Validation of a novel secretion modification region (SMR) of HIV-1 Nef using cohort sequence analysis and molecular modeling

Patrick E. Campbell · Olexandr Isayev · Syed A. Ali · William W. Roth · Ming-Bo Huang · Michael D. Powell · Jerzy Leszczynski · Vincent C. Bond

Abstract The HIV-1 accessory protein Nef plays an active role in the pathogenesis of AIDS by its numerous cellular interactions that facilitate the release of virus particles. This 27 kDa protein is required for maintenance of the viral replication in HIV, and is also known to contribute to immune evasion, blocking of apoptosis in virus-infected cells and enhancement of virus infectivity. Nef has been shown to be secreted and is present on the surface of virus-infected cells. Recent studies from our laboratory have shown that the Nef protein is secreted from *nef*-transfected and HIV-1-infected cells in small exosome-like vesicles (40–100 nm diam.) that do not contain virions. We have identified three amino-terminal domains of Nef as necessary for secretion: (i) the four arginine residues (17,19,21, 22) comprising the basic region; (ii) the phosphofurin acidic cluster sequence (PACS) composed of four glutamic acid residues (61–64); (iii) a previously unknown motif spanning amino acid residues 65–69 (VGFPV) which we named the secretion modification region (SMR). In this study, we have used population-based phylogeny data and sequence analysis to characterize the conservation of the Nef SMR domain that regulates vesicle secretion. We have performed *in silico* computational chemistry analysis involving molecular dynamic structure modeling of mutations in the SMR motif. Sequence analysis of Nef from HIV-1-infected patients, including slow progressors (SP), long term progressors (LTP) and long term non-progressors (LTNP) demonstrated 99 % conservation of the Nef SMR motif. Computational analysis including modeling of wild-type HIV-1 Nef and V66A Nef SMR mutant using structural homology and molecular dynamics of ligand-associated interactions indicated significant structural changes in the Nef mutant, thus supporting the importance of the SMR domain for mediating Nef vesicle secretion.

Keywords HIV-1 protein · Homology modeling · Molecular dynamics · Structure-function relationship

Patrick E. Campbell and Olexandr Isayev contributed equally to this work

P. E. Campbell · S. A. Ali · W. W. Roth · M.-B. Huang ·
M. D. Powell · V. C. Bond (✉)
Department of Microbiology, Immunology and Biochemistry,
Morehouse School of Medicine,
Atlanta, GA, USA
e-mail: vbond@msm.edu

O. Isayev
Department of Chemistry, Case Western Reserve University,
Cleveland, OH 44106, USA

J. Leszczynski
Department of Chemistry, Jackson State University,
Jackson, MS 39217, USA

Present Address:
P. E. Campbell
University of the West Indies,
St. Augustine, Trinidad and Tobago

Present Address:
S. A. Ali
Universiti Sains Malaysia,
Pulau Pinang, Malaysia

Introduction

HIV-1 infection is characterized by a deterioration of host cellular immunity, measured by a steady loss of CD4$^+$ T-cells, eventually resulting in CD4$^+$ T-lymphocyte (T-cell) levels < 200/mm^3, an AIDS-defining condition. The association of the HIV-1 accessory protein, Nef, with HIV-1 disease progression has been thought to be due to its contribution to

increased viral load, with concurrent decrease in CD4$^+$T cell populations. An additional role of Nef in CD4$^+$ T-cell depletion at sites distant from the virus-infected cell may be associated with Nef release from these cells in vesicles, which cause pathological effects on uninfected T-cells. Studies in non-human primates showed that systemic CD4$^+$ T-cell loss resulting from direct viral infection is not sufficient to cause CD4+ T-cell depletion [1], suggesting that killing of "bystander" T-cells is central to HIV disease progression. Supporting evidence for the role of Nef in HIV-1 disease is highlighted by members of the Sydney Blood Bank Cohort SBBCS who were infected with a Nef-deleted strain of HIV, [2] by transgenic mouse studies showing AIDS-like pathologies when Nef is expressed [3, 4] and by rhesus macaques which exhibit no Simian AIDS (SAIDS) when infected with Nef$^-$-minus SIV, but progressing to SAIDS pathology upon the regeneration of *nef* .[5–7] The role of Nef in CD4$^+$ T-cell depletion has also been demonstrated with soluble Nef/Nef-antibody complexes which have been shown to be cytotoxic to T helper cells [8, 9] and apoptotic to lymphoid cells [10, 11].

Intracellular Nef protein binds to a number of host cellular factors and acts as an adaptor in generating protein interactions associated with cellular pathogenesis [12]. The secreted form of Nef [9] has been shown to be present on the surface of infected cells [13, 14], and clinical studies has shown its presence in the serum of infected patients [9]. Some of the biological activities that these secreted proteins possess include: a) the stimulation of HIV transcription in monocytic cells [15], b) the induction of apoptosis in non-infected cells [8, 9, 16, 17], c) the disruption of normal hematopoiesis in bone marrow progenitor cells [18], d) the activation of STAT1 and STAT3 in primary human monocytes/macrophages [19, 20], leading to interference with STAT-regulated intracellular signaling involving the release of chemokines and cytokines.

Nef has been shown to be released from transfected and infected cells, in a high molecular weight format in exosome-like vesicles [21]. This extracellular Nef causes apoptosis in bystander T-cells, mediated via the CXCR4 receptor. Nef can be taken up by CXCR4-minus cells by a yet to be determined mechanism [16, 21]. Earlier studies performed in our laboratory have shown that the secretion of Nef from T-cells is under the control of three separate motifs in the amino-terminal portion of the protein, which are necessary for secretion of Nef from HIV-infected cells. This includes a novel five-amino-acid domain, the SMR, or secretion modification region [22]. In this report, we have generated a representative three dimensional model of HIV-1 Nef, which was used to predict structural alterations in a HIV-1 Nef SMR-V66A mutant. This molecular dynamic modeling approach was used for binding-site analysis, specifically related to amino acid distances from the core of putative binding sites in or near the SMR motif. This approach yielded accurate results that corroborated existing *in vivo* methodologies.

Methods

Nef protein sequence alignments

Consensus Nef amino acid sequences for each HIV-1 clade (A through O) were determined by alignment of individual Nef variant sequences downloaded from the HIV Sequence Database [23], using using the algorithms in GENEious Pro 4.0.2 (Biomatters Ltd, Auckland, NZ). Specifically, alignments were generated using a Blosum62 Cost Matrix, with a gap opening penalty = 12 and gap extension penalty = 3. The 13 HIV-1 clade consensus sequences thus determined were then submitted for alignment in Geneious Pro 4.02, using the same parameters. The Nef sequences from patient cohorts [24–27] were also aligned in Geneious Pro using the same algorithm.

Generation of three dimensional (3D) models

The Nef amino acid sequence used for structural analysis was from the Clade B variant NL4-3, (PubMed ID CAA41585). This was used to generate RCSB Protein Database (PDB) structures which were used as templates for generation of 3D models: (A) PDB ID 1QA5 – the Nef anchor domain: residues 2 through 57, NMR derived, and (B) PDB ID 2NEF – a nearly full-length Nef protein, missing sequences 2–39, and 159–173). Homology model generation was done using Geno3D [28]. Subsequent MD simulations were carried out using the AMBER molecular modeling program suite (version 8.0) [29] using the GENO3D homology generated model.

Molecular dynamics simulations

For all of our simulations, AMBER FF99SB force field was utilized [30–32]. Water was used as the surrounding environment for the protein MD simulations. The Nef protein was neutralized with Na$^+$ prior to simulation. For the explicit solvent simulations, the system was solvated with the 3-site TIP3P water molecules in a rectangular box, with a 10 Å buffer between the Nef protein and the edges of the box. During the course of the simulations, the potential energy (PE) and other conserved quantities were closely monitored to ensure proper equilibration.

Prior to the MD simulation, the protein underwent minimization in two phases: phase one, where water was minimized and phase two, where water, protein and neutralizing Na$^+$ underwent minimization, both sets using Cartesian restraints for the solute. The system was equilibrated in two stages, bringing the temperature up from a low of 100 K to 300 K over 100 ps, with a time step of 2 fs, using Berendsen coupling algorithm to maintain constant temperature. The time constant used for temperature coupling was 2 ps and a periodic boundary with constant volume and no

pressure control was employed, together with a SHAKE algorithm to constrain only bonds involving hydrogen. Once equilibration was achieved the pressure was held constant (1 atm) with isotropic position scaling until the water density equilibrated to 1 g/ml. Molecular dynamic simulation was carried out for periods up to 50 ns. Backbone integrity was judged with RMSD analysis. Vizualization of the PDB images and manipulation of the data to create varying representations of the protein data was done using CHIMERA from UCSF [33].

Ramachandran plot analysis

The alpha carbon (Cα) is the most important locus for the evaluating distortion of covalent geometry in protein structures, as it joins side chain with backbone and responds to both, especially regarding their compatibility. For Ramachandran analysis of backbone suitability [34], RAMPAGE uses density-dependent smoothing for 81,234 non-Gly, non-Pro and non-Pre-Pro areas and regions that are allowed but disfavored. Variance measurements on the basis of a manually curated set of high-quality protein structures and a number of filters (such as B-factor cutoff and van der Waals clashes) is used to develop reference phi/psi plots for Gly, Pro and other general residue types. These selected amino acid residues are divided into 'favored', 'allowed' and 'outlier' regions, with a corresponding picture of the plot generated by RAMPAGE [35].

Results and discussion

Amino acid variance in the SMR

The SMR, or secretion modification region of HIV-1 Nef, consisting of amino acid residues VGFPV, (residues 66–70, see Fig. 1). This is one of three sequence motifs we identified as necessary for secretion of HIV-1Nef protein from Nef-expressing cells [22]. If this motif functions in HIV-1 pathogenesis as we have proposed, the amino acid sequence should be highly conserved across all groups of HIV-1. Therefore, we aligned the relevant Nef sequences from all ten HIV-1 subtypes, or clades. The alignments were generated from a pool of 1600 representative sequences from the AIDS Sequence Database [23]. The majority of entries are from subtypes A, B, C and the circulating recombinant forms (CRF). Table 1 shows the variance of the SMR Domain in the HIV-1 Group M reference sequences from the HIV Sequence Database compendium [23]. These reference sequences represent a survey of HIV-1 Nef variation among different subtypes. The reference sequences from all ten subtypes of HIV-1 were aligned as described in Methods. Analysis of variation was performed for the five amino acids of the SMR (VGFPV)

within each subtype. Subtypes A, B, C, and the circulating recombinant forms (CRF) account for more than 90 % of the reference sequences in the database. The variation in the SMR amino acids among these sequences is 3 %. Variation in these reference sequences reflects the approximate level of variation one might observe in HIV-positive individuals.

### Population studies on selected cohorts

A better way to survey HIV-1 Nef variation should be to examine sequences in HIV-1-positive patients. The alignment of HIV-1 Nef sequences from HIV-1 cohorts from the USA, Italy, Australia and Japan [30–33] is shown in Fig. 2. Collectively, this represents 347 HIV-1 infected individuals classed as rapid progressors, long term progressors or long-term non-progressors, with an age range of 4 – 65 years. Of these representative sequences, there were only three examples of amino acid substitutions in the SMR, with changes from Val-66 to Glycine (Italian Cohort), Pro-69 to Serine (US cohort) and Val-70 to Glycine (US cohort). Figure 2 shows the consensus for each cohort is VGFPV. This represents a 99.14 % SMR conservation across the 347 patient samples, thus indicating the likelihood that an individual HIV-1 Nef isolate has the VGFPV motif is greater than 99 %, and affirming that the SMR is highly conserved among HIV-1 patients. This evidence indicates the relevance of these domains, particularly in a virus that displays high sequence variability. The SMR conservation will be the subject of our future works.

Three dimensional modeling of Nef

The RMSD values of variance between backbone Cα coordinates of a reference compared to all other structures recorded over the duration of the MD. For our purposes, the RMSD reference geometries chosen corresponded to the lowest energy structure conformation obtained from 50 ns MD for both wild-type (WT) and V66A (MT) HIV-1 Nef simulations. The WT standard was obtained after 48.24 ns of MD, while the mutant standard was revealed after 11.58 ns of MD. Examination of the RMSD values (Fig. 3a) for the WT indicates an area of stability after ~20 ns averaged through the last half of the trajectory.

The mutant RMSD examination (Fig. 3a) reveals a more erratic pattern over the course of the 50 ns MD. The structure of lowest energy is attained after 11.58 ns, oscillates over the next 13 ns from a low of 2.4 Å to a high of 7 Å. After this period, oscillation of RMSD values continues over the next 24 ns MD, with a high of 8.2 Å, to a low of 5 Å after 45 ns MD. This comparatively erratic pattern in MD for the WT was achieved by the substitution of a single amino acid in the SMR. Overall, if calculated per residue RMSF and averaged over the whole trajectory, outer loops and random coil regions exhibit high mobility and significant flexibility (see loops colored in blue on Fig. 4). In

Fig. 1 The structure and amino acid sequence of HIV-1 Nef. **a** Three-dimensional model, which shows a subset of relevant, previously reported structural as well as functional motifs. **b** Sequence of HIV-1 Nef showing structural domains required for cellular interactions. The basic amino acid 1 and 2 motifs (BAA-1, BAA-2), helix-1 and helix-2, membrane targeting domain, PACS, and SMR motifs were investigated in the previous study (22)
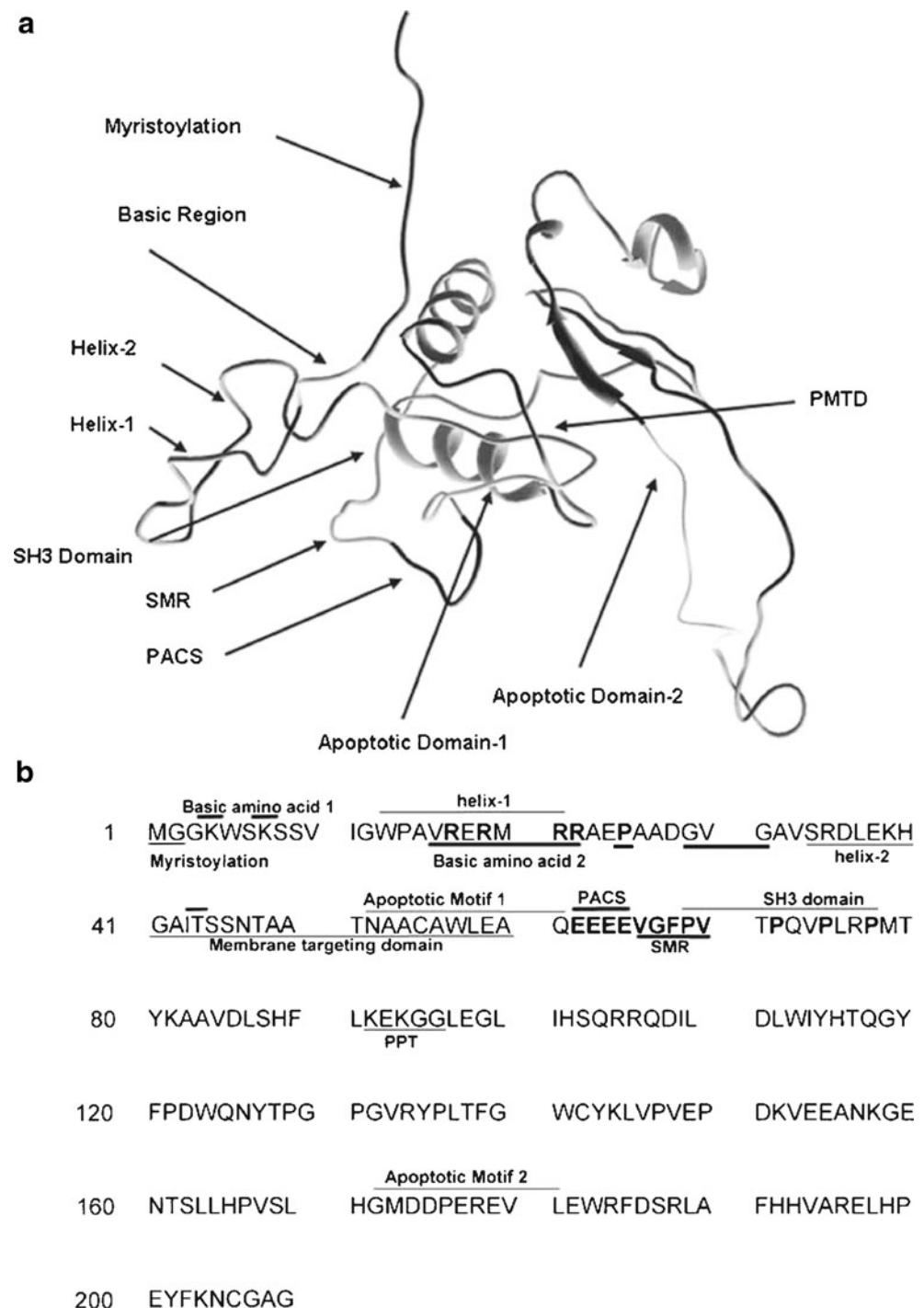
contrast, core regions are much more rigid with RMSD smaller than 3-4 Å (see red residues, Fig. 4).

The Ramachandran plot analysis shows the favored location of amino acids comprising the SMR motif in both WT and V66A mutant (MT) forms. The WT 0 ns and 50 ns MD simulation analysis shows all the SMR residues are located in the 'favored position' (Fig. 3b). For the V66A mutant at 0 ns MD, amino acids in positions one and four of the SMR motif (Ala-66 and Pro-69) have been moved from the favored

position to the allowed position. At 30 ns MD, the mutant SMR has amino acids in positions one and three (Ala- 66 and Gly-67) placed in the allowed region. Overall, these results indicate that at 0 ns and 30 ns MD analysis points, the WT has all of the SMR amino acid residues in the favored region. As for the V66A mutant, only three out of the five residues are located in this area. It is also important to note that there is variability in the V66A amino acid movement durjng the couse of MD: Pro-69 and Ala-66 for 0 ns MD, and Ala-66 and Gly-67 after 30 ns

Table 1 SMR amino acid variation among HIV-1 subtype sequences (23)

Subtype (number entries)	A (111)	B (600)	C (384)	CRF (363)	D (76)	F (12)	G(19)	H (8)	J (2)	K (2)
Val-66[a]	1	3	2	0	0	0	0	0	0	0
Gly-67	1	4	1	1	0	0	0	0	0	0
Phe-68	1	2	2	3	0	0	0	0	0	0
Pro-69	0	2	4	1	0	0	0	0	0	0
Val-70	1	7	5	3	0	0	0	1	0	0

[a] Number of variants at each position in the SMR

MD. In both instances, although Ala-66 is located in the allowed region, it is never found in the favored region.

We also performed RMSD analysis using the homology models (WT and MT) as the reference for each respective MD simulation (Fig. 3c). Higher results indicate that some of the predicted regions were poorly predicted by the homology model, especially Apoptotic domain, Myristoylation region, helices 1 and 2. Therefore, the majority of RMSD difference and RMSF fluctuations originates from compaction of these regions and transition from random coil-like structures to helix-like ones.

For both WT and V66A MT Nef variants, the location of the SMR domain is favorably placed on the outer surface of the 3D structure, ideally unencumbered and available to interact with either solvent or other protein structures. For the WT (Table 2), the location of the SMR is brought into closer contact with the rest of the protein as seen with the progression of the MD time from 0 – 50 ns. The distance measurement from the N-terminal amino acid to the first amino acid of the SMR (Val-65) shows a decrease of 41.5 Å from 52.5 Å at 0 ns, through to 11.0 Å at 50.0 ns MD (Table 2). Additionally, the distance of the amino acid located at the bottom of the WT Nef 'arm', Gly-41 , to the C terminus protein N_{204} reveals an increase in length of 3 Å – a minimal change recorded over the 50 ns MD.

The evidence for a more compact structure is reinforced by the data which indicates an increase in the number of helices for the WT Nef from seven to 13 after 50 ns MD (Table 2). These observations all support the notion of increased thermodynamic stability as indicated by the significant decrease in the total PE of the structure during the simulation. The magnitude of this difference, approximately 850 kcal mol^{-1} after 48.24 ns MD.

For the V66A mutant, with the SMR sequence AGFPV, the distance measurement from the N-terminal amino acid residue to the SMR residue Ala-66 shows a decrease of 27.4 Å from 54.6 Å at 0 ns MD to 27.2 Å after 30 ns MD (Table 3). The measure from the bottom of the (MT) Nef 'arm' at residue Gly-40, to the C terminal residue N_{204} reveals a change of only 0.6 Å after the MD run, indicating little variation in core positioning over the MD.

As was observed for the WT, a more thermodynamically stable structure is attained for V66A mutant after the MD,
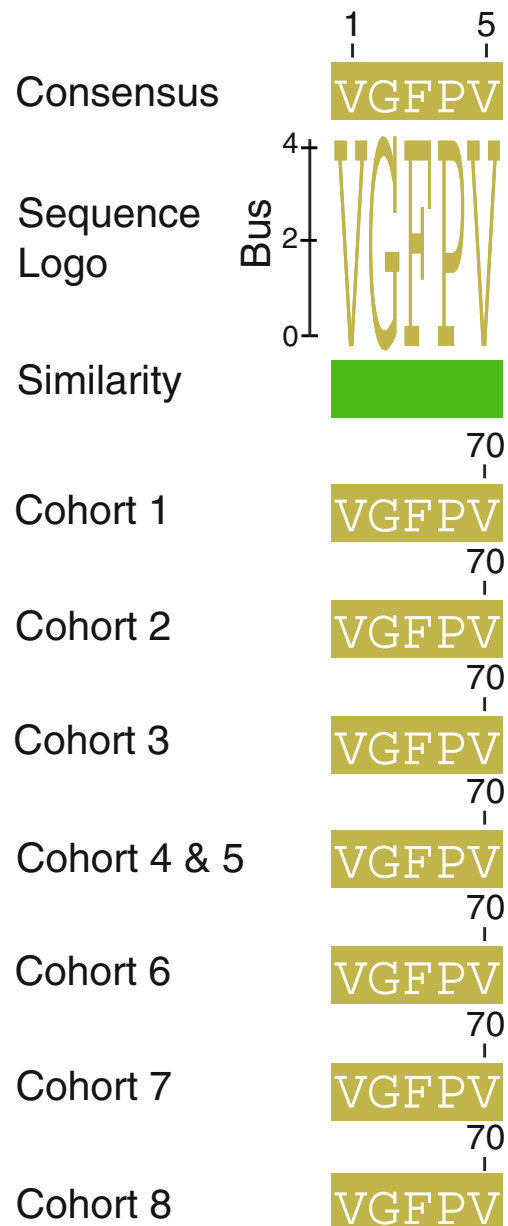

Fig. 2 Alignment of SMR domain in eight HIV-1 patient cohorts. A global alignment with free end gap variants was performed as described in Methods. The consensus SMR amino acid sequence from each cohort was then aligned using the same parameters
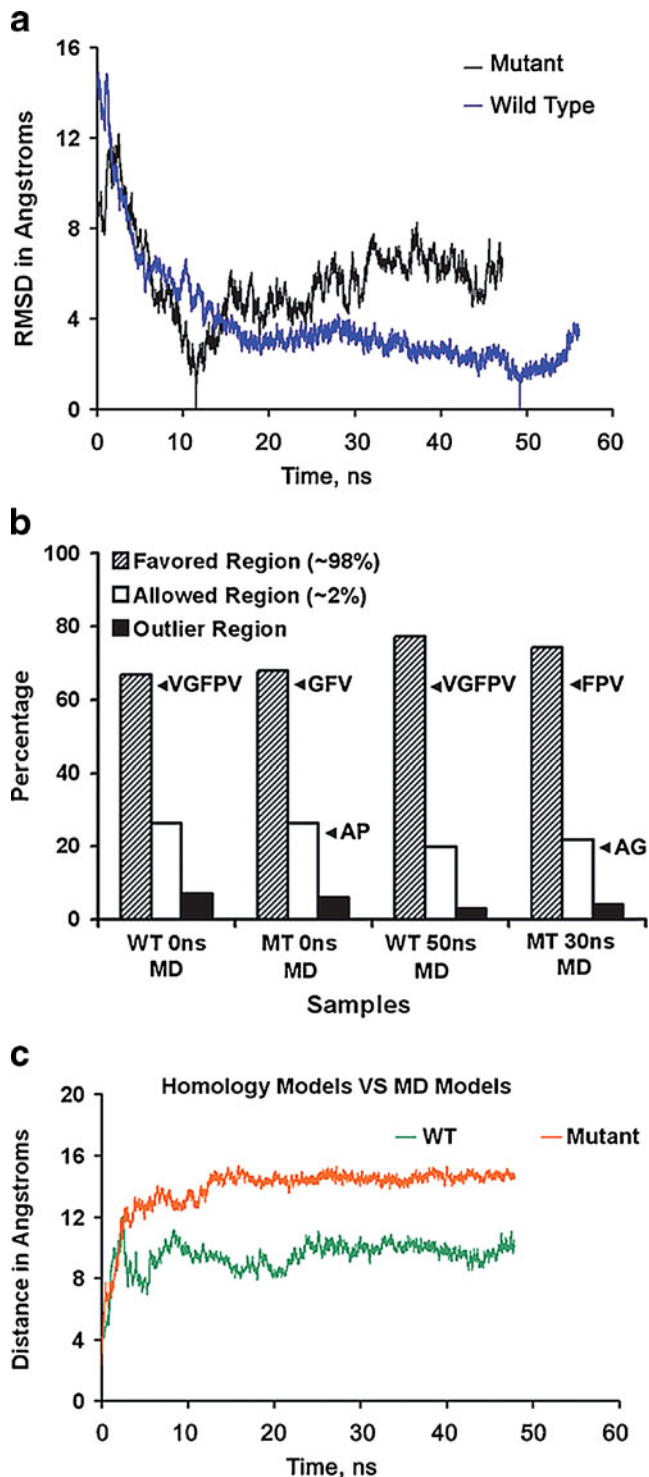
Fig. 3 **a** Root mean square deviation (RMSD) comparison of wild type and mutant HIV-1 Nef. **b** Ramachandran plot analysis of WT and mutant HIV-1 Nef showing region location of amino acids in SMR. **c** RMSD comparison of WT and MT HIV-1 Nef using the respective homology model as the standard for comparison

with an increase of five helices (three to eight) after simulation completion. Comparison of homology model to the structure obtained after 50 ns MD indicates an overall

increase in the thermodynamic stability of the structure. An increase in helix number from seven to 13 over the 50 ns of MD supports this fact.

Examination of putative binding sites

The amino acids comprising the SMR domain are believed to play a role in vesicle secretion. For this reason, we examined their location in putative binding sites on the 3D model of Nef, both in the WT and the MT structures. The putative binding sites, representing favorable areas on the surface of the HIV-1 Nef for protein-protein interactions, were chosen by the criteria of surface exposure for binding, electrostatic favorability for binding and the location of conserved domains. Preliminary results indicated a consistently favorable score for one of the amino acid residues outside of the SMR, T-44 (Table 4). The electrostatic potential surface and the lipophilicity values for our model of the HIV-1 Nef are shown in Fig. 5. The SMR residues are located in a region of medium polarity. In contrast, the T_{43} residue is located at the interface of the highly polar pocket on the back site of PACS sequence. Other polar areas are located on the opposite side of the protein, especially the N- and C-terminal residues.

The unique behavior of Thr-44 was recorded together with the SMR domain residues for the homology model (Table 4). Subsequent MD was performed and the pocket data measurements were now recorded for the SMR amino acid residues and the Thr-44 residue (Fig. 7a). For the wild type Nef SMR, five putative binding pockets were generated by the software MOLEGRO, of which only one was in proximity to the amino acid residues of the SMR as calculated by distance measurements taken from the calculated core of the pocket. The results which are displayed as a function of the measured distance from the putative pocket center, show that for the original homology model, Pro-69 and Val-66 are the closest to the center (8 Å) followed by Gly-67 (9 Å) and Thr-44 , Phe-68 and Val-70 , all at 12 Å.

Of the following MD snapshots taken at 10 ns, 30 ns and 50 ns, Val-66 is always closest to the center of the putative pockets on the 3D image generated for each time phase. The SMR motif amino acid residue Gly-67 is always second to Val66 at times 10, 30 and 50 ns. One of the other SMR residues, Phe-68 , (tied for fourth out of five in the ranking of distance from the center of the binding site at time t=0 ns) is third in distance measured from the core of the pocket at t=10 ns, but is beyond the cutoff point of 20 Å in both measure times of t=30 ns and 50 ns respectively. Pro-69, which was closest to the core of the pocket at time t=0 ns, is fourth out of five in the distance ranking at time t=10 ns, further than 20 Å at t=30 ns, and furthest away from the three qualifying aa for time t=50 ns. On the other hand, Val-70 is furthest away from the core at time t=0 ns and 10 ns,

Fig. 4 Structures of WT and mutant NEF HIV-1. Residue coloring by the backbone RMSD (Å) was averaged over the whole MD trajectory
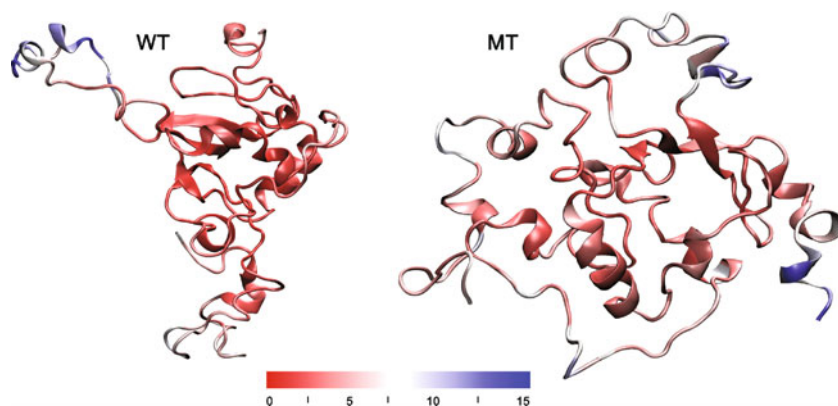
while its distance at times t=30 ns and t=50 ns is beyond the 20 Å cutoff point.

These results indicate the importance of Val-66 and Gly-67 in the SMR domain, as they were present in each of the four time points of MD sampling (time t=0 ns, 10 ns, 30 ns, 50 ns). With the one exception of Gly-67 at t=10 ns, these two SMR amino acid residues consistently ranked first and second in proximity to the core of the putative proximally located SMR pocket. Of the two, Val-66 would be considered the more important due to its consistent location closer to the center of the putative pocket in all time frames under examination. While Pro-69 ranked consistently lower compared with Phe-68, the fact that it appeared closer to the center of the pocket at time t=50 ns (which represents a more stable thermodynamic structure than at time t=30 ns), would give it a rank just ahead of Phe-68. Residue Val-70 would be the next in importance relative to its distance location from the putative binding pocket. Based on these data obtained from the binding pocket measurements, a listing of SMR residues from highest to lowest importance would thus be Val-66, Gly-67, Pro-69, Phe-68 and Val-70. As a result of its closer distance to the putative core at increasing thermodynamically-stable times of 30 ns and 50 ns, Thr-44 would rank ahead of Val-66 in importance.

For the V66A mutant SMR, five putative pockets were also generated by MOLEGRO, of which only one was in a correct proximity for the amino acids of the SMR (Fig. 6b). The results displayed as a function of distance from the center of the putative pocket, show that for the initial homology model, Ala-66 is the second closest to the center

(11 Å) after Trp-44 (6 Å). This is then followed by Val-70 and Gly-67 (12 Å), Pro-69 (13 Å) and Phe-68 (16 Å). The MD snapshots taken at the thermodynamically most favorable time of 30 ns indicate that the order of importance (high to low) for the amino acids of the SMR were Gly-67, Ala-66, with Phe-68 and Pro-69 tied for third. Val-70 was further than the 20 Å cutoff point at time t=30 ns and did not record a score. It is important to note however, that Trp-44 was not only the closest to the pocket core after 30 ns of MD, but was present for all of the simulation check points of times t=0, 10 and 30 ns. The ranking by importance therefore, of all the V66A mutant SMR amino acid residues (highest to lowest) would thus be Gly-67, Ala-66, Phe-68, Pro-69 and Val-70. Significantly, the Trp-44 residue outranked each of the SMR residues at all the recorded time intervals.

The 3D structures obtained after 50 ns of MD for both WT and mutant SMR motifs, proved to be accurate in predicting certain phenomena associated with the Nef protein. The structural components (Fig. 7a and b) indicated the favorable locations of the SMR motif for possible interactions with binding sites located on the surface of the protein. A more thermodynamically favorable structure was obtained after the MD simulations, as indicated from potential energy (PE) measurements, and increases in both helix number and selected distance measurements when compared to the original homology model, (t=0 ns MD). The correctness of our hypothesis is justified by the conformational stability of the Nef during MD (Fig. 3a and c), and corroborated by Ramachandran plot analysis that placed the critical aa of the SMR into the expected regions of the plot. The RMSD indicated greater

Table 2 Structural and energetic variations of wild type HIV-1 Nef during MD

		WT no MD	WT 10.00 ns MD	WT 31.14 ns MD	WT 50.00 ns MD
1	Lowest PE structure, NS	0.00	0.00	0.00	48.24
2	Distance, G1-V65,	52.5	16.8	11.5	11.0
3	Distance, G40-N204,	39.3	40.0	42.2	42.3
4	Helices (number)	7	10	13	13
5	Total PE (kcal/mol)	−172.704.86	−173.466.74	−173.467.52	−173.541.33

Table 3 Structural and energetic variations of HIV-1 Nef V66A mutant during MD

		MT no MD	MT 10 ns MD	MT 26.3 ns MD
1	Lowest PE structure	0.00	0.00	11.58
2	Distance, G1-A65,	54.6	59.6	27.2
3	Distance, G40-N204,	39.4	35.4	38.8
4	Helices, (number)	3	8	8
5	Total PE (kcal/mol)	−203,283.97	−204,161.17	−204,254.08

stability of the WT model over its mutant counterpart, while the distance measurements (Tables 2 and 3), solidified the WT amino acids Val-66 and Gly-67 as the two most important amino acid residues in the putative binding pocket containing the SMR motif.

Conclusions

The HIV-1 Nef protein has been shown to interact with host proteins in order to regulate several host cell functions. Our laboratory and others have shown that Nef is secreted from the human host cells, a process which is mediated by at least three motifs [16, 22, 36]. One of these motifs is the novel SMR, a five-amino acid domain that is absolutely required for Nef protein secretion [22]. The SMR domain was identified by using a series of Nef genetic mutations in functional assays that measure Nef secretion from bacteria and human cells.

We hypothesized that comparative computer modeling analysis of the conserved SMR motif of HIV-1 Nef in both wild-type and V66A mutant proteins would yield valuable information concerning the effect of SMR mutations upon Nef protein structure; reasoning that, since the Nef SMR motif is highly conserved among HIV-1 isolates, its sequence might be critical for the structural integrity and function of the HIV-1 Nef protein.

To address the issue of SMR conservation, we performed alignments of HIV-1 reference sequences from the HIV Sequence Database at Los Alamos National Laboratory [23]. These alignments showed a high degree of conservation (approx. 97 %) in the SMR motif across HIV-1 subtypes. This motif is one of several well-conserved motifs in the Nef protein, nearly all of which have been proposed linked to interaction with host cell proteins.

If the SMR serves a function in the host cell, it should also be conserved in isolates from HIV-1 infected individuals.

Therefore we examined Nef sequences in 347 patient samples taken from: 1) cohorts from globally separated areas (US, Japan, Italy, Australia & Sweden) 2) cohorts of varying degree of disease progression (HIV-1 slow progressors, rapid progressors, and long-term non-progressors, or elite controllers). The comparison of these patient cohort sequences indicated approximately 99 % conservation of the SMR motif, meaning the VGFPV sequence was present in 99 % of isolates, and provides compelling evidence for the importance of the SMR in the viral life cycle and pathogenesis.

Having evidence that the SMR is one of the most conserved motifs in HIV-1 Nef, we next turned our attention to the structural characteristics of the SMR, as examined using computer modeling and molecular dynamic simulation. For this purpose, we generated a full length Nef model using Geno3D modeling software, which utilized the Nef sequence (ID CAA41585) to: (i) identify homologous proteins with known 3D structures by using PSI-BLAST; (ii) provide all potential templates through a convenient user interface for target selection; (iii) perform the alignment of both query and subject sequences; (iv) extract geometrical restraints (dihedral angles and distances) for corresponding atoms between the query and the template; (v) perform the 3D construction of the protein by using a distance geometry approach.

This three-dimensional model was used to examine putative protein binding sites in Nef, since these conserved residues are most likely to be found at important functional sites. To further identify the SMR as a potential site for interaction with binding partners, the shape and solvent accessibility of the SMR motif was examined to differentiate potential binding sites from the remainder of the protein surface. Additionally, the interface with polar residues and hydrophobic/hydrophilic clusters would also be revealed by this 3D examination of the SMR relative to the rest of the Nef protein structure. Our results indicated that the SMR motif was in an electrostatically and morphologically

Table 4 Distances of SMR domain residues from center of binding pocket in the homology model of HIV-1 Nef

	T-44	E-63	E-64	E-65	V-66	G=67	F-68	P-69	V-70
VGFPV	6	13	10	9	5	7	11	8	12
VGYPV	5	15	10	10	7	8	12	9	13
LGYPV	6	11	7	7	5	9	13	12	16
T-43 → A-43	7	13	9	7	5	6	10	9	12

Fig. 5 Electrostatic potential and lipophilicity mapped on VDW HIV-1 Nef surface. Values for electrostatic potential (left panel) and lipophilicity (right) of individual amino acid residues are shown, with higher value residues in red and low values in blue. Regions of interest (PACS, SMR and the Trp-43 residue) are marked
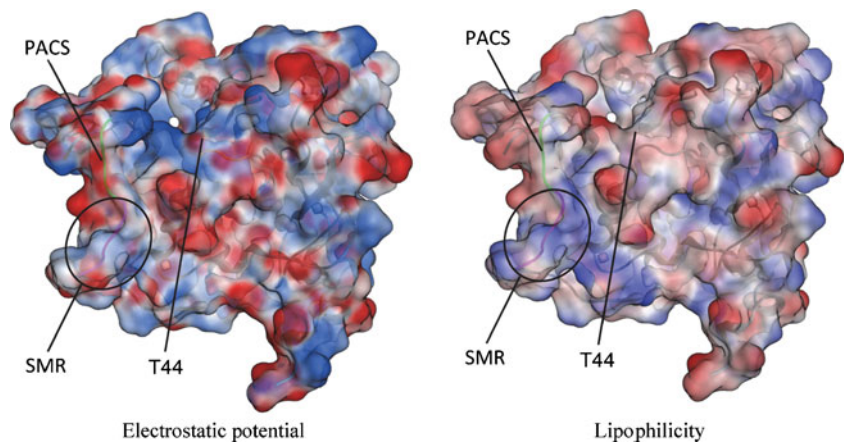
favorable position to act as a functional binding site. Further quantification of the surrounding Nef amino acids in closest proximity to the core of the SMR domain revealed that Thr-44 and Val-66 were consistently closest to the pocket core, even in 3D-generated Nef mutants. We interpreted these results to indicate the importance of these two amino acid residues for the proper functioning of the SMR binding site. We had already created HIV-1 Nef SMR mutant constructs and tested

them in our secretion assay [22] and this new *in silico* data confirmed data already generated by secretion assays of those Nef mutants. However, the information on the potential
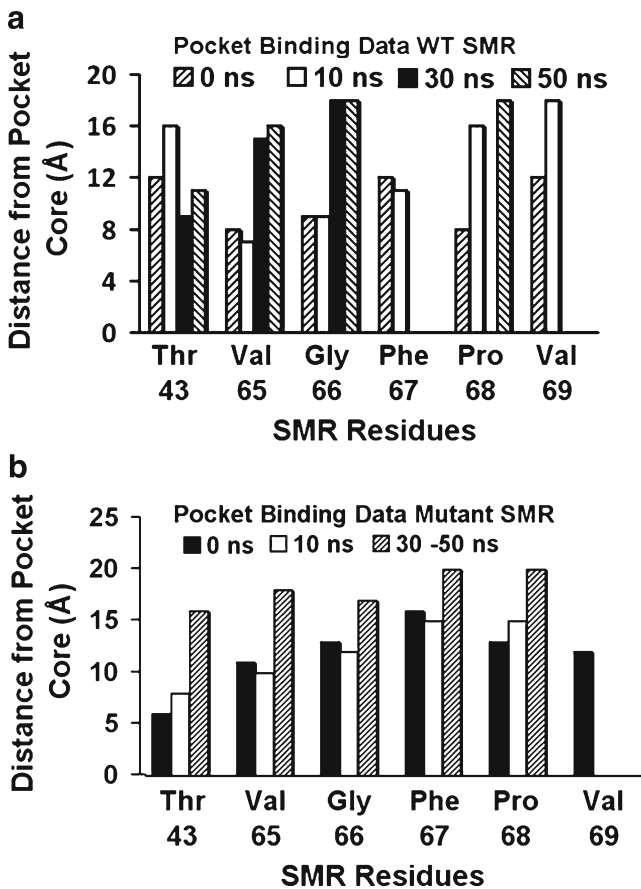
Fig. 6 a Distance measurements for key amino acids in a putative binding pocket of WT HIV-1 Nef **b** Distance measurements for key amino acids in a putative binding pocket of HIV-1 Nef SMR domain mutant
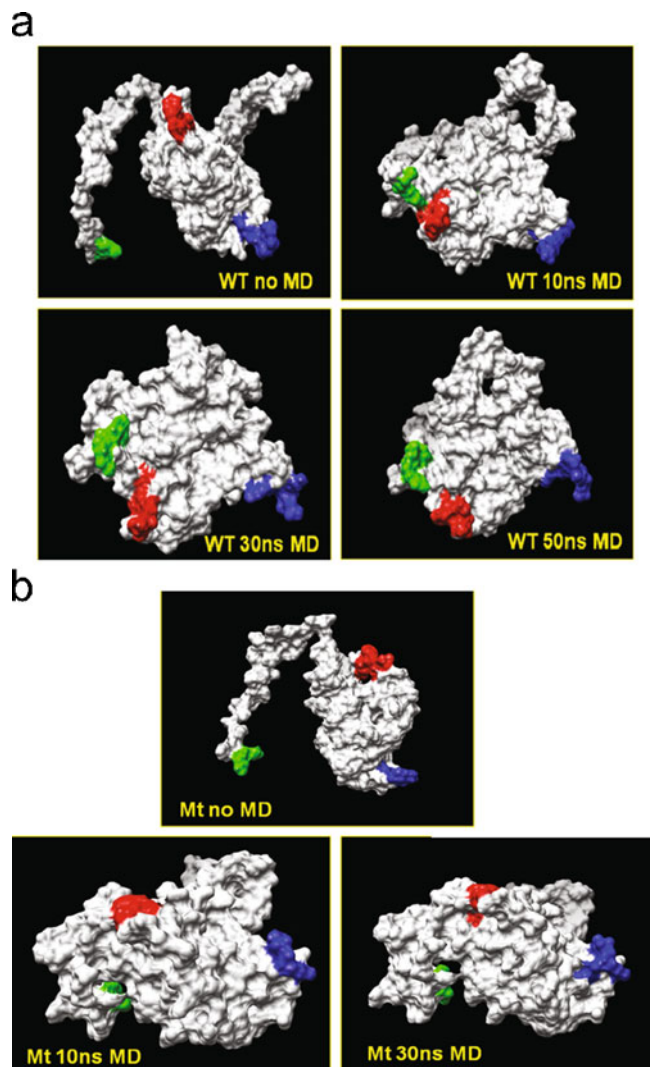
Fig. 7 Spatial positions of important Nef regions at 0, 10, 30 and 50 ns. **a** WT HIV-1 Nef. **b** MT HIV-1 Nef

importance of Thr-44 led us to create a Nef clone with a T44A mutation and test it in the secretion assay [22]. Interestingly, this mutation also disrupted the ability of Nef to induce its own secretion. This is particularly encouraging, since we have not only confirmed a specific amino acid residue within the SMR (Val-66) as necessary for the proper structure of HIV-1 Nef and for Nef secretion: but we also identified an amino acid residue beyond the SMR motif (Thr-44), that is also important for Nef secretion from Jurkat T-lymphocytes.

To enhance the accuracy of our computer-generated 3D model, we performed molecular dynamic (MD) optimization on the structure followed by a repeat of the distance measure to the core of the SMR, in order to validate the results obtained pre-MD (see Table 4).

Next, to verify our MD models, we also analyzed them by random coil comparisons (RCC), distance measurements, root mean square deviation (RMSD) and Ramachandran analysis (RA). The RCC of the backbone structure of the MD models with those that were randomly generated solely using the Nef sequence gave the expected non-alignment with the structures. Both measurements of PE at selected distances between chosen Nef amino acids and changes in helix numbers indicated that the 50 ns MD model was more energetically stable than the initial pre-MD structure. These *in silico* observations were further supported by the RMSD analysis which showed that the WT SMR (VGFPV) MD-generated structure fell within accepted limits (less than 4 Å) while the mutant Nef V66A SMR structure was outside this accepted range, as compared to the lowest energy structure generated by MD experiments. Ramachandran plot analysis revealed that for the allowable areas (psi vs. phi plot) for the amino acids comprising the SMR (plus Thr-44), our WT MD SMR had a higher proportion of amino acids within the favored region, while the mutant (Val-65 to Ala substitution) had a greater number of amino acid residues outside the accepted range. The significance of this observation lies in the fact that a single Val-to-Ala substitution at the first position of the SMR produced a significant change in the entire conformation of the SMR and therefore, of the Nef protein. After the MD was completed, the pocket measurement data was repeated and the data supported the results that were obtained before MD. Thus the MD was able to accurately simulate the forces acting on the protein (in water), and to create a more energetically favorable structure compared with the homology model as verified, and to confirm the identification of Thr-44 and Val-66 obtained from the homology model.

Overall, the high degree of conservation of the SMR domain across HIV-1 quasispecies variants does indicate its significance as an evolutionarily conserved domain with functional activity. The results obtained *in silico* for the 3D models for both the WT and MT HIV-1 Nef indicate their suitability for use as analytical tools to facilitate the design of subsequent *in vivo* studies for mutational analysis of HIV-

1 Nef. The fact that the homology and MD models gave similar results is quite encouraging for future research studies in this area. More analysis of the functional significance of other Nef motifs involved in interactions with host proteins is warranted.

Our *in silico* data also showed the success of the 3D model in predicting the behavior of the SMR domain in MD over a time interval of 50 ns. It also demonstrated the usefulness of the homology model and putative binding pocket generated by the model by accurately predicting amino acids that could be chosen for *in vivo* mutational analysis. More remarkable is that the homology and MD models could be used for identification of an amino acid residue distant from previously identified ones with a significant effect on the functional outcome in the Nef secretion assay.

As discussed earlier, the primary purpose of this study was to create a model for HIV-1 Nef that would help us to examine how the SMR might interact with cellular proteins. The next task will be the identification of protein component(s) that interact with the SMR binding pocket, causing Nef to be packaged in vesicles. Such proteins would presumably reside in the multivesicular body or other parts of the secretory pathway. Once SMR-binding proteins are identified, they could also become candidates for similar modeling or MD analysis.

Following its secretion from HIV-1-infected cells, Nef exerts effects on non-infected bystander T-lymphocytes which include disruption of normal cellular growth and metabolism leading to apoptosis, or programmed cell death. This process may account for a portion of the loss in CD4+ T-lymphocytes during HIV-1 infection, a key component of HIV-1 pathology. It therefore follows that an inhibitor of Nef secretion might decrease the rate of T-lymphocyte depletion in HIV-1-infected individuals. Thus, an accurate picture of the portion of HIV-1 Nef surrounding the SMR motif, coupled with the principles of rational drug design based on use of computer-developed structural models could lead to development of a new class of antiviral therapeutic agents.

References

1. Smith SM (2006) The pathogenesis of HIV infection: stupid may not be so dumb after all. Retrovirol 3:60–65

2. McIntyre LB, Geczy AF, Dyer WB, Learmont JC, Sullivan JS (1999) The Sydney Blood Bank Cohort: a case–control study using a transfused HIV-1 seronegative group. Ann Epidemiol 9:436–440

3. Hanna Z, Priceputu E, Hu C, Vincent P, Jolicoeur P (2006) HIV-1 Nef mutations abrogating downregulation of CD4 affect other Nef functions and show reduced pathogenicity in transgenic mice. Virol 346:40–52. doi:10.1016/j.virol.2005.10.010

4. Priceputu E, Rodrigue I, Chrobak P, Poudrier J, Mak TW, Hanna Z, Hu C, Kay DG, Jolicoeur P (2005) The Nef-mediated AIDS-like disease of CD4C/human immunodeficiency virus transgenic mice is associated with increased Fas/FasL expression on T cells and T-cell death but is not prevented in Fas-, FasL-, tumor necrosis factor receptor 1-, or interleukin-1beta-converting enzyme-deficient or Bcl2-expressing transgenic mice. J Virol 79:6377–6391

5. Simard MC, Chrobak P, Kay DG, Hanna Z, Jothy S, Jolicoeur P (2002) Expression of simian immunodeficiency virus nef in immune cells of transgenic mice leads to a severe AIDS-like disease. J Virol 76:3981–3995

6. Hanna Z, Kay DG, Rebai N, Guimond A, Jothy S, Jolicoeur P (1998) Nef harbors a major determinant of pathogenicity for an AIDS-like disease induced by HIV-1 in transgenic mice. Cell 95:163–175

7. Klotman PE, Notkins AL (1996) Transgenic models of human immunodeficiency virus type-1. Curr Top Microbiol Immunol 206:197–222

8. Fujii Y, Otake K, Tashiro M, Adachi A (1996) Human immunodeficiency virus type 1 Nef protein on the cell surface is cytocidal for human CD4+ T cells. FEBS Lett 393:105–108

9. Fujii Y, Otake K, Tashiro M, Adachi A (1996) Soluble Nef antigen of HIV-1 is cytotoxic for human CD4+ T cells. FEBS Lett 393:93–96

10. Okada H, Takei R, Tashiro M (1997) Nef protein of HIV-1 induces apoptotic cytolysis of murine lymphoid cells independently of CD95 (Fas) and its suppression by serine/threonine protein kinase inhibitors. FEBS Lett 417:61–64

11. Okada H, Takei R, Tashiro M (1997) HIV-1 Nef protein-induced apoptotic cytolysis of a broad spectrum of uninfected human blood cells independently of CD95(Fas). FEBS Lett 414:603–606

12. Roeth JF, Collins KL (2006) Human immunodeficiency virus type 1 Nef: adapting to intracellular trafficking pathways. Microbiol Mol Biol Rev 70:548–563

13. Fujii Y, Otake K, Fujita Y, Yamamoto N, Nagai Y, Tashiro M, Adachi A (1996) Clustered localization of oligomeric Nef protein of human immunodeficiency virus type 1 on the cell surface. FEBS Lett 395:257–261

14. Zeigler ZR, Rosenfeld CS, Andrews DF III, Nemunaitis J, Raymond JM, Shadduck RK, Kramer RE, Gryn JF, Rintels PB, Besa EC, George JN (1996) Plasma von Willebrand Factor Antigen (vWF: AG) and thrombomodulin (TM) levels in Adult Thrombotic Thrombocytopenic Purpura/Hemolytic Uremic Syndromes (TTP/HUS) and bone marrow transplant-associated thrombotic microangiopathy (BMT-TM). Am J Hematol 53:213–220

15. Varin A, Manna SK, Quivy V, Decrion AZ, Van Lint C, Herbein G, Aggarwal BB (2003) Exogenous Nef protein activates NF-kappa B, AP-1, and c-Jun N-terminal kinase and stimulates HIV transcription in promonocytic cells. Role in AIDS pathogenesis. J Biol Chem 278:2219–2227

16. Huang MB, Jin LL, James CO, Khan M, Powell MD, Bond VC (2004) Characterization of Nef-CXCR4 interactions important for apoptosis induction. J Virol 78:11084–11096

17. James CO, Huang M-B, Khan M, Garcia-Barrio M, Powell MD, Bond VC (2004) Extracellular Nef protein targets CD4+ T cells for apoptosis by interacting with CXCR4 surface receptors. J Virol 78:3099–3109

18. Calenda V, Graber P, Delamarter JF, Chermann JC (1994) Involvement of HIV nef protein in abnormal hematopoiesis in AIDS: in vitro study on bone marrow progenitor cells. Eur J Haematol 52:103–107

19. Federico M, Percario Z, Olivetta E, Fiorucci G, Muratori C, Micheli A, Romeo G, Affabris E (2001) HIV-1 Nef activates STAT1 in human monocytes/macrophages through the release of soluble factors. Blood 98:2752–2761

20. Percario Z, Olivetta E, Fiorucci G, Mangino G, Peretti S, Romeo G, Affabris E, Federico M (2003) Human immunodeficiency virus type 1 (HIV-1) Nef activates STAT3 in primary human monocyte/macrophages through the release of soluble factors: involvement of Nef domains interacting with the cell endocytotic machinery. J Leukocyte Biol 74:821–832

21. Campbell TD, Khan M, Huang M-B, Bond VC, Powell MD (2008) HIV-1 Nef protein is secreted into vesicles that can fuse with target cells and virions. Ethn Dis 18:S2–S9

22. Ali SA, Huang M-B, Campbell PE, Roth WW, Campbell T, Khan M, Newman G, Villinger F, Powell MD, Bond VC (2010) Genetic characterization of HIV type 1 Nef-induced vesicle secretion. AIDS Res Hum Retroviruses 26:173–192

23. Leitner T, Foley B, Hahn B, Marx P, McCutchan F, Mellors J, Wolinsky S, Korber B, (2007) HIV Sequence Compendium 2006/2007. Los Alamos National Laboratory, NM 07–

24. van Marle G, Henry S, Todoruk T, Sullivan A, Silva C, Rourke SB, Holden J, McArthur JC, Gill MJ, Power C (2004) Human immunodeficiency virus type 1 Nef protein mediates neural cell death: a neurotoxic role for IP-10. Virol 329:302–318

25. Rhodes DI, Ashton L, Solomon A, Carr A, Cooper D, Kaldor J, Deacon N (2000) Characterization of three nef-defective human immunodeficiency virus type 1 strains associated with long-term nonprogression. Australian Long-Term Nonprogressor Study Group. J Virol 74:10581–10588

26. Catucci M, Venturi G, Romano L, Valensin PE, Zazzi M (2000) Analysis of the HIV-1 nef gene in five intravenous drug users with long-term nonprogressive HIV-1 infection in Italy. J Med Virol 60:294–299

27. Yamada T, Iwamoto A (2000) Comparison of proviral accessory genes between long-term nonprogressors and progressors of human immunodeficiency virus type 1 infection. Arch Virol 145:1021–1027

28. Combet C, Jambon M, Deleage G, Geourjon C (2002) Geno3D: automatic comparative molecular modelling of protein. Bioinforma 18:213–214

29. Case DA, Cheatham TE III, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) The Amber biomolecular simulation programs. J Comput Chem 26:1668–1688

30. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins 65:712–725

31. Price DJ, Brooks CL III (2004) A modified TIP3P water potential for simulation with Ewald summation. J Chem Phys 121:10096–10103

32. Jorgensen WL, Tirado-Rives J (2005) Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. Proc Natl Acad Sci USA 102:6665–6670

33. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera–a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612

34. Petsko GA, Dagmar R (2003) Protein structure and function. Wiley-Blackwell, London

35. Lovell SC, Davis IW, Arendall WB III, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC (2003) Structure validation by Calpha geometry: phi, psi and Cbeta deviation. Proteins 50:437–450

36. Raymond AD, Campbell-Sims TC, Khan M, Lang M, Huang MB, Bond VC, Powell MD (2011) HIV type 1 Nef is released from infected cells in CD45+ microvesicles and is present in the plasma of HIV-infected individuals. AIDS Res Hum Retroviruses 27:167–178. doi:10.1089/aid.2009.0170