pathogens that are traditionally regarded as being restricted to the aerial plant tissues have recently been reported to infect the roots of their hosts[23,24] (*Leptosphaeria maculans*, which causes stem canker of brassicas and *Cercospora beticola*, which causes leaf spot disease of sugar beet). Taken together these observations suggest that soil-borne inoculum and root infection may be significant components of the life cycles of diseases that are traditionally regarded as only affecting above-ground parts of plants. This has important implications for the development of new strategies for plant breeding and disease control. □

Methods

DNA manipulations, fungal growth conditions and plant infection experiments

Standard molecular biology procedures were followed for cloning and enzymatic manipulations with DNA[25]. Amino acid sequence comparisons and alignments were carried out using vector NTI Suite8 (InforMax, Invitrogen). Fungal strains were cultured and maintained as described[14]. The plant material used was rice (*Oryza sativa* L.) cultivars Nipponbare (ssp. *japonica*) and CO39 (ssp. *indica*), barley (*Hordeum vulgare*) cultivar Golden Promise and wheat (*Triticum aestivum*) cultivar Riband. Leaf and root infection assays were carried out as described[10,14].

Construction of SGFP binary vectors and *Agrobacterium*-mediated transformation

A modified pCAMBIA1300 (CAMBIA, Canberra, Australia) binary vector containing a hygromycin resistance gene and the *SGFP* gene (encoding a GFP variant that contains a serine-to-threonine substitution at amino acid 65) from pCT74 (ref. 26) was constructed. The resulting vector (pCAMBgfp) was introduced into *Agrobacterium tumefaciens* strain AGL-1 and transformed into *M. grisea* as described[27]. Transformants expressing SGFP were selected under ultraviolet light. A second binary vector expressing SGFP and containing the *ILV1* gene (which confers resistance to sulfonylurea) (pSULF) was also constructed by ligating *Xho*I/*Eco*RI-digested pCB1532 (ref. 28) into pCAMBIA1300. The *SGFP* gene from pCT74 was then ligated into *Sal*I/*Bam*HI-digested pSULF to give pSULFgfp.

Microscopy

Infected root samples were stained with chlorazole black E as described[29]. Confocal optical section stacks of infected plant material were collected using a Leica TCS-NT confocal microscope. SGFP fluorescence was detected with a 515 nm bandpass emission filter and autofluorescence of the plant cell walls with a 595 nm bandpass emission filter.

Construction of *MgFOW1* gene replacement vector

The candidate *FOW1* homologue *MgFOW1* (annotated at locus n. MG07201.4, chromosome I, contig 2.1337 in the *M. grisea* genome database, http://www.broad.mit.edu/annotation/fungi/magnaporthe/) was amplified with primers 5Mg-FOW (5′-CGGTGTTCTCCTGCAGTACTACG-3′) and 3Mg-FOW (5′-CTATGCC TCTGGTTCTATACCG-3′) from *M. grisea* Guy11 strain. The sequence of the 3.7 kilobase (kb) amplification product was identical to that in the genome database (from *M. grisea* 70-15 strain). The polymerase chain reaction (PCR) product was digested with *Pst*I and *Kpn*I and the resulting 3.4 kb restriction fragment cloned into pBluescriptSK$^-$ (Stratagene). A 170 bp *Xho*I–*Hpa*I fragment within the predicted ORF was replaced by the ~2.6 kb *Xho*I–*Sma*I fragment from pCAMBgfp containing the hygromycin resistance cassette and the *SGFP* gene. An ~5.2 kb *Pst*I–*Kpn*I fragment encompassing the 2.6 kb *Xho*I–*Sma*I fragment and flanking DNA sequences was digested with *Age*I and subcloned into pCAMBIA1300. This gene replacement construct was then introduced into the *M. grisea* strain Guy11 by *Agrobacterium*-mediated transformation. The transferred DNA insertion sites were confirmed by sequencing the PCR fragments amplified using the primer pairs: 5Mg-FOW and MgFOW2 (5′- CTTCTTGCCGGTGACATCGCG-3′), 5Mg-FOW and 5-HYGR (5′-GCCGATAGTGGAAACCGACGC-3′), 5-NOS (5′- CTAGA TCCGATGATAAGCTGTC-3′) and 3Mg-FOW (see above).

For complementation experiments the 3.4 kb *Pst*I–*Kpn*I fragment containing the intact *MgFOW1* gene was blunt-ended, cloned into *Sma*I-digested pSULF and introduced into the *ΔMgFOW1* mutants by *Agrobacterium*-mediated transformation. Transformants were selected with sulfonylurea (Greyhound, UK) and analysed by PCR with the primers MgFOW9 5′-GCAGTACTGTGATGACTAGG-3′ and MgFOW6 (5′-GGTTTCGAGTG CCGCTGTCG-3′), which yielded a ~2.8 kb amplification product fragment derived from the *MgFOW1* locus.

Received 28 April; accepted 20 July 2004; doi:10.1038/nature02880.

1. Mendgen, K., Hahn, M. & Deising, H. Morphogenesis and mechanisms of penetration by plant pathogenic fungi. *Annu. Rev. Phytopathol.* **34,** 367–386 (1996).
2. Tucker, S. L. & Talbot, N. J. Surface attachment and pre-penetration stage development by plant pathogenic fungi. *Annu. Rev. Phytopathol.* **39,** 385–418 (2001).
3. Talbot, N. J. On the trail of a cereal killer: exploring the biology of *Magnaporthe grisea*. *Annu. Rev. Microbiol.* **57,** 177–202 (2003).
4. International Rice Blast Genome Consortium. RICEBLAST.ORG, http://www.riceblast.org/ (2003).
5. Cannon, P. F. The newly recognized family Magnaporthaceae and its interrelationships. *Systema Ascomycetum* **13,** 25–42 (1994).
6. Hornby, D. (ed.) *Biology and Control of Take-all* (Academic Press Inc., London, 1981).
7. Hornby, D. (ed.) *Take-all Disease of Cereals: a Regional Perspective* (CAB International, Wallingford, 1998).
8. Landschoot, P. J. & Jackson, N. *Magnaporthe poae* sp. nov., a hyphopodiate fungus with a *Phialophora* anamorph from grass roots in the United States. *Mycol. Res.* **93,** 59–62 (1989).
9. Scott, D. B. & Deacon, J. W. *Magnaporthe rhizophila* sp. nov, a dark mycelial fungus with a Phialophora conidial state, from cereal roots in South-Africa. *Trans. British Mycol. Soc.* **81,** 77–81 (1983).
10. Dufresne, M. & Osbourn, A. E. Definition of tissue-specific and general requirements for plant infection in a phytopathogenic fungus. *Mol. Plant Microbe Interact.* **14,** 300–307 (2001).
11. De Jong, J. C., McCormack, B. J., Smirnoff, N. & Talbot, N. J. Glycerol generates turgor in rice blast. *Nature* **389,** 244–245 (1997).
12. Howard, R. J. & Valent, B. Breaking and entering: host penetration by the fungal rice blast pathogen *Magnaporthe grisea*. *Annu. Rev. Microbiol.* **50,** 491–512 (1996).
13. Rodrigues, F. A., Benhamou, N., Datnoff, L. E., Jones, J. B. & Belanger, R. R. Ultrastructural and cytochemical aspects of silicon-mediated rice blast resistance. *Phytopathology* **93,** 535–546 (2003).
14. Talbot, N. J., Ebbole, D. J. & Hamer, J. E. Identification and characterization of *MPG1*, a gene involved in pathogenicity from the rice blast fungus *Magnaporthe grisea*. *Plant Cell* **5,** 1575–1590 (1993).
15. Chumley, F. G. & Valent, B. Genetic analysis of melanin deficient, nonpathogenic mutants of *Magnaporthe grisea*. *Mol. Plant Microbe Interact.* **3,** 135–143 (1990).
16. Valent, B., Farrall, L. & Chumley, F. *Magnaporthe grisea* genes for pathogenicity and virulence identified through a series of backcrosses. *Genetics* **127,** 87–101 (1991).
17. Mitchell, T. K. & Dean, R. A. The cAMP-dependent protein kinase catalytic subunit is required for appressorium formation and pathogenesis by the rice blast pathogen *Magnaporthe grisea*. *Plant Cell* **7,** 1869–1878 (1995).
18. Xu, J. R., Urban, M., Sweigard, J. A. & Hamer, J. E. The *CPKA* gene of *Magnaporthe grisea* is essential for appressorial penetration. *Mol. Plant Microbe Interact.* **10,** 187–194 (1997).
19. Inoue, I., Namiki, F. & Tsuge, T. Plant colonization by the vascular wilt fungus *Fusarium oxysporum* requires *FOW1*, a gene encoding a mitochondrial protein. *Plant Cell* **14,** 1869–1883 (2002).
20. Farman, M. L. *et al.* Analysis of the Structure of the AVR1–CO39 avirulence locus in virulent rice-infecting isolates of *Magnaporthe grisea*. *Mol. Plant Microbe Interact.* **15,** 6–16 (2002).
21. Tian, D., Traw, M. B., Chen, J. Q., Kreitman, M. & Bergelson, J. Fitness costs of R-gene-mediated resistance in *Arabidopsis thaliana*. *Nature* **423,** 74–76 (2003).
22. Ou, S. H. *Rice Diseases* (Commonwealth Mycological Institute, Kew, Surrey, UK, 1985).
23. Sosnowski, M., Ramsey, M., Murray, G., Scott, E. & Wilmshurst, C. Symptoms of blackleg (*Leptosphaeria maculans*) on the roots of canola in Australia. *Plant Pathol.* **50,** 808 (2001).
24. Vereijssen, J., Schneider, H. J. M. & Termorshuizen, A. J. Possible root infection of *Cercospora beticola* in sugar beet. *Eur. J. Plant Pathol.* **110,** 103–106 (2004).
25. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: a Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989).
26. Lorang, J. M. *et al.* Green fluorescent protein is lighting up fungal biology. *Appl. Environ. Microbiol.* **67,** 1987–1994 (2001).
27. Rho, H. S., Kang, S. & Lee, Y. H. *Agrobacterium tumefaciens*-mediated transformation of the plant pathogenic fungus, *Magnaporthe grisea*. *Mol. Cells* **12,** 407–411 (2001).
28. Sweigard, J. A., Chumley, F., Carroll, A., Farrall, L. & Valent, B. A series of vectors for fungal transformation. *Fungal Genet. Newslett.* **44,** 52–53 (1997).
29. Resendes, C. M., Geil, R. D. & Guinel, F. C. Mycorrhyzal development in a low nodulating pea mutant. *New Phytol.* **150,** 563–572 (2001).
30. Gangopadhyay, S. & Row, K. V. K. Perennation of *Pyricularia oryzae* briosi *et* cav. in sclerotial state. *Int. J. Trop. Plant Dis.* **4,** 187–192 (1986).

Acknowledgements We thank the *M. grisea* community for providing fungal strains; L. Ciuffetti for the pCT74 reporter vector, and J. Deacon and R. Gutteridge for providing the images for Fig. 1l and 1m, respectively. A.S. has been supported by a Marie Curie fellowship from the European Community and by the Gatsby Charitable Foundation.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.E.O. (annie.osbourn@sainsbury-laboratory.ac.uk).

Structural basis for packaging the dimeric genome of Moloney murine leukaemia virus

Victoria D'Souza & Michael F. Summers

Howard Hughes Medical Institute and Department of Chemistry and Biochemistry, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, Maryland 21250, USA

All retroviruses specifically package two copies of their genomes during virus assembly, a requirement for strand-transfer-mediated recombination during reverse transcription[1,2]. Genomic RNA exists in virions as dimers, and the overlap of RNA elements that promote dimerization and encapsidation suggests

that these processes may be coupled[3–5]. Both processes are mediated by the nucleocapsid domain (NC) of the retroviral Gag polyprotein[3]. Here we show that dimerization-induced register shifts in base pairing within the Ψ-RNA packaging signal of Moloney murine leukaemia virus (MoMuLV) expose conserved UCUG elements that bind NC with high affinity (dissociation constant = 75 ± 12 nM). These elements are base-paired and do not bind NC in the monomeric RNA. The structure of the NC complex with a 101-nucleotide 'core encapsidation' segment of the MoMuLV Ψ site[6] reveals a network of interactions that promote sequence- and structure-specific binding by NC's single CCHC zinc knuckle. Our findings support a structural RNA switch mechanism for genome encapsidation, in which protein binding sites are sequestered by base pairing in the monomeric RNA and become exposed upon dimerization to promote packaging of a diploid genome.

The MoMuLV is a prototypical retrovirus widely used in human gene therapy trials and extensively studied as a model for retrovirus assembly and genome encapsidation[7]. The 350-nucleotide MoMuLV packaging signal (Ψ, Fig. 1a) forms a structurally organized domain that independently directs dimerization and packaging of the viral RNA[8]. The secondary structure of the Ψ site changes on dimerization[9], and it has been suggested that RNA conformational changes may help regulate genome packaging and other genome-related replication events[9,10]. Dimerization is promoted by two stem loops that form intermolecular duplexes (DIS-1, A204–G229; and

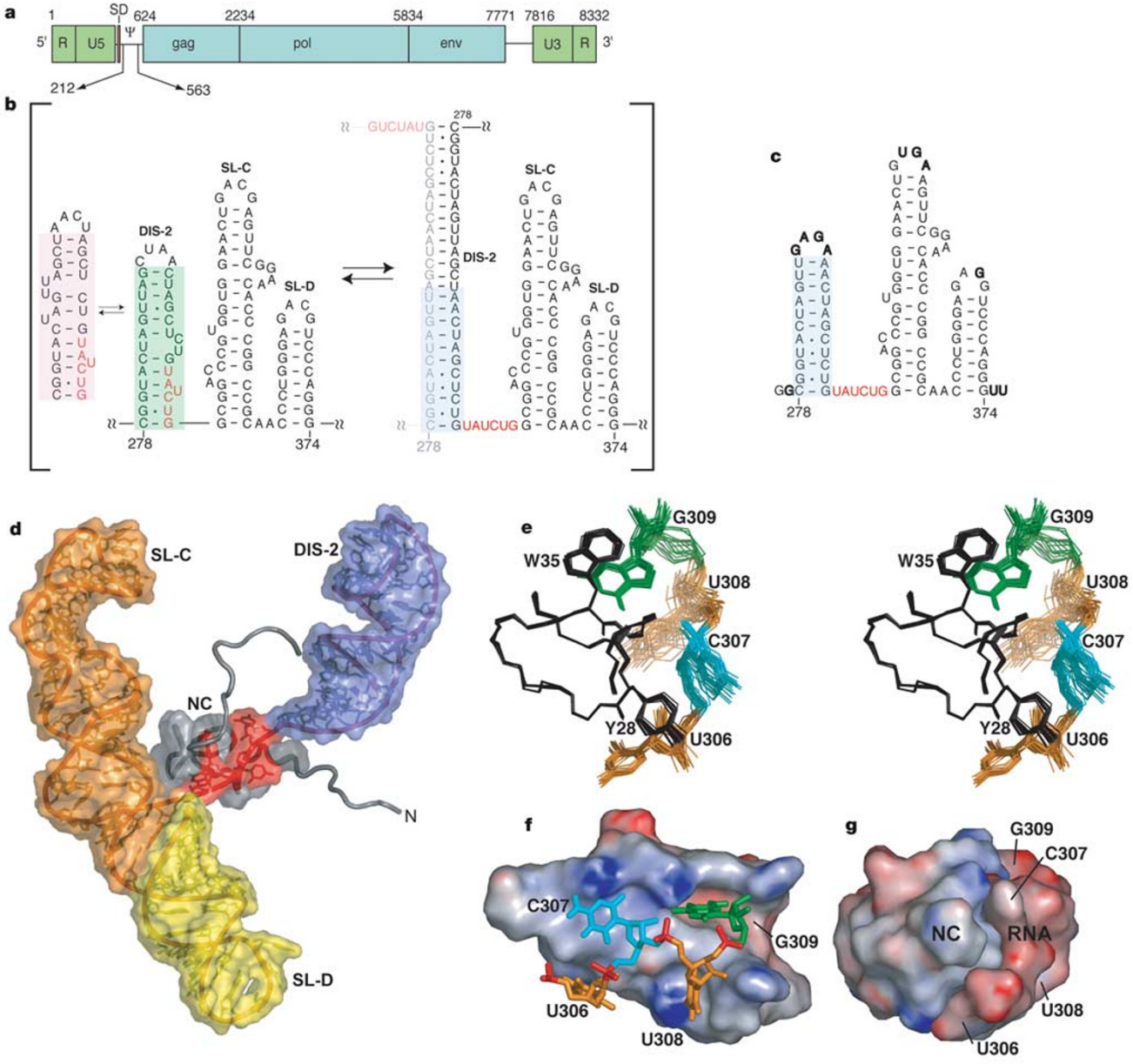

Figure 1 MoMuLV Ψ site and structure of the NC–mΨ^{CES} complex. **a**, Unspliced MoMuLV genome showing the location of the Ψ packaging signal. **b**, Secondary structure of the core encapsidation signal (Ψ^{CES}). DIS-2 exists in two alternate monomeric conformations (shaded red and green), and undergoes a frame shift upon dimerization (shaded blue) that exposes a UAUCUG element (red). **c**, Secondary structure of mΨ^{CES} with non-native nucleotides shown in bold. DIS-2 base pairings match those of the dimeric form of native Ψ^{CES}. **d**, Representative NC–mΨ^{CES} structure DIS-2 (blue), SL-C (orange) and SL-D (yellow), the UCUG segment (red) and NC (grey). **e**, Stereo image showing the best-fit backbone superposition of the CCHC zinc knuckle for the 20 calculated NC–mΨ^{CES} structures. **f**, Interactions between the zinc knuckle (coloured according to electrostatic surface potential) and the U306CUG element: U (orange), C (cyan) and G (green). **g**, Surface representations of the zinc knuckle and U306CUG segment showing complementarity of shape and electrostatic potential at the protein–RNA interface.

DIS-2, C278–G309)[9,11,12] and two additional stem loops that can form 'kissing complexes' through intermolecular base pairing of their tetraloop nucleotides (SL-C, G310–C352; and SL-D, C355–G374)[13]. Although all four stem loops are required for efficient encapsidation, fragments of Ψ containing SL-C, SL-D and portions of DIS-2 are capable of independently directing the packaging of heterologous RNAs into virus-like particles[6]. These residues comprise a 'core encapsidation signal' (Ψ^{CES}; Fig. 1b; ref. 6) that can bind NC stoichiometrically and with high affinity[14]. The DIS-2 stem loop of Ψ^{CES} undergoes a register-shift in base pairing upon dimerization[9], in which residues U304–G309 are base paired in the monomer but not in the dimer (Fig. 1b). To gain insights into the molecular determinants of genome packaging, we determined the structure of the complex between NC and a 101 nucleotide Ψ^{CES} mutant, engineered to remain monomeric in solution but retain internal base pairing of the dimer[14,15] ($m\Psi^{CES}$; Fig. 1c).

The NC–$m\Psi^{CES}$ structure demonstrated an unexpected binding mode, in which the single CCHC zinc knuckle of NC interacts with the UCUG linker that connects DIS-2 to SL-C (Fig. 1d–f). The binding interface exhibits significant complementarity of shape and charge (Fig. 1g). U306 and C307 pack against the side chain of Tyr 28, and U308 packs against the side chains of Ala 27 and Leu 21. The binding of these three nucleotides appears to be promoted by the following direct or water mediated interactions: U306-5′-phosphodiester to Tyr 28-OH, U306-O4 to Lys 42-NH_3^+, C307-N3 and/or -O2 atoms to Lys 42-NH_3^+, U308-O2′ to Arg 18 guanidinium group, U308-O4 to Lys 30-NH_3^+ and U308-5′-phosphodiester to Lys 37-NH_3^+. The G309 nucleotide base fits deeply into a pocket defined by the side chains of Leu 21, Ala 27, Trp 35 and Ala 36, and forms hydrogen bonds with backbone NH and O atoms located at the bottom of the pocket (G309-O6–Ala 27-NH; G309-O6–Ala 36-NH; G309-N1H–Gln 25-O; G309-NH21–Gln 25-O; Fig. 2a). Similar interactions with exposed guanosines have been observed in human immunodeficiency virus Type-1 (HIV-1) NC–RNA[16,17] and zinc knuckle–DNA[18] complexes (Fig. 2b), although for HIV-1 NC, tight binding requires two adjacent CCHC zinc knuckles, each of which interacts directly with only a single nucleotide base (G)[16,17]. Also, the HIV-1 pockets are 'pre-formed', being derived from residues on the surface of the folded zinc knuckle domain. In contrast, the guanosine binding pocket of the MoMuLV zinc knuckle is only partially formed in the absence of $m\Psi^{CES}$, with residues Arg 18–Asp 24 being disordered in the free protein and folding only on binding to the RNA. The remaining residues of the amino-terminal tail (Ala 1–Arg 17), and the carboxy-terminal tail (Arg 44–Leu 56), remain disordered on binding and do not interact tightly or specifically with the RNA. These findings are consistent with mutagenesis studies showing that basic residues between Arg 16 and Pro 43 are required for efficient genome packaging, whereas those in the flexible regions of NC (including the entire flexible portion of the C-terminal tail) are disposable[4,19].

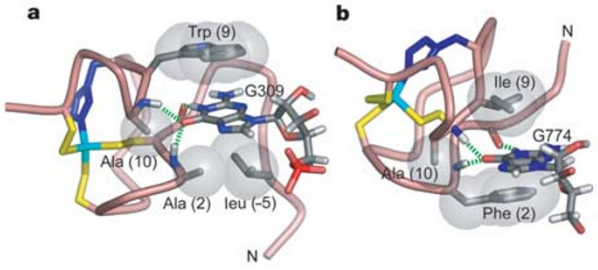

Figure 2 Comparison of the guanosine binding sites of the MoMuLV and HIV-1 zinc knuckles. **a**, MoMuLV zinc knuckle in NC–$m\Psi^{CES}$. **b**, HIV-1 zinc knuckle in NC–SL3 (ref. 16). For comparison, the numbering scheme used for both zinc knuckles begins with the first cysteine labelled Cys 1 (this figure only). The zinc atom (cyan) and cysteine (yellow) and histidine (blue) side chains are shown.

NMR chemical shifts and nuclear Overhauser effect (NOE) cross peak patterns of the DIS-2, SL-C and SL-D stem loops were not perturbed by NC binding, and their structures are identical to those observed for the free $m\Psi^{CES}$ RNA[15]. To determine if transient interactions with these elements might contribute to binding, NC titration experiments were performed with truncated forms of $m\Psi^{CES}$. Tight binding (dissociation constant (K_d) < 200 nM) was observed for all fragments that included the linker residues, including the shortest fragment tested, r-UAUCUG (K_d = 75 ± 12 nM; Supplementary Fig. S1). For comparison, NC binds native Ψ^{CES} with a dissociation constant of 132 ± 55 nM (ref. 14). Binding was not observed for any RNAs that lacked the linker[14].

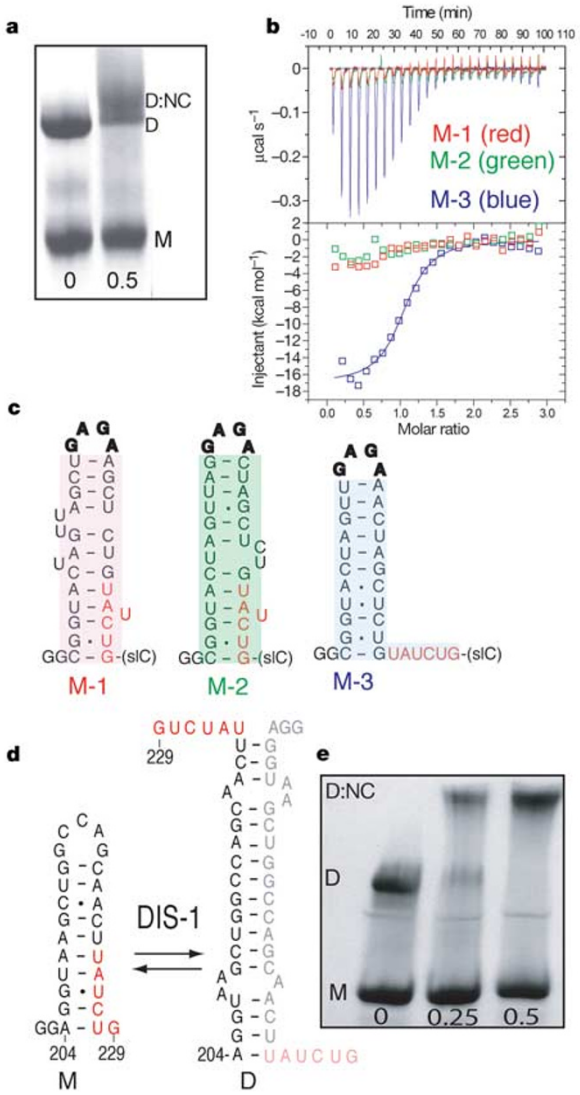

Figure 3 NC binds to dimeric forms of DIS-1 and DIS-2. **a**, Native PAGE data obtained for an RNA construct corresponding to residues G276–C352 of native Ψ^{CES} (DIS2–C) in the absence (left) and presence (right) of NC (0.5 equivalents). The titration was performed under equilibrium conditions of monomeric (M) and dimeric (D) species, and a band shift was observed only for the dimeric species. **b**, **c**, ITC NC titration results obtained for DIS2–C constructs containing GNRA-type mutations (bold) that stabilize base pairings of the monomeric (M-1 and M-2) and dimeric (M-3) Ψ site. All three constructs are monomeric under experimental conditions, based on native gel electrophoresis. Base pairings were confirmed by 2D ^{1}H NMR. Colour shadings and secondary structures correspond to those shown in Fig. 1b for the native Ψ site. NC binds M-3 with significant affinity (K_d = 173 ± 32 nM), but does not interact tightly with M-1 or M-2. **d**, Secondary structures determined by 2D ^{1}H NMR for a 28 nucleotide DIS-1 RNA in its monomeric (M) and dimeric (D) forms. **e**, Native PAGE data showing that only the dimeric form of DIS-1 binds NC.

In addition, NMR chemical shifts and NOE cross peak patterns in two-dimensional (2D) NOESY data obtained for the NC–r(UAUCUG) complex matched the data obtained for NC–mΨ^{CES} complex (Supplementary Figs S2 and S3). These data indicate that the UCUG linker element, which is conserved among the murine C-type retroviruses[15], is both necessary and sufficient for high affinity NC binding. DIS-2 contains a second UCUG sequence that is also conserved (U300CUG), but this element is sequestered by intramolecular base pairing and stacking interactions and does not interact with NC.

UCUG sequences are unusually abundant in the region between the primer binding site (PBS) and Gag initiation codon of the mammalian C-type retroviruses, occurring at a frequency of approximately 1 in 50 nucleotides. By comparison, these sequences occur at a frequency of 1 in 225 nucleotides in the coding and LTR (long terminal repeat) regions of the MoMuLV and other retroviral genomes reported in the NIH NCBI Gene Bank. Thirteen UCUG sequences are present in the 418-nucleotide segment between the PBS and Gag start sites of the MoMuLV genome, accounting for more than 12% of the nucleotides in this region. Five of these sequences are located within the 55 nucleotides between Ψ and the Gag initiation codon, and this may explain why 'Ψ+' RNAs (that is, RNAs with a 3′-extended Ψ site) package heterologous RNAs with greater efficiency than Ψ-only RNAs[20]. Not all of the UCUG segments are likely to bind NC owing to their participation in intramolecular interactions (as observed for U300CUG). However, the high abundance of these sequences is consistent with proposals that the NC domains of assembling Gag molecules bind cooperatively to multiple sites within the Ψ-RNA packaging signal. No UCUG sequences were found between the PBS and Gag initiation codons of representative lentivirus (HIV-1) and D-type retrovirus (Mason-Pfizer monkey virus) genomes, and fewer than three UCUG sequences were observed in representative B-type (mouse mammary tumour virus), avian C-type (Rous sarcoma virus) and HTLV/BLV (human T-lymphotrophic virus-I) retrovirus genomes. The NC proteins of these retroviruses contain two CCHC zinc knuckle domains with amino acid compositions that differ significantly from those of the mammalian C-type retroviruses. These differences may explain why chimaeric HIV-1 virions containing the MoMuLV NC domain are able to efficiently package the MoMuLV genome but unable to package the HIV-1 genome[21], and are consistent with the observation that retroviruses of a given genus are often capable of packaging each other's genomes[22].

Previous chemical accessibility mapping and free energy calculations suggested that base pairing within the DIS-2 stem loop of the intact MoMuLV Ψ site changes upon dimerization[9]. NMR data obtained for RNAs containing the native DIS-2 sequence confirmed the predicted structures shown in Fig. 1b (Supplementary Fig. S4). The fact that residues U306–G309 are internally base-paired in the monomeric hairpin and exposed in the duplex[9] raised the possibility that NC might bind preferentially to the duplex. To test this hypothesis, a 77 nucleotide RNA corresponding to stem loops DIS-2 and SL-C (DIS2–C) of the native Ψ^{CES} RNA was prepared for NC-binding studies. Titration of DIS2–C with NC resulted in a small but distinguishable electrophoretic band shift for the dimeric species, but did not affect the mobility of the monomer (Fig. 3a; ref. 15). To quantitatively test for structure-dependent NC binding, isothermal titration calorimetry (ITC) data were obtained for mutants of DIS2–C containing GNRA sequences that specifically stabilize base pairings observed in the monomeric and dimeric Ψ site (see Figs 1b and 3b, c). As shown in Fig. 3b, RNAs with base pairings of the monomer (M1 and M2 in Fig. 3c) did not bind NC, whereas the construct with base pairings of the dimer (M3) bound NC with high affinity. These findings confirm that NC binds only to the dimeric form of DIS–2, in which the U306CUG element is exposed.

The MoMuLV Ψ site contains a second palindromic segment that promotes genome dimerization *in vitro*[11,12] and encapsidation *in vivo*[12], and this segment (DIS-1, A204–G229) has also been predicted to undergo a register shift in base pairing on conversion from a monomeric hairpin to a dimeric duplex[11]. DIS-1 also contains a 3′-UCUG sequence that could potentially bind NC. A 28-nucleotide RNA with sequence of DIS-1 (G202–G229, Fig. 3d) forms monomers and dimers at RNA concentrations above 0.1 mM, and titration of DIS-1 with NC resulted in a significant electrophoretic band shift for the dimeric (but not the monomeric) species (Fig. 3e). 2D NOESY experiments confirmed that the UCUG segment is base-paired in the hairpin and exposed in the duplex, and that NC binding is mediated by interactions between the CCHC zinc knuckle and the exposed UCUG element of the duplex (data not shown). Thus, DIS-1 exhibits dimerization-dependent structural and NC-binding behaviour very similar to that observed for DIS-2.

Our findings support a structural RNA switch mechanism for genome encapsidation, in which conserved UCUG elements associated with DIS-1 and DIS-2 are sequestered in the monomeric RNA and become exposed to promote high affinity NC binding on dimerization (Fig. 4). In both cases, asymmetry within the pseudo-palindromic stem loops allows dimerization-induced register shifts that re-optimize base pairing in the dimer and expose the protein binding site. These findings are consistent with studies suggesting that genome dimerization and encapsidation events are intimately coupled[5,7,23–27]. Conserved UCUG elements downstream of SL-D that exhibit dimerization-dependent chemical reactivity (U510CUG and U559CUG)[9] may also bind NC in a dimerization-dependent manner. Further studies of these and the five UCUG elements in the segment downstream of Ψ that is required for optimal packaging efficiency[20] are warranted. Dimerization-dependent conformational changes have recently been detected in the HIV-1 Ψ site[10], and it seems that at least one high affinity NC binding site that is present in the dimeric Ψ site (stem loop SL2)[17] exists in a significantly different conformation in the monomeric RNA. Conformational RNA switch mechanisms may therefore be commonly used by retroviruses to promote specific packaging of a diploid genome. □

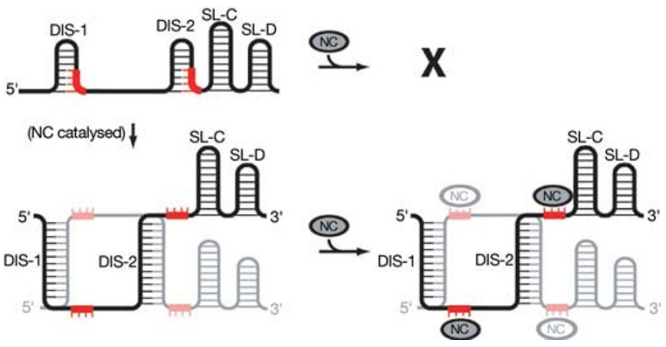

Figure 4 Structural changes in the Ψ-site that seem to serve as a switch for the selective binding of NC to the dimer. Conserved UCUG bases (red) are base paired in the monomeric state and become exposed for NC binding upon dimerization. Dimerization may promote exposure of additional downstream UCUG elements to enhance the specific packaging of a diploid genome.

Methods

RNA design

The GACG tetraloops of SL-C and SL-D were mutated to GUGA and GAGG, respectively, to eliminate intermolecular kissing interactions that can lead to aggregation at concentrations required for NMR studies[14]. These mutations do not affect the structures of the stem loops or the NC binding properties of the RNA[14,15]. In addition, the palindromic AGCU segment of DIS-2 was replaced by GAGA, a GNRA-type tetraloop that stabilizes hairpin structures[28]. This mutation stabilizes base pairings in the stem of DIS-2 that match those observed for the native, dimeric Ψ site (see Fig. 1b, c). The resulting mΨ^{CES} RNA remains monomeric at concentrations well above 1.0 mM (based on native

gel electrophoresis and NMR) and binds NC with the same affinity and stoichiometry observed for the native Ψ^{CES} RNA[14].

Sample preparation

MoMuLV NC protein and RNA constructs were prepared as described[14,15]. RNAs of 35 nucleotides or less were obtained from Dharmacon and purified by denaturing gel electrophoresis. Samples for all NMR, ITC and polyacrylamide gel electrophoresis (PAGE) measurements were prepared in Tris-HCl buffer (10 mM at pH 7.0, 10 mM NaCl, 0.1 mM $ZnCl_2$ and 0.1 mM β-mercaptoethanol).

NC binding experiments

ITC data (VP-ITC calorimeter, MicroCal Corp.) were measured at 30 °C. Exothermic heats of reaction were measured for 25 injections of NC protein (80 μM) into 1.4 ml of RNA (5.0 μM). Binding curves were analysed by nonlinear least squares fitting of the baseline corrected data to a single binding site model as described[14]. RNAs for PAGE experiments (0.5 mM) were heated for 2 min at 90 °C and cooled on ice before addition of NC. Samples were loaded onto native polyacrylamide gels and electrophoresed at 4 °C in Tris-borate buffer (45 mM Tris base, 45 mM boric acid, pH 8.3), and gels were stained with Stains-All (Sigma).

NMR spectroscopy and signal assignments

NMR data (Bruker DRX spectrometer, 800 MHz ^{1}H, T = 15, 25 and 35 °C) were obtained from a combination of two-, three- and four-dimensional NOESY data, for samples with combinations of natural abundance and ^{15}N-,^{13}C-labelled NC and both ^{15}N-,^{13}C-labelled and perdeuterated/selectively protonated RNA. RNA signals were assigned as described for the free mΨ^{CES} RNA[15]. Protein backbone signals were assigned using standard triple resonance methods, and side chain signals were assigned from 3D and 4D ^{15}N-, ^{13}C-, and ^{15}N/^{13}C-edited NOESY data (see refs 16, 29, 30, and citations there-in). Titration of mΨ^{CES} with NC resulted in NMR chemical shift changes and intermolecular ^{1}H–^{1}H NOEs for nucleotides U306–G309. All other mΨ^{CES} NMR signals were unaffected by NC binding. Intermolecular NOEs were readily assigned from 2D NOESY data obtained for samples containing nucleotide-specifically protonated RNA (G^H-mΨ^{CES}, A^H-mΨ^{CES}, U^H-mΨ^{CES}, C^H-mΨ^{CES}; containing protonated G, A, U and C, respectively, with the remaining nucleotides perdeuterated), in combination with 3D ^{13}C-edited NOESY data obtained for RNA complexes with ^{15}N-,^{13}C-labelled NC. Some RNA NMR signals were significantly broadened on titration with NC, and in these cases, intermolecular NOEs were identified or confirmed on the basis of exchange-mediated intermolecular NOEs with the unbound RNA in samples containing a ~30% excess of RNA (see Supplementary Fig. S2). Subsequent to solving the structure of the NC–mΨ^{CES} complex, 2D NOESY data were obtained for the NC complex with r-UAUCUG (Supplementary Fig. S3), which enabled confirmation of assignments made for NC–mΨ^{CES} but did not lead to the identification of additional intermolecular NOEs. Backbone NH signals for residues of the N- and C-terminal tails (Ala 1–Arg 17 and Arg 44–Leu 56, respectively) were sensitive to exchange with H_2O protons and exhibited random coil chemical shifts (as also observed for the Cα carbons).

Structure calculations

Upper interproton distance bounds of 2.7, 3.3 and 5.0 Å were employed for NOE cross peaks of strong, medium and weak intensity respectively, which were qualitatively assessed following intensity normalization of the different NOE data sets. Structures were calculated in torsion angle space with CYANA (http://www.las.jp/index_eg.html) starting from random initial angles (involving both protein and RNA), and stem loops DIS-2, SL-C and SL-D were refined independently using dipolar couplings obtained for the isolated stem loops. This approach is valid because the NMR chemical shifts and NOE cross peak patterns and intensities of the isolated stem loops and those in the NC–mΨ^{CES} complex were indistinguishable[15]. Statistical information and superposition images are provided in Supplementary Table S1 and Fig. S5, respectively. Structure figures were generated with PyMOL (http://pymol.sourceforge.net).

Received 8 July; accepted 16 August 2004; doi:10.1038/nature02944.

1. Varmus, H. E. Form and function of retroviral proviruses. *Science* **216,** 812–820 (1982).
2. Hu, W. S. & Temin, H. M. Retroviral recombination and reverse transcription. *Science* **250,** 1227–1233 (1990).
3. Prats, A.-C. *et al. Cis* elements and *trans*-acting factors involved in dimer formation of murine leukemia virus RNA. *J. Virol.* **64,** 774–783 (1990).
4. Housset, V., De Rocquigny, H., Roques, B. P. & Darlix, J.-L. Basic amino acids flanking the zinc finger of Moloney murine leukemia virus nucleocapsid protein NCp10 are critical for virus infectivity. *J. Virol.* **67,** 2537–2545 (1993).
5. Paillart, J.-C., Shehu-Xhilaga, M., Marquet, R. & Mak, J. Dimerization of retroviral RNA genomes: an inseparable pair. *Nature Rev. Microbiol.* **2,** 461–472 (2004).
6. Mougel, M. & Barklis, E. A role for two hairpin structures as a core RNA encapsidation signal in murine leukemia virus virions. *J. Virol.* **71,** 8061–8065 (1997).
7. Berkowitz, R., Fisher, J. & Goff, S. P. RNA packaging. *Curr. Top. Microbiol. Immunol.* **214,** 177–218 (1996).
8. Mann, R., Mulligan, R. C. & Baltimore, D. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell* **33,** 153–159 (1983).
9. Tounekti, N. *et al.* Effect of dimerization on the conformation of the encapsidation psi domain of Moloney murine leukemia virus RNA. *J. Mol. Biol.* **223,** 205–220 (1992).
10. Huthoff, H. & Berkhout, B. Two alternating structures of the HIV-1 leader RNA. *RNA* **7,** 143–157 (2001).
11. Oroudjev, E. M., Kang, P. C. E. & Kohlstaedt, L. A. An additional dimer linkage structure in Moloney murine leukemia virus RNA. *J. Mol. Biol.* **291,** 603–613 (1999).
12. Ly, H. & Parslow, T. G. Bipartite signal for genomic RNA dimerization in the Moloney murine leukemia virus. *J. Virol.* **76,** 3135–3144 (2002).
13. Kim, C.-H. & Tinoco, I. Jr. A retroviral RNA kissing complex containing only two G–C base pairs. *Proc. Natl Acad. Sci. USA* **97,** 9396–9401 (2000).
14. D'Souza, V. *et al.* Identification of a high-affinity nucleocapsid protein binding site within the Moloney murine leukemia virus Ψ-RNA packaging signal. Implications for genome recognition. *J. Mol. Biol.* **314,** 217–232 (2001).
15. D'Souza, V., Dey, A., Habib, D. & Summers, M. F. NMR structure of the 101 nucleotide core encapsidation signal of the Moloney murine leukemia virus. *J. Mol. Biol.* **337,** 427–442 (2004).
16. De Guzman, R. N. *et al.* Structure of the HIV-1 nucleocapsid protein bound to the SL3 Ψ-RNA recognition element. *Science* **279,** 384–388 (1998).
17. Amarasinghe, G. K. *et al.* NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the Ψ-RNA packaging signal. *J. Mol. Biol.* **301,** 491–511 (2000).
18. Schuller, W., Dong, C.-Z., Wecker, K. & Roques, B.-P. NMR structure of the complex between the zinc finger protein NCp10 of Moloney murine leukemia virus and the single-stranded pentanucleotide d(ACGCC): Comparison with HIV-NCp7 complexes. *Biochemistry* **38,** 12984–12994 (1999).
19. Rein, A., Harvin, D. P., Mirro, J., Ernst, S. M. & Gorelick, R. J. Evidence that a central domain of nucleocapsid protein is required for RNA packaging in murine leukemia virus. *J. Virol.* **68,** 6124–6129 (1994).
20. Bender, M. A., Palmer, T. D., Gelinas, R. E. & Miller, A. D. Evidence that the packaging signal of Moloney murine leukemia virus extends into the *gag* region. *J. Virol.* **61,** 1639–1646 (1987).
21. Berkowitz, R. D., Ohagen, A., Hoglund, S. & Goff, S. P. Retroviral nucleocapsid domains mediate the specific recognition of genomic viral RNAs by chimeric Gag polyproteins during RNA packaging *in vivo*. *J. Virol.* **69,** 6445–6456 (1995).
22. Yang, S. & Temin, H. M. A double hairpin structure is necessary for the efficient encapsidation of spleen necrosis virus retroviral RNA. *EMBO J.* **13,** 713–726 (1994).
23. Hibbert, C. S., Mirro, J. & Rein, A. mRNA molecules containing MLV packaging signals are encapsidated as dimers. *J. Virol.* (in the press).
24. Levin, J. G., Grimley, P. M., Ramseur, J. M. & Berezesky, I. K. Deficiency of 60 to 70S RNA in murine leukemia virus particles assembled in cells treated with actinomycin D. *J. Virol.* **14,** 152–161 (1974).
25. Sakuragi, J.-I., Shioda, T. & Panganiban, A. T. Duplication of the primary encapsidation and dimer linkage region of Human immunodeficiency virus type 1 RNA results in the appearance of monomeric RNA in virions. *J. Virol.* **75,** 2557–2565 (2001).
26. Torrent, C., Bordet, T. & Darlix, J.-L. Analytical study of rat retrotransposon VL30 RNA dimerization *in vitro* and packaging in murine leukemia virus. *J. Mol. Biol.* **240,** 434–444 (1994).
27. Méric, C. & Goff, S. P. Characterization of Moloney murine leukemia virus mutants with single-amino-acid substitutions in the Cys-His box of the nucleocapsid protein. *J. Virol.* **63,** 1558–1568 (1989).
28. Antao, V. P., Lai, S. Y. & Tinoco, I. Jr. A thermodynamic study of unusually stable RNA and DNA hairpins. *Nucleic Acids Res.* **19,** 5901–5905 (1991).
29. Wüthrich, K. *NMR of Proteins and Nucleic Acids* (John Wiley, New York, 1986).
30. Kay, L. E., Clore, G. M., Bax, A. & Gronenborn, A. M. Four-dimensional heteronuclear triple-resonance NMR spectroscopy of interleukin-1β in solution. *Science* **249,** 411–414 (1990).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements We thank the HHMI staff and colleagues at UMBC for technical support. This research was supported by a grant from the NIH to M.F.S.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.F.S. (summers@hhmi.umbc.edu).

Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins

Lars Ferbitz[1], Timm Maier[1], Holger Patzelt[2], Bernd Bukau[2], Elke Deuerling[2] & Nenad Ban[1]

[1]*Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule Hönggerberg (ETH Zürich), HPK Gebäude, CH-8093 Zürich, Switzerland*
[2]*Zentrum für Molekulare Biologie (ZMBH), Universität Heidelberg, Im Neuenheimer Feld 282, D-69120 Heidelberg, Germany*

During protein biosynthesis, nascent polypeptide chains that emerge from the ribosomal exit tunnel encounter ribosome-associated chaperones, which assist their folding to the native state[1,2]. Here we present a 2.7 Å crystal structure of *Escherichia coli* trigger factor, the best-characterized chaperone of this type, together with the structure of its ribosome-binding