

The nucleocapsid protein as a target for novel anti-HIV drugs

Andrew G Stephen, Alan Rein, Robert J Fisher & Robert H Shoemaker
SAIC-Frederick & National Cancer Institute, USA



Combination chemotherapy with regimens containing reverse transcriptase and protease inhibitors (HAART) has resulted in major reductions in the morbidity and mortality associated with AIDS. Curative therapy, however, is limited by the

development of drug resistance and persistence of infection in a latent form. One approach to overcoming the problem of resistance is to develop novel agents against additional targets that can be added to combination regimens. The HIV-1 nucleocapsid protein may be particularly important in this context, as mutations in this protein result in greatly reduced viral fitness.

The National Cancer Institute at Frederick has been a major center for HIV research since the late 1980s. The first large-scale screen for anti-HIV drugs was established by Weislow et al in 1989 as part of the Developmental Therapeutics Program at Frederick and yielded discoveries that resulted in 2 of the 17 FDA approved drugs for treatment of patients with AIDS, ie 3TC and carbovir (Ziagen).

A nucleus of investigators interested in basic science, vaccine development, and chemotherapy evolved at NCI-Frederick and has made seminal contributions in these areas. The expertise and key reagents available has enabled development of a molecular-targeted screen with the potential to identify novel inhibitors of various functions of the nucleocapsid protein. This strategy contrasts with earlier, cell-based assays that were not mechanism-directed.

The nucleocapsid protein

All retroviruses encode a Gag polyprotein that is produced in the host cell during the later stages of infection. This protein recognizes specific sequence elements of the viral RNA that direct the viral encapsidation process and particle assembly. After budding from the cell membrane, the virus

undergoes maturation, a process induced by the proteolytic processing of Gag by the viral protease into matrix (MA), capsid (CA) and nucleocapsid (NC) proteins, which then rearrange to form the mature infectious particle.

In HIV-1, NC is involved in multiple functions throughout the life cycle of the

virus (Figure 1). As a domain of Gag, NC promotes the packaging of the viral genome, a process that requires the recognition of a specific viral packaging sequence (Ψ element). This Ψ element is about 120 nucleotides long and consists of 4 stem loop structures. The NC domain is required for the annealing of the tRNA^{Lys,3}

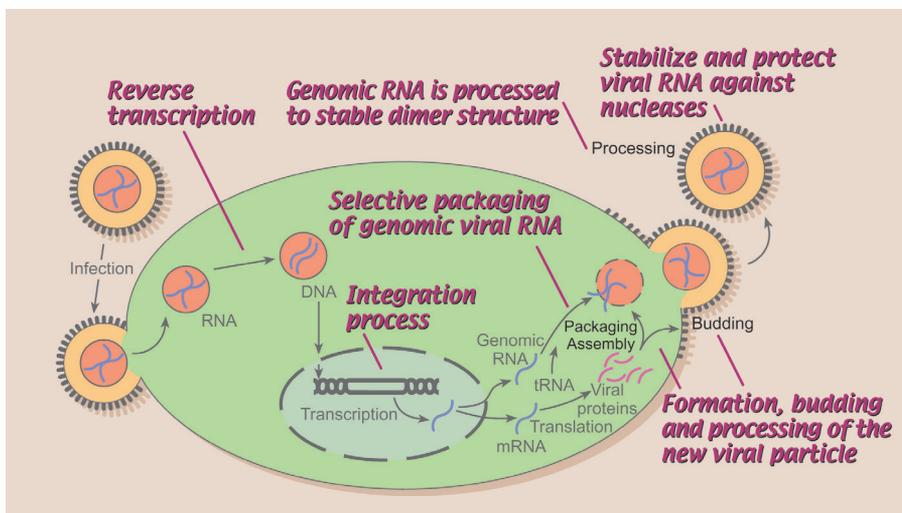


Figure 1. Functions of NC and the NC domain of Gag during the HIV-1 replication cycle. NC improves the efficiency of reverse transcription, processes genomic RNA into a stable dimer structure and enhances the integration process. The NC domain of Gag is involved in the selective packaging of the genomic viral RNA and anneals a cellular tRNA to the genomic RNA. A domain of Gag NC also plays a role in the formation of budding and processing of the new viral particle. Finally NC stabilizes and protects the viral RNA against nucleases (adapted from Drullenne and Roques (2000) *Drug News Perspect* 13(6)).

to the primer binding site. After cleavage, NC binds non-specifically to the viral genome to form the nucleocapsid core and protect the genome from nucleases. When the particle infects a new host cell, NC is a crucial co-factor in reverse transcription and integration. We have found that it binds with low nanomolar affinity to the repeating sequence (UG)_n, but the biological significance of this high affinity is not clear at present.

NC also has several roles as a nucleic acid chaperone, many of which are required for efficient reverse transcription. NC limits the pausing of reverse transcriptase by removing regions of secondary structure and promotes effective strand transfer. In addition NC interacts with integrase and favors the integration of the provirus into the DNA of host cells. NC also stabilizes the formation of the genomic viral RNA into a dimer.

Structure

NC is a basic protein (pI 10-11) comprising only 55 amino acids. Its major structural feature is the two zinc fingers (Cys-X₂-Cys-X₄-His-X₄-Cys, where X is a variable amino acid) that are analogous to those found in many cellular DNA binding proteins. The Cys and His residues coordinate zinc with extremely high affinity to produce these stable structural domains. The remaining portions of the protein are punctuated with a large number of basic amino acids (5 basic residues within the 14 amino acids at the N-terminus and 3 residues within the 7 amino acid linker sequence that separates the two fingers).

Michael Summers' group at the University of Maryland Baltimore County determined two structures of NC bound to either stem loop 2 or stem loop 3 of the Ψ element. The zinc fingers bind to exposed guanosine bases and the N-terminal region forms a 3₁₀ helix as it interacts with the RNA in both structures. However there are also many differences between the two structures, most notably the orientation of the two zinc fingers differs when bound to the different stem loops. These two structures most likely reflect a 'snapshot' of the many possible conformations that NC can adopt when bound to nucleic acid. The

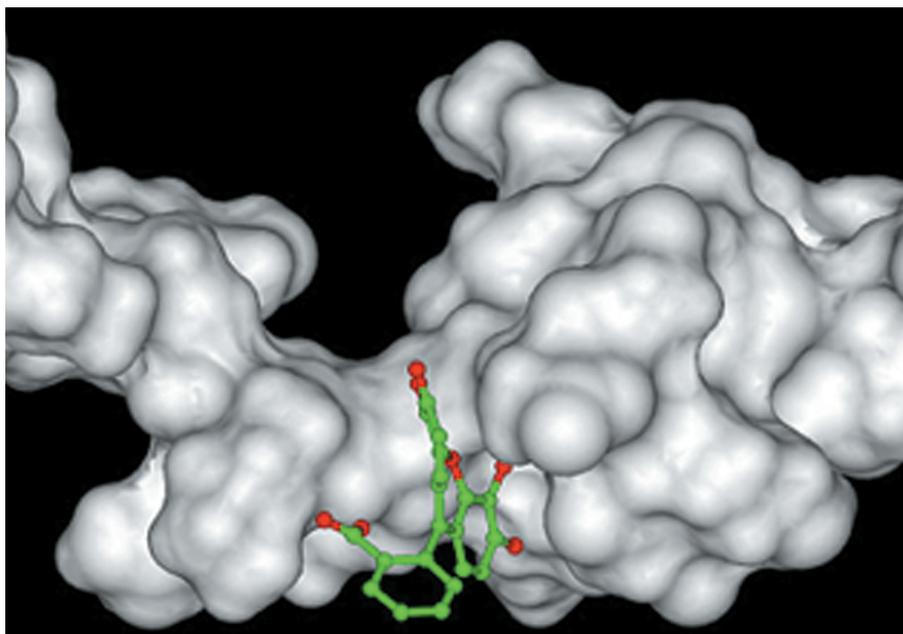


Figure 2. Binding of tetrachlorogallein to the carboxy terminal zinc finger of NC. This view is selected to show the putative binding pocket, which appears to match the shape of the ligand quite well. It should be noted that the bend in the tetrachlorogallein is an artifact of the docking program and the tricyclic ring structure should remain planar *in vivo* (reprinted with permission from *Biochemistry Biophysics Research Communication* (2002) 296:1228-1237).

inherent flexibility of the N-terminal region facilitates the formation of many different structures of NC bound to different RNA targets.

A drug target?

Its multitude of crucial functions in both early and late stages of the viral lifecycle and the relatively low tolerance to point mutations make the NC protein an attractive candidate for drug development. There has been considerable work directed towards identifying compounds that attack the zinc fingers, an effort that was initiated by William Rice and co-workers while at the NCI in Frederick. To date, these studies have identified several different classes of oxidizing agents that are able to attack the sulfhydryl group of the Cys residue and eject the zinc, rendering the zinc fingers and by extension, the whole protein, inactive. The most recent derivatives of these zinc ejectors are the pyridinioalkanoyl thioester (PATE) group of compounds that specifically target NC zinc fingers, are viricidal, and show good water solubility. Using a similar approach, The AIDS Vaccine Program at the NCI-Frederick has used aldrithiol (a commercially available disulfide) to inactivate zinc fingers in NC in

whole HIV-1 viruses as a source of material for large-scale viral protein purification or as vaccine and immunological reagents.

Other approaches to anti-NC compounds have included peptidomimetics of the three dimensional structure of NC. These cyclic peptides show antiviral activity, and although they do not directly bind to NC they are able to inhibit some of the annealing functions of NC.

Non-redox small molecule inhibitors

In 1998, we observed a high-affinity interaction between NC and repeating bases of thymidines and guanines, and realized that a small molecule inhibitor of this interaction could not only help in further understanding the mechanisms of NC binding to nucleic acids but could also have potential as a drug lead. Using highly pure and extremely well characterized NC protein provided by Robert Gorelick of the AIDS Vaccine Program at NCI-Frederick, an ELISA-based approach was developed, using NC immobilized by adsorption onto a polystyrene 96-well plate. A biotinylated oligonucleotide, d(TG)₄, was added to the NC in the presence or absence of 10 μM of compound, and oligo binding monitored using a luminescence detection system.

Any positives were verified using a surface plasmon resonance (SPR) secondary assay in which d(TG)₄ was immobilized on a sensor chip, then exposed to NC that had been preincubated with increasing amounts of the active compound.

The NCI has assembled a repository of over 140,000 compounds that have been submitted for screening purposes, from which a subset of ~2000 compounds (the 'Diversity Set') represents a broad range of the 3D pharmacophores it contains. Screening of the latter using our assay identified 26 inhibitors of which 5 had a related fluorescein-like structure, suggesting a potential pharmacophore to inhibit NC binding to nucleic acids.

An additional 63 fluorescein-like compounds were then identified in the larger compound repository. The SPR assay we used measures the ability of the compounds to inhibit the binding of NC to an immobilized oligo, but does not demonstrate direct binding of the compound to the NC protein. By measuring quenching of the single tryptophan found in NC we assayed the direct binding of the fluorescein-like compounds to NC. For high-affinity binders, we saw a good correlation between NC inhibitor activity and direct binding activity. Our lead compound (NSC-119889 or tetrachlorogallein) inhibited NC binding to nucleic acid stoichiometrically, whereas fluorescein showed no activity.

Antiviral activity

Concurrent with these studies, the 63 related compounds were submitted for antiviral testing using a cytoprotection assay based on infected or non-infected CEM-SS cells (a human T-cell line). In this manner we could rapidly evaluate any cytoprotective or toxic effects of the compounds. In general, more antiviral activity was observed with compounds that bound with the highest affinity to NC protein, while tetrachlorogallein analogs missing two hydroxyls in positions 4' and 5' in the xanthenyl ring structure did not bind NC.

Molecular modeling studies of tetrachlorogallein predicted that the compound would bind electrostatically to the amide nitrogen of Gly³⁵, with other contacts at the carbonyl oxygens of Gly⁴⁹ and Lys³³ (Figure 2). The modeling program also pre-

dicted that tetrabromofluorescein would have a much weaker affinity to NC, as observed experimentally. We concluded that these compounds bind near the first zinc finger and somehow alter its conformation, severely reducing the ability of the protein to bind effectively to nucleic acids.

Second-generation inhibitors

As a continuation of this work we sought to identify compounds active against NC when it was part of the Gag polyprotein, specifically inhibitors of the assembly of Gag into virus particles. This was possible since the Gag polyprotein, in the presence of nucleic acid and under the correct buffer conditions, will assemble into virus-like particles.

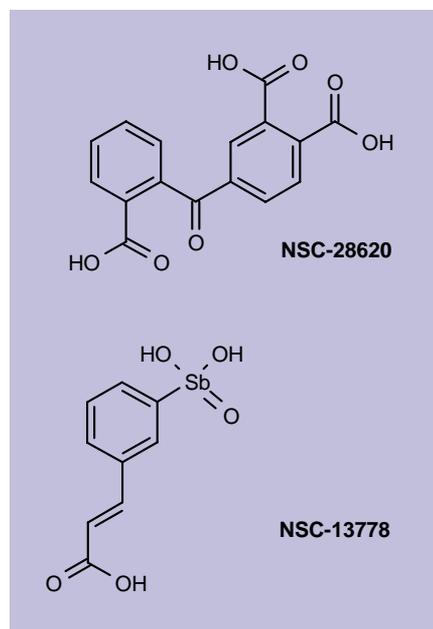


Figure 3. Second-generation nucleocapsid inhibitors.

The NC domain of Gag is presumed to play a key function when Gag binds to nucleic acid during this assembly process, and accordingly, several of our tetrachlorogallein analogs were able to disrupt the virus-like particles. Further screening of the NCI Diversity Set was carried by incubating test compounds (10 μ M) for 30 min with virus-like particles assembled from biotinylated tRNA, which identified a further 36 compounds as positives. A secondary assay monitoring the disruption of particles assembled in vitro reduced the number of positives to 6, representing

several chemotypes. As expected, tetrabromogallein identified in our NC screen, was among these positives, in addition to several sulfonic acid-containing compounds. Of particular interest to us were two compounds, NSC-13778 and NSC-28620 (Figure 3) that had also been identified in our primary NC screen.

NSC-13778 in particular represented an unusual class of organoantimony compounds that we believed warranted further study. In order to generate structure-activity information, 23 compounds related to NSC-13778, and a group of corresponding structures containing arsenic, were identified in the NCI repository and assayed for disassembly and antiviral activity. Again, compounds that were active in particle disassembly were also able to protect cells against infection with HIV-1. NSC-28620 also protected cells from HIV-1 infection and data are currently being evaluated from a structurally related series.

It should be pointed out that neither NSC-13778 nor -28620 was active in a DNA-ethidium bromide displacement assay, further supporting the idea these compounds are binding to the protein. In order to investigate this further, we established a fluorescence anisotropy assay where d(TG)₁₀ was labeled at the 3' end with fluorescein and the binding of NC to d(TG)₁₀ was monitored by a decrease in the rotation of the labeled oligo, manifested as an increase in the fluorescence anisotropy. The two compounds were relatively ineffective at releasing d(TG)₁₀ bound to NC. However if NC with zinc fingers inactivated by EDTA was used, then these compounds were quite effective at releasing the bound oligo from NC ($K_d \sim 400$ nM for both NSC-13778 and NSC-28620). Evaluation of compounds related to NSC-13778 also identified some with K_d values of 150 nM. In addition, those compounds that had antiviral and disassembly activity also were the most active in the NC fluorescence anisotropy assay.

Further work

Detailed studies of the gallein-type nucleocapsid inhibitors has revealed pH-dependencies for activity that substantially limit their value as drug development leads, so

we are currently focused on evaluating the antiviral activity of the other chemotypes identified. A substantial body of structure-activity information is available for the organoantimony lead NSC-13778 that supports the need for the pentavalent antimony atom in contrast to other atoms such as arsenic. We plan to measure the direct binding of these lead molecules to NC immobilized on a BIAcore sensor surface to determine association and dissociation rates. Using this same system, we can also determine how the different analogs bind to serum proteins, further equipping us to optimize these leads as potential therapeutics.

Currently we are seeking CRADA (Collaborative Research and Development Agreement) partners with medicinal and

synthetic chemical expertise to work with us towards identification of a clinical development candidate. In addition, we hope that identification of novel classes of inhibitors of NC will aid in understanding how this small protein performs such a wide variety of functions during the viral lifecycle.

Acknowledgement

Work reported here resulted from the efforts of a team of dedicated investigators. We would like to acknowledge the contributions of Shizuko Sei, Paula Roberts, Ya-Xiong Feng, Mike Currens, Paul Klausmeyer, Rhone Akee, Thomas McCloud, John Cardellina, David Covell, Rob Gorelick, Lou Henderson, Karen Worthy, and Bruce Crise. We thank

Dr Edward Sausville for ongoing support of this work. Supported in part by NCI Contract No NO1-CO-12400.

Andrew G Stephen¹, Alan Rein², Robert J Fisher¹, and Robert H Shoemaker³

¹Protein Chemistry Laboratory, SAIC-Frederick Inc

²HIV Drug Resistance Program, Center for Cancer Research

³Screening Technologies Branch, Developmental Therapeutics Program National Cancer Institute at Frederick Frederick

MD 21702

USA

Email: shoemaker@dtpax2.ncifcrf.gov

FURTHER READING

Amarasinghe GK *et al* (2000) **NMR structure of the HIV-1 nucleocapsid protein bound to the stem-loop SL2 of the Ψ -RNA packaging signal. Implications for genome recognition.** *Journal of Molecular Biology* **301**:491-511.

De Guzman RN *et al* (1998) **Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element.** *Science* **279**:384-388.

Fisher RJ *et al* (1998) **Sequence-specific binding of human immunodeficiency virus type 1 nucleocapsid protein to short oligonucleotides.** *Journal of Virology* **72**:1902-1909.

Sausville EA, Shoemaker RH (2000) **Role of the National Cancer Institute in AIDS-related drug discovery.** *Journal of the National Cancer Institute Monograph* **28**:55-57.

Stephen AG *et al* (2002) **Identification of HIV-1 nucleocapsid protein: nucleic acid antagonists with cellular anti-HIV activity.** *Biochemistry Biophysics Research Communication* **296**:1228-1237.

Weislow OS *et al* (1989) **New soluble-Formazan assay for HIV-1 cytopathic effects: Application to high-flux screening of synthetic and natural products for AIDS-antiviral activity.** *Journal of the National Cancer Institute* **81**:577-586.

Better drug molecules, discovered faster



From molecules designed to target GPCRs, through lead optimisation based on pharmacological knowledge, Peakdale offers you a drug discovery service to optimise your hits and reduce attrition rates in the clinic.

www.peakdale.co.uk