C-terminal activation domain.
With this new structural knowledge, it becomes possible to design rationally mutants with desired properties. In particular, it would be desirable to create mutant Rev variants that do not form a functional complex.

A new method to probe structural contacts

Using a set of established observations on the stability of mutant proteins, we have designed a new strategy to obtain structural information about proteins using molecular recombination techniques. We were able to confirm a number of general predictions by carrying out a series of experiments involving mutants containing both single and double amino acid changes. Upon introduction of subtle changes, the protein stays active but becomes sensitive to further mutations of some selected residues, presumably those residues which are structurally close together in the functional conformation of the protein.

This new method offers an alternative way to probe protein structures in cases where standard techniques fail. This is possible because our technique does not require homogeneous solutions of high concentration and purity (e.g. like NMR) or even crystals of high quality that can diffract well enough to resolve the 3D structure by X-ray crystallography. Our method relies on a functional assay. The quantities of protein required can be extremely small. Therefore, we do not suffer from problems met when expressing unphysiologically large amounts of protein. We also probe the properties of the protein under physiological conditions. Such in vivo structure determination, as tested here by Rev's trans-activation capacity, is not yet possible using other methods.

Because of the experimental complexity that limits the number of mutants that can realistically be tested at present, this approach relies on structural models that provide a limited set of structural alternatives to be tested. With the present progress in technology and robotics, however, it is conceivable that the number of different experiments will in the future no longer be a limitation. Instead, complete mutational scans through a protein will be available. With such data, in principle, one can imagine a complete structure determination where structurally close pairs of residues are identified by functional data obtained for carefully designed series of mutants.

References