Identification of the B cell superantigen-binding site of HIV-1 gp120

(human antibodies/immunoglobulin variable genes)

SAOUSSEN KARRAY AND MONCEF ZOUALI

Département d'Immunologie, Institut Pasteur, 28 rue du Dr. Roux, 75015 Paris, France

Communicated by Lewis H. Sarett, Viola, ID, December 2, 1996 (received for review October 12, 1996)

ABSTRACT Previous studies showed that the gp120 envelope protein of HIV-1 is able to crosslink membrane IgM on normal human B cells and to induce their activation in a V_H3 **immunoglobulin gene-family-specific manner. Because this VH gene family is the largest in the human repertoire, this superantigen (SAg) property is thought to have deleterious consequences for the host, including a progressive decline of B cells with progression of the HIV-1-induced disease. Here, we have identified the sequence motifs on gp120 involved in SAg binding to normal Igs. We show that this SAg-binding activity is present in gp120s from highly divergent isolates of HIV-1 belonging to clades derived from various geographical origins, and that carbohydrate residues are not essential for its expression. The SAg-binding site is formed by protein sequences from two regions of the gp120 molecule. The core motif is a discontinuous epitope spanning the V4 variable domain and the amino-terminal region flanking the C4 constant domain. The most critical residues appear to be Leu395– Asp³⁹⁷ and Ile425–Gln427. Residues from the C2 constant domain (positions 252–272) also seem to play an accessory role in SAg binding of gp120 to normal human Igs. These findings are important in the design of a successful gp120 based vaccine against HIV-1.**

After HIV infection, strong humoral and cellular immune responses are developed to various viral antigens, including the envelope glycoprotein gp120. The sequence of this molecule can be subdivided into five variable regions (V1–V5) with 25% or fewer conserved residues, and five constant regions (C1– C5) with greater than 75% conservation (1). By virtue of its accessibility to immune effectors, this glycoprotein has been shown to elicit virus-neutralizing antibodies in a variety of experimental systems and to bind a portion of the neutralizing antibodies present in sera from HIV-1-infected subjects (2). Remarkably, the important immunological determinants of gp120 are present on both its constant and variable regions. Because of its strong immunogenicity and its binding to CD4, the viral coreceptor, gp120 is considered a potential subunit vaccine candidate against this virus (3).

However, in addition to its role in triggering potentially protective responses, gp120 is known to participate in responses detrimental to the host, for example, by triggering autoimmune responses (4), by inhibiting natural killer activity (5), by enhancing HIV infectivity (6), and by exhibiting superantigen (SAg)-like properties for B cells (7). More specifically, gp120 binds a subpopulation of B cells from normal individuals by means of immunoglobulins of the V_H3 gene family and selectively induces Ig secretion by V_H3^+ B cells (7).

Copyright $@$ 1997 by The NATIONAL ACADEMY OF SCIENCES OF THE USA 0027-8424/97/941356-5\$2.00/0

PNAS is available online at **http://www.pnas.org.** Abbreviation: SAg, superantigen.

This SAg-like property is thought to be responsible for the increase in V_H3 ⁺ B cells during early stages of the infection followed by their progressive decline during advanced stages of the disease (8, 9). Given the severity of this B cell depletion, we have characterized the epitopes on the gp120 molecule engaged in B cell SAg interactions.

MATERIALS AND METHODS

Recombinant Proteins. Recombinant gp120s from different HIV-1 isolates were obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, National Institute of Allergy and Infectious Diseases, National Institutes of Health from MicrogeneSys (West Haven, CT) (gp120_{CM}, gp120_{LAV}, gp120_{MN}) and from K. Steimer (Chiron) (gp120_{SF2}). Recombinant gp160 proteins were provided by M.-P. Kieny (Transgene, Strasbourg, France). Their characteristics are summarized in Table 1.

Synthetic Peptides. Peptides covering the entire sequence of gp120 were obtained from the Agence Nationale de la Recherche sur le SIDA (Paris), and through the AIDS Research and Reference Reagent Program. All peptides were soluble at neutral pH, except peptides SP9045, SP89247, and SP89049, which had to be solubilized in an acidic buffer. Their sequences are presented in Table 2.

Human Monoclonal Igs. Human monoclonal IgG and IgM expressing variable regions of the V_H3 gene family were purified from the serum of patients with multiple myeloma or Waldenström macroglobulinemia, as described previously (11).

Binding of Igs to HIV-1 Envelope Proteins. Direct binding was assessed by enzyme-linked immunosorbent assay (ELISA). Polystyrene microtiter plates (Maxisorp F96, Nunc) were incubated overnight at $+4^{\circ}$ C with 100 μ l of either rgp120 or rgp160 antigens (50 ng per well) diluted in borate-buffered saline (pH 8.4). Plates were washed with PBS (pH 7.2)

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked ''*advertisement*'' in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Numbering is according to Ratner *et al*. (10). Domain assignment is according to ref. 2. Each amino acid is represented by a single-letter code that is the first letter of its name, except for arginine (R), asparagine (N), glutamine (Q), glutamic acid (E), lysine (K), phenylalanine (F), tryptophan (W), and tyrosine (Y) .

containing 0.1% Tween-20 (Prolabo, Paris), and nonspecific sites of the wells were blocked with PBS containing 1% (wt/vol) BSA for 2 h at 37° C. A 100- μ l volume of the test antibody diluted in PBS/BSA was added to each well and, after 2 h of incubation at 37° C, the plates were washed again with PBS/Tween. Bound antibodies were revealed with an alkaline phosphatase anti-immunoglobulin conjugate diluted (1:1000) in PBS/BSA. After a further incubation for 1 h at 37° C, the plates were washed with PBS/Tween and a volume of 100 μ l per well of the alkaline phosphatase substrate *p*-nitrophenyl phosphate (Sigma) (1 mg/ml) diluted in 0.05 M carbonate buffer, pH 9.5, containing 2 mM $MgCl₂$, was added. Absorbance was recorded at 405 nm with a multiscan automatic plate reader (MR 5000, Dynatech).

Competitive Immunoassays. Plates were coated overnight at 4° C with recombinant proteins, and, after several washings with PBS/Tween, they were blocked with PBS/BSA. The test Ig at a dilution that achieved 50% of its maximal binding to the coated protein was mixed with various concentrations of the competitor peptide solution. After an overnight incubation at 48C, the mixtures were transferred to protein-coated wells. The ELISA then was continued as described above. The percent inhibition of Ig binding to its target protein by the competitor peptide was calculated according to the formula

 $%$ inhibition = 100

$$
\times \left[1 - \left(\frac{A_{405} \text{ in the presence of the competitor peptide}}{A_{405} \text{ in the absence of the competitor peptide}}\right)\right].
$$

RESULTS

Recombinant gp120s from Divergent HIV-1 Isolates Display SAg-Like Activity. The gp120 molecule is a SAg for normal B cells, and it interacts, through this SAg-binding site, with V_H3^+ Igs from normal subjects (7). To obtain further information on the gp120 SAg-binding site, we used two monoclonal Igs (IgM LAN and IgG SAU) isolated from two individuals who have not been exposed to HIV-1. First, we performed direct binding studies with rgp120 and rgp160 from various isolates. Fig. 1 shows the dose-dependent binding of monoclonal IgG SAU to divergent rgp120s. It also demonstrates that glycosylated and unglycosylated forms of gp120 from different clades exhibit the SAg-binding property, implying that the activity does not depend on glycosylation. Similar results were obtained with IgM LAN (not shown). To probe the effect of protein denaturation on the SAg-binding site, gp120_{SF2} was heated at 100°C in 0.1% sodium dodecyl sulfate. This treatment did not affect the SAg-binding activity. These data suggest that the SAg site is present on gp120s from sequence-divergent isolates and that it is not particularly sensitive to gp120 conformation, in that it survives heat denaturation.

The B Cell SAg-Binding Site Maps to Distant Motifs of gp120. Next, we undertook a series of inhibition experiments aimed at precisely mapping the Ig-binding site on gp120. For this purpose, we tested the inhibitory capacity of a panel of synthetic peptides encompassing the whole sequence of the gp120 molecule. When tested at equimolar concentrations in a competitive assay using monoclonal IgM LAN, only peptides SP93855, SP89261, and to a lesser degree, SP89045, had an inhibitory activity (Fig. 2). All other peptides completely failed to induce inhibition, even at high concentrations. Similar results were obtained with monoclonal IgG SAU (not shown). To gain further insight into the complexity of the binding site, we performed crosscompetition experiments with the three positive peptides used in the different possible combinations. Individually, when tested at a 5 μ M concentration, peptides SP93855 and SP89261 produced 80% and 75% inhibition, respectively. When both peptides were simultaneously used in an equimolar mixture, they had no synergistic or additive effect on Ig SAg binding to gp120. A similar trend was seen in inhibition experiments with other combinations of the peptides. These results imply that the binding site probably is centered on the V4 variable domain (peptides SP93855 and SP89261) of gp120 with contribution from the C2 constant domain (peptide SP89045), and that it is clearly not sensitive to the presence of N-linked glycosylation sites in these areas of the molecule.

Delineation of Critical Amino Acid Positions by Mutagenesis. To further characterize the SAg-binding site, we determined whether rgp120 binding was affected by amino acid substitutions. First, the fact that peptide SP89045 (residues 252-272) was inhibitory of SAg activity, and peptide 1960 (residues 241-260) was inactive, suggests that amino acids 252-260 are not engaged in gp120–Ig interaction. We therefore synthesized M1 mutant peptide, where the polar amino acids $G\ln^{263}$ and Asn²⁶⁷ were replaced by the apolar amino acid Ala.

FIG. 1. SAg binding of IgG SAU to rgp120s from different HIV-1 isolates. Microtiter wells were coated with 50 ng of rgp120s, and Ig binding was revealed by an alkaline phosphatase-labeled anti-IgG conjugate. The results are from one representative experiment. gp120- $SF2$ ng refers to an unglycosylated form of $gp120_{SF2}$ that was heat inactivated.

FIG. 2. Inhibition of SAg binding by synthetic peptides from distant sites of gp120. Dilutions of the synthetic peptides were mixed with an equal volume of IgM LAN at a concentration that achieved 50% of maximal binding. After incubation, the mixtures were transferred to $gp120_{MN}$ -coated wells, and residual binding was revealed as indicated in the text. The results are from one representative experiment.

Second, for peptide SP93855, which maps to the V4 domain, sequence comparison of the gp120 from the various isolates tested revealed that, despite the sequence variability of this domain, some positions are conserved (Fig. 3). We have therefore used $\overline{M}2a$ mutant peptide, where amino acids Leu³⁹⁵, Phe³⁹⁶, and Asn³⁹⁷ were replaced by Ala. Finally, for SP89261 we designed two mutant peptides, M3a and M3b, where the conserved amino acids Ile^{419} and Leu^{421} , and Ile^{425} , His⁴²⁶, and Gln⁴²⁷ were replaced by Ala, respectively. Competition experiments showed that only the amino acid substitutions in M2a and M3b mutant peptides significantly impaired or abolished SAg binding (Fig. 4), implying that $Leu³⁹⁵ – Asn³⁹⁷$ and $Ile⁴²⁵ –$ Gln⁴²⁷ amino acid positions are critical for gp120 SAg binding.

DISCUSSION

A definitive assignment of Ig interaction sites with foreign ligands requires information on the three-dimensional structure derived, for example, from crystallographic data for Ig–ligand complexes. In most cases, however, epitopes can be identified by indirect means such as reactivity with phylogenetically related ligands, or by protein modification using either chemical procedures or site-directed mutagenesis (12). For infectious agents exhibiting antigenic variation and evolving at a high rate, such as HIV-1, study of proteins from various isolates provides a convenient means for identifying critical residues necessary for immune recognition. In this report, we have applied this approach to map the SAg-binding site of gp120 for normal human Ig. This was possible because HIV-1 isolates from different infected individuals vary in primary sequences. In addition, there is a deeper level of variation between isolates from different phylogenetically defined clades, with nine clades (A through I) identified as a result of coordinated studies of HIV-1 diversity (13). Using gp120s from different clades, we have shown that the SAg-binding activity is well conserved, a feature reminiscent of the conservation of the CD4-binding site of gp120 between different viral clades (14). That gp120s from different clades exhibit the SAg property is not inherently surprising. HIV-1 clades are defined on the basis of primary sequence, whereas SAg binding is a much more complex process that depends on interactions with motifs that are likely to be influenced by the tertiary and quaternary structures of the protein (see below). It will be important to see whether human monoclonal antibodies elicited in response to HIV infection interfere with the SAgbinding activity.

Our studies with gp120 synthetic peptides indicate that the SAg-binding region of gp120 is composed of two discrete

FIG. 3. Sequence comparison of the gp120 molecules from different isolates. Shown are the locations of the variable regions. The sequences corresponding to the peptides that inhibit gp120 SAg binding are boxed. Gaps are introduced to maximize homology and dots indicate homology. Numbering is according to Ratner *et al.* (10).

binding sites, which allow for the simultaneous binding of V_H 3⁺ Ig molecules. We have shown that residues at the base of the V4 domain and the amino-terminal region flanking the C4 domain of gp120 (positions 392-434) are critical for binding. It is remarkable that this latter domain is also the target of other molecular interactions. Discontinuous epitopes overlapping the C4 region are the target of a large proportion of

..--.....T.....N............E........R-----------

CAR

anti-HIV antibodies from most seropositive subjects, and neutralizing antibodies bind linear or discontinuous C4 epitopes (2). This domain also participates, with the C2 and C3 domains of gp120, in forming the CD4-binding domain to the CD4 coreceptor (2), implying that the C4 domain is accessible for molecular interactions. It is of interest that the C4 epitope associated with the SAg activity also contributes to the CD4

FIG. 4. Mutagenesis analysis of the sequence motifs involved in SAg binding. Dilutions of the wild-type and mutant peptides (M1, M2a, M3a, and M3b) were mixed with an equal volume of IgM LAN at a concentration that achieved 50% of maximal binding. After incubation, the mixtures were transferred to gp120_{MN}-coated wells, and residual binding was revealed as indicated in the text. Represented is the inhibition achieved by the peptides used at a 5μ M concentration. The results are from one representative experiment.

binding site and the mutations that alter SAg activity within this site also may affect CD4 binding to gp120 (2). In addition, our data suggest that SAg binding requires contributions of residues from the C2 domain (positions 252-272).

On the basis of their ability to be presented in the form of short peptides, we have categorized the SAg-binding site as being composed of two widely separated linear segments. Our data show that the core motif is a discontinuous epitope spanning the V4 domain and the amino-terminal region flanking the C4 domain, with a contribution from a segment of the distant C2 domain. The finding that two motifs located on different domains of the gp120 molecule are involved in SAg binding would suggest that they are located in close proximity in the native structure. There is, however, strong evidence demonstrating physical interactions between distant domains of the gp120 molecule, including the V2 and V3 domains (15), the V2 and C4 domains (16), the V3 and C4 domains (17), and all three domains (15). These interdomain interactions suggest that other regions of the molecule distant from the linear sequences we have identified might be indirectly implicated in optimal expression of the SAg-binding motif.

Although the secondary structure of gp120 has not been elucidated, it has been proposed that it may contain as many as six α -helices and a number of β -strands (1, 2). To provide a theoretical model of the secondary structure of the SAgbinding site and its adjacent areas, we analyzed the gp120 segments between residues 252-272 and 392-434 with regard to the distribution of polar and nonpolar residues, and α -helices and β -strands. Amino acid analysis of the V4 loop domain in the seven isolates derived from distinct geographic origins shows that the amino acids adjacent to the amino-terminal Cys are quite conserved (Fig. 3). In addition, a previous study predicted that residues 394-412 in this domain comprise a β -turn region, with amino acids that are hydrophilic and flexible and, hence, have a higher probability of being exposed at the surface of the gp120 molecule (1). These features may account for our finding that mutations at the V4 epitope (residues 392-414) significantly affect SAg binding. On the other hand, the $V4/C4$ epitope (residues 414-434) contains a potentially surface-exposed α -helix (2). Changing three amino acids of this helix dramatically impairs SAg binding (Fig. 4), a finding in agreement with another study (18). Thus, the results presented herein support the notion that the core SAg binding is mediated by two discrete, but separate, sites that map to amino acid positions 392-434. The most critical residues at the amino terminus appear to be Leu³⁹⁵-Asp³⁹⁷. At the carboxyl terminus Ile⁴²⁵, His⁴²⁶, and Gln⁴²⁷ are also critical. In addition, residues from the C2 domain (positions 252-272) seem to play an accessory role in gp120 SAg binding to normal human Igs.

These findings are potentially important in the design of an efficient gp120-based vaccine against HIV-1. The SAg-binding site must endow gp120 with the ability to crosslink membrane IgM on normal human B cells to induce their activation in a V_H 3 gene-family-specific manner (7). Because this V_H family is the largest in the human repertoire (19, 20), this property is thought to have deleterious consequences for the immune system, including a progressive decline of B cells with progression of the disease (9, 21) and formation of complementactivating immune complexes with subsequent activation of the complement cascade (22, 23). These potentially deleterious SAg effects might be offset by the use of a gp120 vaccine devoid of its SAg-binding site.

We are indebted to Drs. J. P. Bouvet (Institut Pasteur, Paris) and P. Lambin (Institut National de Transfusion Sanguine, Paris) for the monoclonal antibodies, the Agence Nationale de Recherche sur le SIDA (Paris) for the HIV synthetic peptides, and Dr. M.-P. Kieny (Transgène, Strasbourg, France) for HIV recombinant proteins. We thank Dr. G. Bentley (Institut Pasteur, Paris) for helpful discussions. This work was supported by grants from the Institut Pasteur and the Fondation pour la Recherche Médicale. S.K. is an Investigator of the Centre National de la Recherche Scientifique, and M.Z. is an Investigator of the Institut National de la Santé et de la Recherche Médicale.

- 1. Modrow, S., Hahn, B. H., Shaw, G. M., Gallo, R. C., Wong-Staal, F. & Wolf, H. (1987) *J. Virol.* **61,** 570–578.
- 2. Poignard, P., Klasse, P. J. & Sattentau, Q. J. (1996) *Immunol. Today* **17,** 239–246.
- 3. Levy, J. A. (1993) *Microbiol. Rev.* **57,** 183–289.
- 4. Ziegler, J. L. & Stites, D. P. (1986) *Clin. Immunol. Immunopathol.* **41,** 305–313.
- 5. Robinson, J., Edward, W., Mitchell, W. M., Chambers, W. H., Schuffman, S. S., Montefiori, D. C. & Oeltman, T. N. (1988) *Hum. Pathol.* **19,** 535–540.
- 6. Robinson, W. E. J., Montefiori, D. C. & Mitchell, W. M. (1988) *Lancet* **i,** 790–794.
- 7. Berberian, L., Goodglick, L., Kipps, T. J. & Braun, J. (1993) *Science* **261,** 1588–1591.
- 8. Berberian, L., Shukla, J., Jefferis, R. & Braun, J. (1994) *J. Acquired Immune Defic. Syndr.* **7,** 641–646.
- 9. Zouali, M. (1995) *Immunol. Today* **16,** 399–405.
- 10. Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, B., Josephs, S. F., Doran, E. R., Rafalski, J. A., Whitehorn, E. A., Baumeister, K., Ivanoff, L., Petteway, S. R., Jr., Pearson, M. L., Lautenberger, J. A., Papas, T. K., Ghrayeb, J., Chang, N. N., Gallo, R. C. & Wong-Stall, F. (1985) *Nature (London)* **313,** 277–284.
- 11. Silverman, G. J., Pire`s, R. & Bouvet, J. P. (1996) *J. Immunol.* **157,** 4416–4502.
- 12. Benjamin, D. C., Berzofsky, J. A., East, I. J., Gurd, F. R. N., Hannum, C., Leach, S. J., Margoliash, E., Michael, J. G., Miller, A., Prager, E. M., Reichlin, M., Sercarz, E. E., Smit-Gill, S. J., Todd, P. E. & Wilson, A. C. (1984) *Annu. Rev. Immunol.* **2,** 67–101.
- 13. Gao, F., Morrison, S. G., Robertson, D. L., Thornton, C. L., Craig, S., Karlsson, G., Sodroski, J., Morgado, M., Galvao-Castro, B., von Briesen, H., Beddows, S., Weber, J., Sharp, P. M., Shaw, G. M. & Hahn, B. H. (1996) *J. Virol.* **70,** 1651–1657.
- 14. Moore, J. P., McCutchan, F. E., Poon, S. W., Mascola, J., Liu, J., Cao, Y. & Ho, D. D. (1994) *J. Virol.* **68,** 8350–8364.
- 15. Choe, H. R. & Sodroski, J. (1995) *J. Virol.* **69,** 2801–2810.
- 16. Moore, J. P., Sattentau, Q. J., Yoshiyama, H., Thali, M., Charles, M., Sullivan, N., Poon, S. W., Fung, M. S., Traincard, F., Pinkus, M., Robey, G., Robinson, J. E., Ho, D. D. & Sodroski, J. (1993) *J. Virol.* **67,** 6136–6151.
- 17. Moore, J. P., Thali, M., Jameson, B. A., Vignaux, F., Lewis, G. K., Poon, S. W., Charles, M., Fung, M. S., Sun, B., Durda, P. J., Akerblom, L., Wahren, B., Ho, D. D., Sattentau, Q. J. & Sodroski, J. (1993) *J. Virol.* **67,** 4785–4796.
- 18. Goodglick, L., Zevit, N., Neshat, M. S. & Braun, J. (1995) *J. Immunol.* **155,** 5151–5159.
- 19. Zouali, M. (1994) *Nat. Genet.* **7,** 118–120.
- 20. Cook, G. P. & Tomlinson, I. M. (1995) *Immunol. Today* **16,** 237–242.
- 21. Goodglick, L. & Braun, J. (1994) *Am. J. Pathol.* **144,** 623–636.
- 22. Goodglick, L., Townsley-Fuchs, J., Fairhurst, R. M., Neshat, M., Marcelo, D., Mosier, D. E. & Braun, J. in *Human B-Cell Superantigens*, ed. Zouali, M. (Landes, Austin, TX), pp. 145–160.
- 23. Karray, S., Juompan, L. & Zouali, M. (1996) in *Human B-Cell Superantigens*, ed. Zouali, M. (Landes, Austin, TX), pp. 161–177.