Journal of Structural Biology xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

# A new HIV-1 Rev structure optimizes interaction with target RNA (RRE) for nuclear export

Norman R. Watts[a,1], Elif Eren[b,1], Xiaolei Zhuang[a,1], Yun-Xing Wang[c], Alasdair C. Steven[b], Paul T. Wingfield[a,*]

[a] Protein Expression Laboratory, NIAMS, NIH, Bethesda, MD 20892, USA
[b] Laboratory of Structural Biology Research, NIAMS, NIH, Bethesda, MD 20892, USA
[c] Structural Biophysics Laboratory, NCI, NIH, Frederick, MD 21702, USA

ARTICLE INFO

*Keywords:*
HIV-1
Nuclear export complex
Rev
Rev response element
RRE
Specificity checkpoint

ABSTRACT

HIV-1 Rev mediates the nuclear export of unspliced and partially-spliced viral transcripts for the production of progeny genomes and structural proteins. In this process, four (or more) copies of Rev assemble onto a highly-structured 351-nt region in such viral transcripts, the Rev response element (RRE). How this occurs is not known. The Rev assembly domain has a helical-hairpin structure which associates through three (A-A, B-B and C-C) interfaces. The RRE has the topology of an upper-case letter A, with the two known Rev binding sites mapping onto the legs of the A. We have determined a crystal structure for the Rev assembly domain at 2.25 Å resolution, without resort to either mutations or chaperones. It shows that B-B dimers adopt an arrangement reversed relative to that previously reported, and join through a C-C interface to form tetramers. The new subunit arrangement shows how four Rev molecules can assemble on the two sites on the RRE to form the specificity checkpoint, and how further copies add through A-A interactions. Residues at the C-C interface, specifically the Pro31-Trp45 axis, are a potential target for intervention.

## 1. Introduction

HIV/AIDS is projected by the World Health Organization to be among the leading contributors to global burden of disease by 2030 (Mathers and Loncar, 2006) (for 2015 updates see http://www.who.int/healthinfo/global_burden_disease/projections/en/). Despite the unquestionable success of combination antiretroviral therapies (cART) there remains a need for alternative means of intervention. In HIV-1, viral replication is regulated by two proteins; Tat (trans-activator protein) which stimulates transcription from the long terminal repeat, and Rev (regulator of expression of virion proteins) which controls the nuclear export of unspliced and partially-spliced transcripts required for the production of viral genomes and structural proteins. Rev binds to a 351-nt region within viral transcripts, known as the Rev response element (RRE), coupling them to Crm1, a nuclear export factor for certain host proteins and for RNAs (Karn and Stoltzfus, 2012). The importance of this interaction as a potential point of intervention has long been recognized (Baba, 2004; Giver et al., 1993; Xiao et al., 2001), but it has remained largely untargeted due in part to a lack of structural information about both Rev and the RRE.

Rev is a small (13 kDa) protein with two domains; an N-terminal domain (residues 1–65) with a simple anti-parallel helix-loop-helix hairpin-like structure, and a C-terminal domain (residues 66–116) that is generally thought to be unstructured (Daugherty et al., 2010a; DiMattia et al., 2010; Watts et al., 1998). The N-terminal domain engages in three types of homotypic interactions termed A-A, B-B and C-C. The first two involve the pairing of monomers through the A and B hydrophobic faces located on either side of the open ends of the hairpins (heterotypic A-B interactions have never been observed), whereas C-C interactions involve association between the closed (loop) ends of the hairpins. Observation of these three interfaces in numerous crystal structures (Daugherty et al., 2010a; DiMattia et al., 2010; Jayaraman et al., 2014), as well as in several cryo-EM helical reconstructions of tubular Rev polymers with different helical lattices (DiMattia et al., 2016), has revealed wide variation in the crossing-angles of the dimers. In the case of A-A dimers, the angle can vary, seemingly almost continuously, between 96° and 140° (DiMattia et al., 2016) whereas B-B dimers have been observed at angular settings of 50° and 120° (Daugherty et al., 2010a; Jayaraman et al., 2014). This variability is facilitated by the highly hydrophobic and rather featureless nature of

* Corresponding author.
*E-mail address:* pelpw@helix.nih.gov (P.T. Wingfield).
[1] These authors contributed equally.

https://doi.org/10.1016/j.jsb.2018.03.011
Received 12 February 2018; Received in revised form 27 March 2018; Accepted 28 March 2018

these interfaces. Unlike the A-A and B-B interfaces, the C-C interface involves the proline-rich loop of one monomer hooking into a groove formed by the corresponding region of the opposing one (DiMattia et al., 2016). As pairing involves Pro-Trp stacking interactions and the formation of specific hydrogen bonds, the range of crossing angles at C-C interfaces is more restricted. A-A and B-B interactions are the dominant ways in which Rev monomers associate. By comparison, the C-C interface involves a smaller area (and a smaller calculated free energy of formation) and was initially considered a mere crystal contact. However, the repeated observation in several different crystal settings and helical lattices implied a functional role (DiMattia et al., 2016). It is these variable interactions of the N-terminal domain, and the apparently disordered nature of the C-terminal domain, combined with a strong tendency towards aggregation, that long hindered structure determination. Crystallization of full-length Rev was finally achieved by the use of Fab and scFv chaperones to block the B-faces, but the C-terminal domains remained disordered in the crystal lattices (DiMattia et al., 2010, 2016). In helical reconstructions of full-length Rev, the C-terminal domains appeared only as poorly ordered bi-lobed densities (DiMattia et al., 2016). The N-terminal domain has also been crystallized, either alone or bound to a small RNA, but only when mutated at the A-face (Daugherty et al., 2010a; Jayaraman et al., 2014). In addition to the N- and C-terminal domains, Rev has several functional regions. Pertinent to this discussion is the Arg-rich motif (ARM) in helix 2 of the N-terminal domain and located near the closed end of the hairpin in the folded protein. It is the ARM by which Rev binds to the RRE (Battiste et al., 1996; Karn and Stoltzfus, 2012).

The RRE has been studied intently for 30 years, primarily by biochemical means, but it is still usually depicted either as a two-dimensionally folded RNA or simply as a cartoon. Nothing was known about the three-dimensional structure until recently when a SAXS-based model was advanced showing it as having a topology resembling an upper-case letter A (Fang et al., 2013). A single high-affinity binding site for Rev, located in Stem IIB of the RRE, was identified early on (Malim et al., 1990; Tiley et al., 1992), and this maps to one leg of the A-shaped RNA. The details of how Rev binds to the IIB site have previously been elucidated in a solution NMR study of an ARM peptide bound to a minimal IIB RNA (Battiste et al., 1996). More recently, a second site (IA) was identified in Stem I of the RRE, but this has been less well characterized and no structure has been reported (Daugherty et al., 2008; Jayaraman et al., 2014). This site maps onto the other leg of the A-shaped RNA. A third site has also been proposed at a distal location on Stem I on the basis of time-resolved SHAPE analysis (Bai et al., 2014).

Rev has been reported to bind to the RRE in *ca.* 4–13 copies (Rausch and Le Grice, 2015). However, how Rev assembles on the A-shaped RRE is not known. On the basis of the similar spacing between the two ARMs in A-A dimers (*ca.* 55 Å) and the distance between the IIB and IA sites, it was proposed that a single Rev dimer bridges the two sites and that subsequent dimers add in an analogous manner along the two legs of the A-shaped RRE (Fang et al., 2013). Following the discovery of the C-C interface, an updated model was proposed wherein one Rev dimer binds at each of the IIB and IA sites and these dimers are joined through a C-C interaction (DiMattia et al., 2016). Neither model showed how Rev engages the RRE, or how multiple copies of Rev might assemble given only two (or perhaps three) binding sites on the RRE. Attempts at three-dimensional reconstructions of the Rev-RRE-Crm1 export complex (Booth et al., 2014) have been hampered by the highly flexible nature of the assembly. In cases where traditional methods of structure determination experience difficulty, SAXS, though lower in spatial resolution, can still provide structurally useful information (Hura et al., 2009), with the HIV-1 RRE being a notable example.

Here we describe an X-ray crystal structure of the Rev assembly domain in an entirely new arrangement, obtained without either assembly-impeding mutations or chaperones. Docked with the SAXS-based model of the RRE, it shows how Rev can bind to and assemble on the RRE to form a stable four-Rev specificity checkpoint and also allow for the addition of further subunits through protein-protein interactions.

## 2. Materials and methods

### 2.1. Preparation of $Rev^{1-69}$

Rev has two domains; an ordered N-terminal assembly domain (usually considered to be residues 1–65) and a disordered C-terminal effector domain (residues 66–116), however, in one instance (PDB: 3LPH) the N-terminal domain has been observed to be helically ordered up to Glu69 or possibly Pro70 (Daugherty et al., 2010a). To obtain crystals of the assembly domain we deleted the C-terminal domain. The resulting protein, $Rev^{1-69}$, otherwise had the wild-type sequence and a molecular mass of 8.1 kD. We did not employ any mutations to impede oligomerization, nor did we add a surface entropy reducing mutation, as has sometimes been done (Jayaraman et al., 2014). $Rev^{1-69}$ was expressed in *E. coli*. The cells were suspended in 50 mM Tris, pH 8.0, 0.1 M NaCl, 1 mM EDTA and 2 M urea supplemented with a protease inhibitor cocktail (Roche) then lysed by two passes through a French Press at 12,000 psi with one minute of sonication after each pass to reduce viscosity. The material was clarified by centrifugation in a JA-14 rotor at 13,000 rpm for two hours at 4 °C and then fractionated by DEAE anion exchange chromatography. Flow-through fractions were pooled, subjected to a 40% ammonium sulfate cut, and gel filtered on Superdex 75 (GE Healthcare) equilibrated with 4 M Guanidine HCl. The protein was refolded by sequential dialysis, first against 25 mM HEPES, pH 7.4, 0.15 M NaCl, 2 mM DTT and then 25 mM HEPES, pH 7.4, 0.15 M NaCl, 0.1 M Arginine, 10% glycerol and 2 mM DTT. The protein was concentrated using Millipore Ultra-15 centrifugal filters and any precipitate was removed by bench top centrifugation. The protein concentration was determined using UV spectroscopy: 1 mg/ml has an absorbance of 1.0.

### 2.2. Crystallization, diffraction data collection and structure determination

Crystallization was performed by the hanging drop method. Protein, 1 µl (4–5 mg/ml), was mixed with 1 µl of 0.1 M sodium citrate (pH 5.6) containing 25% 2-methyl-2-propanol; 100 µl buffer was added to plate wells. Crystals were grown at 10 °C over ~2 weeks. Crystals were cryoprotected with 30% 2-methyl-2,4-pentanediol (MPD) and flash-frozen. Data were collected at the Advanced Photon Source (APS) beamline 22ID at 100 °K. Diffraction data were processed using XDS and scaled and merged using XSCALE (Kabsch, 2010). The HIV-1 Rev structure (PDB: 3LPH) (Daugherty et al., 2010a) was used for molecular replacement in Phenix (Adams et al., 2010). The crystal belonged to the $P2_12_12_1$ space group with a single copy of the Rev dimer per asymmetric unit. Iterative rounds of model building were done in Coot (Emsley and Cowtan, 2004) and refinement in Phenix. Interfaces were analyzed with PDBePISA (Krissinel and Henrick, 2007).

### 2.3. Model building

The structure of the RRE, as determined by SAXS, has been described previously (Fang et al., 2013). All modeling of Rev on the RRE was performed with UCSF Chimera (Pettersen et al., 2004).

## 3. Results and discussion

### 3.1. Rev dimerizes in an arrangement opposite to that previously reported

We have determined a crystal structure for the Rev assembly domain ($Rev^{1-69}$) at 2.25 Å resolution (Table 1), revealing a dimer (Fig. 1). Both monomers have a well folded core (residues 8–65) and are essentially superimposable. The two monomers associate *via* their B-faces