PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

SHORT COMMUNICATION

Correlated mutations at gp120 positions 322 and 440: Implications for gp120 structure

Osnat Rosen,[1] Avraham O. Samson,[2] and Jacob Anglister[1]*

[1] Department of Structural Biology, Weizmann Institute of Science, Rehovot 76100, Israel
[2] Department of Structural Biology, Stanford University, Stanford, California 94305

ABSTRACT

Analysis of V3 and C4 sequences of HIV-1 reveals correlated mutations at gp120 positions 322 and 440, and a very strong preference for a positively charged residue at position 440 when position 322 is negatively charged. This observation suggests that these two residues are close to each other and interact electrostatically in R5 viruses. This interaction was used to model V3 in the context of gp120 using NMR data for the V3 loop and the crystal structure of the gp120-core. The interaction between residues 322 and 440 may serve as part of the molecular switch for HIV-1 phenotype conversion.

Proteins 2008; 71:1066–1070.

Key words: HIV; gp120; V3; correlated mutation; structure.

INTRODUCTION

Specific variations in the sequence of the third variable region (V3) of the HIV-1 glycoprotein gp120, involving charged amino acids, correlate with conversion of viral phenotype and coreceptor usage. In general, the V3 sequences of X4 isolates are more positively charged than those of R5 isolates. It has been suggested that the presence of a basic residue at V3 positions 306 or 322 is indicative of X4 and X4R5 viruses (the "11/25 rule").[1] Moreover, conversion of a negatively charged residue at position 322 (residue 25 of V3) to a positively charged residue was sufficient to convert an R5 strain to an X4 strain.[1–3]

Using NMR, our laboratory has reported two conformations of V3 β-hairpins in V3 peptides bound to HIV-1 neutralizing antibodies, leading us to suggest that they correspond to the R5 and X4 conformations of V3.[4,5] In the postulated R5 conformation, residue E322 opposes the conserved arginine at position 304, so that electrostatic attraction stabilizes the conformation.[5] Mutation at position 322 to a positively charged residue produces electrostatic repulsion that may trigger the conversion to the postulated X4 conformation.[5]

NMR and X-ray structures of different V3 peptides bound to HIV-1 neutralizing-antibodies showed that the N-terminal strand and the β turn of V3 interact extensively with the antibodies whereas the C-terminal strand has considerably fewer interactions.[4,6,7] The segment A316-E322 at the V3 C-terminal strand was found to make only a minor contribution to CCR5 binding and viral infection[8] compared with the opposing N-terminal strand (R304-I309). This could be because of partial occlusion of the C-terminal strand of the V3 in gp120 as a result of interactions with other domains of the envelope protein. The NMR studies revealed that while the conformation of the C-terminal is conserved in the two previously discovered conformations of the V3, the

The Supplementary Material referred to in this article can be found online at http://www.interscience.wiley.com/jpages/0887-3585/suppmat/

Grant sponsor: National Institute of Health; Grant number: GM 53329.

*Correspondence to: Jacob Anglister, Department of Structural Biology, Weizmann Institute of Science, Rehovot 76100, Israel. E-mail: jacob.anglister@weizmann.ac.il

Received 16 September 2007; Revised 19 December 2007; Accepted 31 December 2007
Published online 14 February 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21982

N-terminal strand is shifted by one residue with respect to the C-terminal strand. This shift leads to different pairing of the residues in the β-hairpin, a one register shift in the N-terminal strand residues forming hydrogen bond as well as 180° difference in the orientation of the side chains.[4] The conformational flexibility of the N-terminal strand and the conservation of the C-terminal strand are further indications that the C-terminal strand is involved in interactions within gp120 that restrict its flexibility.

Although the V3 region has a dominant role in phenotype conversion, other domains of gp120 may indirectly influence HIV-1 tropism. Statistical analysis of gp120 sequences revealed that the presence of a negatively charged residue at position 440 in the C4 domain of gp120 was correlated with a general increase in the positive charges in the V3 region, resulting in an X4 phenotype.[9,10] On the other hand, R5 viruses required the presence of a positively charged residue at position 440 to maintain a functional, replicating virus.[11] An interaction between V3 and the segment ^{429}EVGKAMYAPP^{438} (ending two residues upstream of residue 440) has been suggested based on immunochemical analysis of the gp120 surface.[12] Although the presence of a charged residue at position 440 was associated with phenotype conversion, this amino acid does not appear to interact directly with either coreceptor. Therefore, residue 440 may exert an indirect influence on the phenotype switch so that the correlation between residue 440 and the phenotype switch is a consequence of interaction between residue 440 and the V3.[9] We sought to learn whether a specific residue in V3 can be pinpointed as being responsible for the correlation between residue 440 and the charge of the V3 region. Since residue 322 was involved in phenotype conversion, this position was the prime candidate for testing. Interestingly, we found a very strong preference for a positively charged residue at position 440 when position 322 is negatively charged. This possible interaction was used to model V3 in the context of gp120. This interaction may serve as part of the molecular switch for HIV-1 phenotype conversion.

METHODS

Search for correlation between residue 322 and 440

To examine whether a correlation exists between the charge of amino acid at position 322 (in V3) and the charge of amino acid at position 440 (in C4), we analyzed all clade B sequences of gp120 that appeared in the Los Alamos data base and included both the V3 and C4 regions. To do so, HIV-1 clade B amino acid sequences available in the database (http://hiv-web.lanl.gov/content/hiv-db/mainpage.html) were downloaded. To identify the relevant gp120 sequences that include both V3 and C4 regions, the entire database was searched for sequences containing specific motifs representing the V3 and C4 regions of gp120. The basis for this search was the presence of conserved (99–100% conservation) residues in the entire V3 region (In particular, eight conserved amino acids were required: C296, R298, P299, G314, G325, R328, A330, and C332). The presence of a specific sequence, ^{435}YAPP^{438} that is conserved in the C4 region was required. This conserved segment is located 100–150 residues downstream of V3. All sequences were analyzed using a Perl script. Since the database is highly inhomogeneous with respect to sequence relatedness, single representatives were used in cases of identical sequences or sequences originating from the same individual. Three types of amino acid groups were defined according to charge: (+) for arginine and lysine; (−) for aspartate and glutamate and (0) for all other amino acids. The program finds the number of strains having a specific pair of residues at positions 322 and 440 out of the total number of strains having the same type of residue at position 322. The significance of the correlation between these two positions was calculated using the χ^2 test according to Kass and Horovitz.[13] To learn whether such correlation was or was not limited to clade B, the same procedure was repeated for clades A and C. In addition, to ensure unbiased results, the same calculations were performed for other possible "positive–negative" pairs within V3 and C4, i.e., residues 326/440, 326/444, 322/444, and 306/440.

Modeling the V3$_{JR\text{-}FL}$ into the gp120 core

The V3 structure in the context of gp120 was modeled using Crystallography and NMR System program and NMR data on the V3 conformation.[5] The structural template of the gp120 molcule was that of HIV-1$_{JR\text{-}FL}$ gp120-core with V3 in complex with CD4 receptor and X5 antibody (PDB accession: 2B4C).[7] The interaction between residues 322 and 440 was translated into an assumed inter-residue distance constraint of 3 Å between the carboxyl group of E322 and the N_ε- hydrogen of R440. The position of V3 residues P299-R327 and C4 residue 440 was allowed to change, while the position of all other gp120 residues was fixed (2B4C.pdb chain G residues V84-R298, Q328-I439, and G441-E492). The position of the first three residues of the N-terminal half (C296-R298) and the last four residues of the C-terminal half of the V3 loop (Q328-C331) was fixed since these residues adopt a well defined β-strand structure within 2B4C.pdb. Structures of the V3 loop in the HIV-1$_{JR\text{-}FL}$ gp120-core with V3 were calculated using all NMR constraints applied previously (to solve the structure of V3$_{JR\text{-}FL}$ peptide in complex with 447-52D antibody[4]).

RESULTS AND DISCUSSION

Approximately 3800 sequences contained both the V3 and the C4 segments according to our search require-

Table I
Correlation Between the Charges of V3 Residue 322 and C4 Residue 440 of Clade-B HIV-1

	322 (V3)		
440 (C4)	−	0	+
+	78[a]	54	28
0	17	31	51
−	5	15	21
Total	2642	668	463

[a]The values represent percent of HIV-1 clade B strains that have the indicated charge (positive, neutral, or negative) at residue 440 out of all HIV-1 clade B strains that have the indicated charge at position 322. *P*-value $< 10^{-4}$.

Table III
Correlation Between the Charges of V3 Residue 326 and C4 Residue 440 of Clade-B HIV-1

	326 (V3)		
440 (C4)	−	0	+
+	68[a]	63	60
0	22	34	40
−	10	3	0
Total	3199	569	5

[a]The values represent percent of HIV-1 clade B strains that have the indicated charge (positive, neutral, or negative) at position 440 out of all clade B HIV-1 strains that have the indicated charge at V3 residue 326.

ments. Their analysis revealed a significant correlation (*P*-value $<10^{-4}$) between the type of amino acid at position 322 in V3 and the type of amino acid at position 440 in C4: 70% of HIV-1 clade B isolates contain negatively charged residues at position 322, and of these, 78% have a positively charged residue at position 440, 17% have a neutral amino acid and only 5% have a negatively charged residue (Table I). When the amino acid at position 322 changes from negatively charged to a neutral amino acid (18% of the HIV-1 clade B sequences in this database) the percentage of positively charged residues at position 440 drops to 54%, while 31% are neutral and 15% are negative. Of the 12% of isolates having a positively charged residue at position 322, the percentage of sequences with a positively charged residue at position 440 drops to 28%, while 51% are neutral and 21% are negative (Table I). However, examination of clades A and C failed to find the same pattern (Supplementary Tables SI and SII). The other clade B "positive/negative" pairs that were tested (322/444, 326/440, 326/444, and 306/440) did not show any significant correlation (Tables II–IV). In these pairs, a clear preference for positive charge at position 440 and 444 was evident, regardless of the charge at positions 322 and 326.

The vast majority of gp120 sequences in the database are not characterized according to the phenotype of the virus from which they were derived. Examination of 21 clade B sequences for which the phenotype is known had previously revealed a correlation between the overall charge of V3 and the charge of the residue at position 440.[10] Interestingly, we now find that all strains with a negatively charged residue at position 322 (10 sequences) have a positively charged residue at position 440, and all of them are R5 viruses. Of the sequences having neutral residue at position 322 (six strains), five have neutral residues at position 440 and one has a positively charged residue, and most of these are dual tropic viruses (four of six). In the case of positively charged residue at position 322 (five sequences), three have neutral residues at position 440 and two have a negatively charged residue. Of these five strains, three are X4, and the other two are dual-tropic.

These data together with our results discussed above clearly indicate a very strong preference for a positively charge residue at position 440 when position 322 is negatively charged. This suggests an electrostatic interaction between residue 322 in the C-terminal strand of V3 and the C4 residue 440 in R5 viruses. When the charge at residue 322 changes from negative to positive, we see a drop from 78 to 28% positively charged residues at position 440, indicating that positive charge simultaneously at both positions is unfavorable, appearing in less than 3.5% of viruses. Combination of a charged residue at either 322 or 440 with a neutral residue at the other position occurs in ∼38% of HIV-1 clade B viruses. Our sequence analysis together with that presented by Yama-

Table II
Correlation Between the Charges of V3 Residue 322 and C4 Residue 444 of Clade-B HIV-1

	322 (V3)		
444 (C4)	−	0	+
+	80[a]	85	84
0	18	14	16
−	2	1	0
Total	2642	668	463

[a]The values represent percent of HIV-1 clade B strains that have the indicated charge (positive, neutral, or negative) at position 444 out of all HIV-1 clade B strains that have the indicated charge at V3 residue 322.

Table IV
Correlation Between the Charges of V3 Residue 326 and C4 Residue 444 of Clade-B HIV-1

	326 (V3)		
444 (C4)	−	0	+
+	81[a]	83	100
0	18	17	0
−	1	0	0
Total	3199	569	5

[a]The values represent percent of HIV-1 clade B strains that have the indicated charge (positive, neutral, or negative) at position 444 out of all clade B strains that have the indicated charge at V3 residue 326.

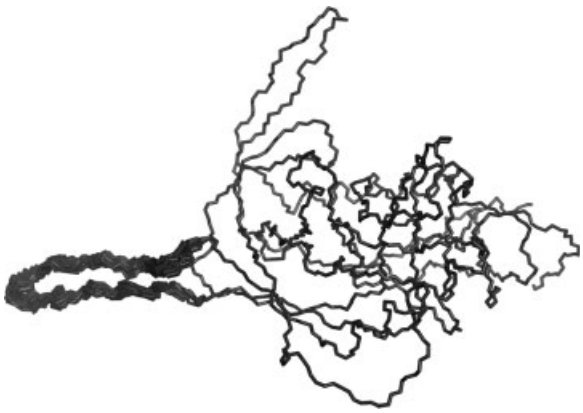

Figure 1

Backbone superposition of the 40 lowest-energy models for the gp120-core containing V3. The models were calculated according to the crystal structure of gp120 core and the base of the V3,[7] the NMR structure of $V3_{JR-FL}$ bound to the HIV-1 neutralizing antibody 447-52D[4] and a distance constraint between residues 322 and 440 deduced from the correlation between these two residues.

guchi-Kabata *et al.*[10] indicate that, while a strong preference for a positive charge at position 322 and negative charge at position 440 cannot be demonstrated for X4 viruses, the presence of two positive charges at both positions is nevertheless very unfavorable.

An interaction between residues 322 and 440 could be one of the anchors responsible for the conservation of the V3 C-terminal strand conformation observed in NMR studies. In the crystal structure of gp120-core with V3 present,[7] the distance between these two residues was 23–24 Å, which is too long to enable electrostatic interactions. However, the segment F317-I323 was not well defined in the crystal structure due to its disorder.[7] Therefore, we modeled the structure of $HIV\text{-}1_{JR-FL}$ based on gp120-core with V3 in complex with CD4 receptor and X5 antibody (PDB accession: 2B4C) and our NMR data for $V3_{JR-FL}$.[5] The backbone superposition of the 40 resulting lowest energy structures is shown in Figure 1. The structures define a β-hairpin consisting of two strands. Most of the residues are in an extended but not ideal β-strand conformation; however, two short well defined antiparallel β-strands are formed (Fig. 2, green arrows, residues S306-I307 and T319-T320). The electrostatic interaction created between V3 and C4 in the calculated gp120-model does not result in clashes with other domains in the crystal structure of the V3-containing gp120-core and is compatible with the NMR structure of V3-peptides bound to HIV-1 neutralizing antibodies.

The use of a distance constraint between residues 322 and 440 despite the observation that they are separated by 23–24 Å in the crystal structure as noted above can be justified by the fact that the crystal structure was obtained using a monomeric gp120 lacking the V1/V2 loops and gp41 as well as the N-linked glycosylation site (301N/Q) in the V3. On the virion's envelope, gp120 appears as a trimer interacting noncovalently with a trimer of gp41. The sequence correlations discovered in the present study should apply to the native gp120/gp41 trimeric structure. A model for the gp120/gp41 trimer,[14] based on the crystal structure of SIV and HIV-1 gp120, suggested interactions between the V3-region of one protomer with the V1/V2-region of an adjacent protomer and exhibits a propeller-like structure. This model explained the contribution of V1/V2 to neutralization resistance to anti-V3 antibodies observed in many primary isolates of HIV-1 and the discovery that some neutralizing antibodies recognize a discontinuous V2–V3 determinant.[15] The interaction of V3 with the V1/V2 region and possibly with gp41 and the lack of the glycosylation in the V3 could easily affect the conformation of V3 especially in those segments that show disorder in the crystal structure.[7] One of the disordered segments is located at the center of the V3 C-terminal strand and includes residue 322.[7] The calculated model that takes into account the crystal structure of the gp120-core and the base of the V3, as well as the NMR structure of V3 segment R304-E322 (which was obtained when bound to antibody that recognize the native gp120 structure, i.e., with glycosylation), can easily accommodate the correlation between residues at positions 322 and 440. Such a model also explains the previous suggestion of proximity between the V3 and the C4 region derived from immunochemical analysis.[12]

The pair of residues (E/D)322 and (R/K)440 could serve as part of the molecular switch for phenotype con-

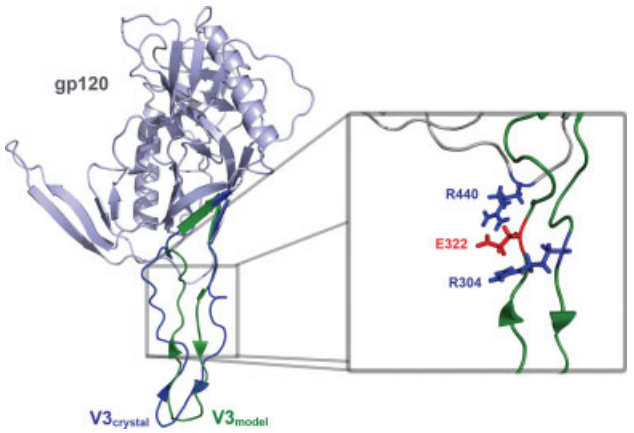

Figure 2

*Ribbon diagram of the superposition of the crystallographic structure of V3-containing gp120-core and lowest-energy NMR structure of V3 modeled onto gp120-core. (**A**) gp120 core of $HIV\text{-}1_{JR-FL}$ (gray),[7] the $V3_{JR-FL}$ structure according to the NMR constraints (green)[5] and the crystallographic structure of $V3_{JR-FL}$ as part of the V3-containing gp120-core (blue).[7] (**B**) Magnification of the V3 NMR-structure modeled onto gp120-core and the C4 segment that potentially interact with V3. The side-chains of R440 within gp120-core and R304 and E322 of the V3 NMR-structure are represented by stick and colored blue and red for positive and negative, respectively.*

version described previously.[4,5] This pair, perhaps in combination with other residues in the V3 C-terminal strand that interact with gp120 residues outside of V3, can serve as a pivot for the conformational switch by restricting the flexibility of the V3 C-terminus. The more flexible N-terminal strand will then orient itself relative to the C-terminal strand in one of two conformations that differ by a one register shift relative to the C-terminal strand to form favorable pair-wise interactions between the two opposing β-strands. Each conformation of the N-terminal strand will result in different surface topology that could influence binding to the chemokine receptors. Interestingly, such a pivot is also found in the β2,β3-hairpin of chemokines and is formed by a disulfide bond between a cysteine residue in the β3-strand of the β2-β3 hairpin and C11. Previously, we have shown an "analogy" between the alternative conformations of V3 and the conformations of the β2,β3-hairpin in the CCR5 chemokines and CXCR4 chemokine.[16] In both cases, a different topology of a β-hairpin surface that potentially interacts with the chemokine receptors was created by a one register shift of the N-terminal strand of the β-hairpin relative to the C-terminal strand.[4] Further studies are required to elucidate how the conformational change in the HIV-1 V3 and in the chemokine β2-β3 hairpin modulates the specific interactions of these molecules with CCR5 and CXCR4 receptors.

ACKNOWLEDGMENTS

We thank Dr. Sandy Livnat for editorial assistance and Orly Noivirt-Brik for bioinformatics and statistical guideness. J.A. is the Dr. Joseph and Ruth Owades Professor of Chemistry.

REFERENCES

1. Fouchier RA, Groenink M, Kootstra NA, Tersmette M, Huisman HG, Miedema F, Schuitemaker H. Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule. J Virol 1992;66:3183–3187.
2. De Jong JJ, De Ronde A, Keulen W, Tersmette M, Goudsmit J. Minimal requirements for the human immunodeficiency virus type 1 V3 domain to support the syncytium-inducing phenotype: analysis by single amino acid substitution. J Virol 1992;66:6777–6780.
3. Resch W, Hoffman N, Swanstrom R. Improved success of phenotype prediction of the human immunodeficiency virus type 1 from envelope variable loop 3 sequence using neural networks. Virology 2001;288:51–62.
4. Rosen O, Chill J, Sharon M, Kessler N, Mester B, Zolla-Pazner S, Anglister J. Induced fit in HIV-neutralizing antibody complexes: evidence for alternative conformations of the gp120 V3 loop and the molecular basis for broad neutralization. Biochemistry 2005;44: 7250–7258.
5. Rosen O, Sharon M, Quadt-Akabayov SR, Anglister J. Molecular switch for alternative conformations of the HIV-1 V3 region: implications for phenotype conversion. Proc Natl Acad Sci USA 2006;103:13950–13955.
6. Stanfield RL, Gorny MK, Williams C, Zolla-Pazner S, Wilson IA. Structural rationale for the broad neutralization of HIV-1 by human monoclonal antibody 447–52D. Structure (Cambridge) 2004;12: 193–204.
7. Huang CC, Tang M, Zhang MY, Majeed S, Montabana E, Stanfield RL, Dimitrov DS, Korber B, Sodroski J, Wilson IA, Wyatt R, Kwong PD. Structure of a V3-containing HIV-1 gp120 core. Science 2005;310: 1025–1028.
8. Suphaphiphat P, Essex M, Lee TH. Mutations in the V3 stem versus the V3 crown and C4 region have different effects on the binding and fusion steps of human immunodeficiency virus type 1 gp120 interaction with the CCR5 coreceptor. Virology 2007;360:182–190.
9. Hoffman NG, Seillier-Moiseiwitsch F, Ahn J, Walker JM, Swanstrom R. Variability in the human immunodeficiency virus type 1 gp120 Env protein linked to phenotype-associated changes in the V3 loop. J Virol 2002;76:3852–3864.
10. Yamaguchi-Kabata Y, Yamashita M, Ohkura S, Hayami M, Miura T. Linkage of amino acid variation and evolution of human immunodeficiency virus type 1 gp120 envelope glycoprotein (subtype B) with usage of the second receptor. J Mol Evol 2004;58:333–340.
11. Carrillo A, Ratner L. Human immunodeficiency virus type 1 tropism for T-lymphoid cell lines: role of the V3 loop and C4 envelope determinants. J Virol 1996;70:1301–1309.
12. Moore JP, Thali M, Jameson BA, Vignaux F, Lewis GK, Poon SW, Charles M, Fung MS, Sun B, Durda PJ, Åkerblom N, Wahren B, Ho DD, Sattentau QJ, Sodroski J. Immunochemical analysis of the gp120 surface glycoprotein of human immunodeficiency virus type 1: probing the structure of the C4 and V4 domains and the interaction of the C4 domain with the V3 loop. J Virol 1993;67:4785–4796.
13. Kass I, Horovitz A. Mapping pathways of allosteric communication in GroEL by analysis of correlated mutations. Proteins 2002;48:611–617.
14. Chen B, Vogan EM, Gong H, Skehel JJ, Wiley DC, Harrison SC. Structure of an unliganded simian immunodeficiency virus gp120 core. Nature 2005;433:834–841.
15. Etemad-Moghadam B, Karlsson GB, Halloran M, Sun Y, Schenten D, Fernandes M, Letvin NL, Sodroski J. Characterization of simian-human immunodeficiency virus envelope glycoprotein epitopes recognized by neutralizing antibodies from infected monkeys. J Virol 1998;72:8437–8445.
16. Sharon M, Kessler N, Levy R, Zolla-Pazner S, Gorlach M, Anglister J. Alternative conformations of HIV-1 V3 loops mimic beta hairpins in chemokines, suggesting a mechanism for coreceptor selectivity. Structure 2003;11:225–236.