Binding of Full-Length HIV-1 gp120 to CD4 Induces Structural Reorientation around the gp120 Core

Ashish,* Renu Garg,† Juan Anguita,‡ and Joanna K. Krueger*

Departments of *Chemistry and †Biology, University of North Carolina, Charlotte, North Carolina 28213; and ‡Department of Veterinary and Animal Sciences, University of Massachusetts, Amherst, Massachusetts 01003

ABSTRACT Small-angle x-ray scattering data on the unliganded full-length fully glycosylated HIV-1 gp120, the soluble CD4 (domains 1–2) receptor, and their complex in solution are presented. Ab initio structure restorations using these data provides the first look at the envelope shape for the unliganded and the complexed gp120 molecule. Fitting known crystal structures of the unliganded SIV and the complexed HIV gp120 core regions within our resultant shape constraints reveals movement of the V3 loop upon binding.

Received for publication 2 June 2006 and in final form 6 July 2006.

Address reprint requests and inquiries to J. K. Krueger, E-mail: jkkruege@uncc.edu.

The binding of HIV-1 and SIV gp120 to its receptor (CD4) and coreceptor on the host cell surface initiates the viral fusion and entry process. Structures of the complex of HIV-1 gp120 core with the soluble CD4 (sCD4) receptor (1,2), of the unliganded SIV gp120 core (3,4), and of the V3-containing HIV-1 gp120 core cocomplexed with both sCD4 and X5 antibody (4) have been solved. The ''core'' region of the HIV gp120 protein, for which crystal structures are known, lacks N- and C-terminal extensions, three variable (V1-V3) loops, and glycosylation (~50%). These loop regions play an important role in molecular recognition and have been the limiting factor in resolving the crystal structure of the full-length gp120 molecule (1–5). Placement of the stem residues for the missing loops within the known structures imply significant conformational changes within gp120 during complex formation. We used solution small-angle x-ray scattering (SAXS) on the full-length HIV-1 gp120, sCD4 (D1-D2), and their 1:1 complex under native-like conditions to model the individual components and the resultant shape changes upon complexation.

The proteins, gp120 and sCD4, were gifted by the National Institutes of Health Research and Reference Reagent Program. SAXS data were collected at beam line X21 at National Synchrotron Light Source (Brookhaven National Laboratory). The wavelength of the beam was 1.24 Å and the ratio of the sample/detector distance to the diameter of the charge-coupled device detector (D/d) was 9.94M. Samples (30 μl) of the individual proteins, their complex, and phosphate buffer saline pH 7.4 were exposed for 60 s, each at 10°C and at a flow-rate of 27 μL/min. SAXS data collected at several concentrations of hen egg white lysozyme (14.2 kDa; ACROS, Morris Plains, NJ) were run in the same sample cell under identical conditions. Images were circularly averaged, buffer-subtracted, and scaled to obtain relative scattering intensity (I) as a function of q (reciprocal space, $q = [4\pi\sin\theta]/\lambda$). All experiments were carried out in duplicate.

Guinier analysis, using PRIMUS (6), illustrate a linear fit to the unliganded gp120 data (Fig. 1 *A*, *inset*), the sCD4 data (Fig. 1 *C*, *inset*), and the SAXS data of their mixture (Fig. 2 *A*, *inset*) over the expected q range (where $q_{max} \cdot R_g \leq 1$). This analysis demonstrates monodispersity of scattering particles in solution. An indirect Fourier transform of the SAXS data, using GNOM45 (7), gave the pairwise distance distribution function in real space, $P(r)$. The beam intensity at zero-angle (I_0) determined from GNOM analysis of the lysozyme data, which had been extrapolated to infinite dilution, was measured at 9.7 ± 0.1. This value was used to calculate the concentrations of the individual proteins (0.95 mg/ml of gp120 and 1.25 mg/ml of sCD4). A 1.3:1 molar mixture of sCD4/gp120 was made by adding 10 μl of sCD4 (46 μM) to 47 μl of gp120 (7.8 μM). I_0 analysis of this mixture predicts that 73% of the proteins had formed a 1:1 complex under these conditions. Assuming formation of a complete 1:1 complex, actual protein concentration for the gp120/sCD4 complex would be maximum 0.95 mg/ml (6.5 μM), leaving 1.73 μM (0.04 mg/ml) of sCD4 uncomplexed in solution. (The uncomplexed sCD4 would contribute only 1.8% to the scattering from the mixture). $P(r)$ analysis of the gp120 data (Fig. 1 *B*) solves for an R_g of 43.6 Å and a D_{max} of 166 Å. Indirect Fourier transformation of the sCD4 data over a q range of 0.01–0.25 Å^{-1} solves for an R_g of 22.3 Å and a D_{max} of 75 Å (Fig. 1 *D*). This $P(r)$ curve exhibits a peak and a shoulder indicative of the expected bilobal shape for the two domains of sCD4. $P(r)$ analysis of the SAXS data for the mixture (q range of 0.01–0.25 Å^{-1}) solves well for an R_g of 44.4 Å with a D_{max} of 165 Å (Fig. 2 *B*). Values for the I_0, R_g, and D_{max} determined from the SAXS data analysis are summarized in Table 1.

GASBOR22IQW (8) was employed for ab initio shape restorations, whereby the protein structure is represented by an ensemble of dummy residues forming a chain-compatible model and tested for best fit to the $I(q)$ scattering data. The models of unliganded gp120, sCD4, and their complex used

doi: 10.1529/biophysj.106.090381

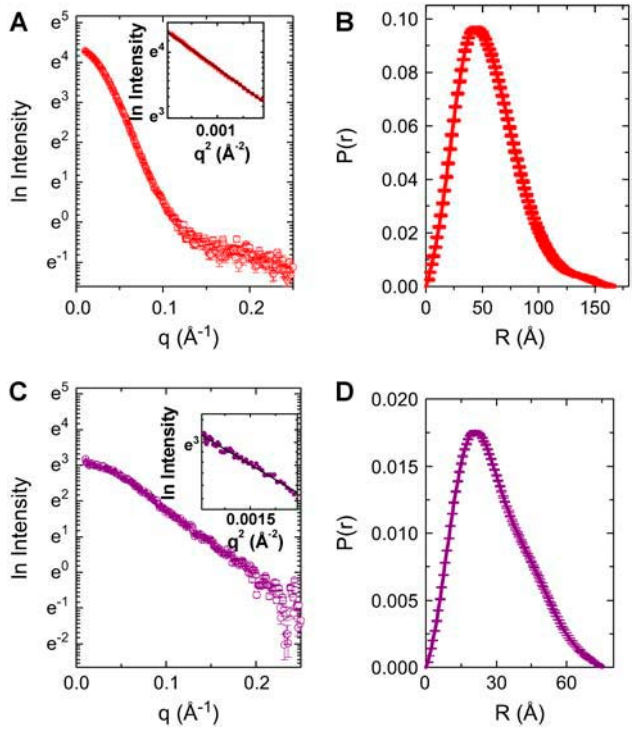

FIGURE 1 SAXS data (*A*) and *P*(*r*) plot (*B*) for unliganded full-length HIV-1 gp120. SAXS data (*C*) and *P*(*r*) curve (*D*) for the uncomplexed CD4 (D1-D2 domains). Linear fits from the Guinier analysis of these data are shown in the insets.

2168, 298, and 2387 dummy residues (packing radius 1.9 Å each), and 1598, 234, and 1598 dummy water molecules, respectively. Each calculation was repeated 10 times, with no shape or symmetry bias (Supplementary Figs. 1–3). For each data set modeled, one representative shape restoration model was selected based on the frequency of occurrence of the major shape features and are shown in Figs. 3 and 4.

The shape of the gp120 molecule in solution is characterized by a dense central core and less dense regions on three sides of the core region (Fig. 3 *A*). Using SUPCOMB13 (9), inertial axes of this gp120 model were overlapped with those for the unliganded SIV gp120 core (Protein Data Bank (PDB) No. 2BF1) (3). The automated superposition of high- and low-resolution models show that SIV gp120 core resides within the dense region of our model (Fig. 3 *B*). SAXS-based model of unliganded gp120 thus displays the volume, which

TABLE 1 Summary of *P*(*r*) analysis using GNOM45

Sample	Molecular mass (kDa)	I_0 Observed	I_0/c Expected	Protein (mg/ml)	R_g (Å)	D_{max} (Å)
Lysozyme	14.2	9.7	–	1	14.6	45
gp120	120	77.5	82.0	0.95	42.4	166
sCD4	26	21.6	17.8	1.25	22.3	75
1:1 complex	146	88.2	95.4	0.88	44.4	165

The q range input for each analysis was from 0.01 to 0.25 Å^{-1}.

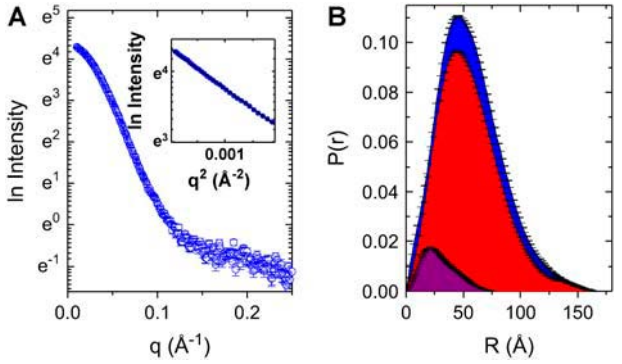

FIGURE 2 (A) SAXS data for the 1.3:1 mixture of sCD4/gp120. Inset shows the linearity of the Guinier plot. (*B*) A comparison of the *P*(*r*) curves calculated from the SAXS data for the unliganded gp120 (*red*), sCD4 (*purple*), and their mixture (*blue*).

the variable domains as well as high levels of glycosylation must be occupying in solution. The SIV gp120 protein (44.3 kDa) used for the crystal structure has ~65% smaller molecular weight than that used in our solution studies. Ab initio structure restoration for sCD4 (Fig. 3 *C*) clearly shows the expected low density between the two domains, also seen in the shape of the $P(r)$ plot. Overlapping the inertial axes of our scattering envelope with those for the unliganded sCD4 (D1-D2) (PDB No. 3CD4) (10) showed excellent agreement in shape and size between the two models (Fig. 3 *D*).

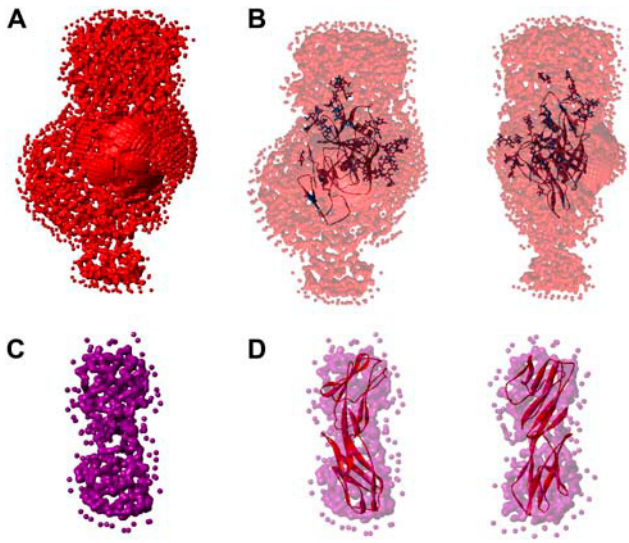

FIGURE 3 Scattering envelopes of (*A*) the full-length unliganded gp120 molecule with all dummy residues (*red*) shown in space-fill mode. (*B*) Two orthogonal views of the superimposition of inertial axes of our SAXS model (*transparent*, *red*) and the crystal structure of glycosylated SIV core, PDB No. 2BF1 (*purple ribbon*). (*C*) Space-filled representation (*purple*) of the scattering envelope that best fit the sCD4 SAXS data. (*D*) Two views of the sCD4 model (*transparent*, *purple*) superimposed with the crystal structure (*red ribbon*) of unliganded CD4, PDB No. 3CD4.

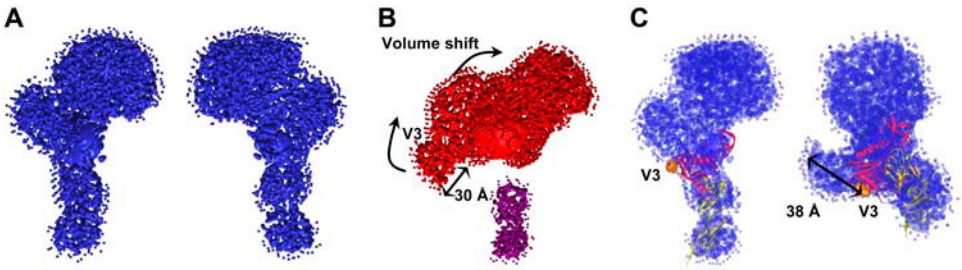

FIGURE 4 (*A*) Two views of the best-fit model of the 1:1 complex between full-length HIV-1 gp120 and sCD4 in solution. Dummy residues (*blue*) are shown in space-fill mode. (*B*) SAXS-based models of unliganded gp120 (*red*) and sCD4 (*purple*). (*C*) Two views of gp120 core (*pink ribbon*) and sCD4 (*yellow ribbon*) crystal structures, extracted from PDB No. 1G9M fit within the best-fit model of the SAXS data on the complex (*transparent, blue*).

The scattering envelope computed for the sCD4/gp120 SAXS data (Fig. 4 *A*) reveals three distinct features: the bilobal shape of the two domains of sCD4, a densely packed region over the top of one lobe of sCD4, and a less dense region that corresponds to the gp120 glycoprotein. The envelope volume is 303,000 Å^3, which agrees well with the sum total envelope volumes calculated for the individual components (25,900 and 276,000 Å^3), indicating that flexibility does not change considerably upon binding. Alignment of inertial axes of only the sCD4 domains from the crystal structures (PDB Nos. 1G9M and 1GC1) with the small bilobal arm protrusion in our SAXS-based model places the gp120 core structure from the crystal structures within the dense region of our model (Fig. 4 *C*). The additional volume around two sides of D1 domain in complex could be arising from the β-strands of gp120 core clamping that lobe as in 1GC1 or 1G9M (1,2). Comparing Fig. 4, *B* and *C*, it is clear that the binding of gp120 core to the apex of D1 domain of CD4 drastically reorients the less dense regions of gp120 not seen previously in crystal structures containing only the cores.

Our SAXS data provide structural constraints for modeling the interaction between the viral envelope protein of HIV-1 and its host cell receptor CD4. We propose that the small lobe sticking out (~30 Å) from the unliganded protein is likely to be the V3 loop of gp120. The small orange volume (Fig. 4 *C*) depicts the four residues where the V3 loop would have been connected in the crystal structures 1G9M or 1GC1. Assuming that this ''lobe'' does represents the V3 loop, then binding to CD4 induces this loop to stick out away from the compact core, even in the absence of a coreceptor. This suggestion is exemplified by an increased exposure (~38 Å) and considerable movement of the small ''V3'' lobe away from the CD4/gp120 binding interface (Fig. 4 *C*, *right side*).

SUPPLEMENTARY MATERIAL

An online supplement to this article can be found by visiting BJ Online at http://www.biophysj.org.

ACKNOWLEDGMENTS

The authors thank the National Synchrotron Light Source (Brookhaven National Laboratory) general user program for beam time allocation, and acknowledge invaluable support from Dr. Lin Yang for data acquisition and processing.

The following reagents were obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, National Institute of Allergy and Infectious Diseases (NIAIDS), National Institutes of Health: sCD4-183 from Pharmacia (Uppsala, Sweden) and HIV-1 SF162 gp120 from the NIAIDS Division of AIDS. This work was supported by the National Science Foundation (Career Award Grant No. MCB-0237676 to J.K.K.) and the National Institutes of Health.

REFERENCES and FOOTNOTES

1. Kwong, P. D., R. Wyatt, J. Robinson, R. W. Sweet, J. Sodroski, and W. A. Hendrickson. 1998. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature*. 393:648–659.
2. Kwong, P. D., R. Wyatt, S. Majeed, J. Robinson, R. W. Sweet, J. Sodroski, and W. A. Hendrickson. 2000. Structures of HIV-1 gp120 envelope glycoproteins from laboratory-adapted and primary isolates. *Structure*. 8:1329–1339.
3. Chen, B., E. M. Vogan, H. Gong, J. J. Skehel, D. C. Wiley, and S. C. Harrison. 2005. Structure of an unliganded simian immunodeficiency virus gp120 core. *Nature*. 433:834–841.
4. Huang, C.-C., M. Tang, M. Y. Zhang, S. Majeed, E. Montabana, R. L. Stanfield, D. S. Dimitrov, B. Korber, J. Sodroski, I. A. Wilson, R. Wyatt, and P. D. Kwong. 2005. Structure of a V3-containing HIV-1 gp120 core. *Science*. 310:1025–1028.
5. Sirois, S., T. Sing, and K. C. Chou. 2005. HIV-1 gp120 V3 loop for structure-based drug design. *Curr. Protein Pept. Sci.* 6:413–422.
6. Konarev, P. V., V. V. Volkov, A. V. Sokolova, M. H. J. Koch, and D. I. Svergun. 2003. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* 36:1277–1282.
7. Svergun, D. I. 1992. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* 25:495–503.
8. Svergun, D. I., M. V. Petoukhov, and M. H. J. Koch. 2001. Determination of domain structure of proteins from x-ray solution scattering. *Biophys. J.* 80:2946–2953.
9. Kozin, M. B., and D. I. Svergun. 2001. Automated matching of high- and low-resolution structural models. *J. Appl. Crystallogr.* 34:33–41.
10. Garlick, R. L., R. J. Kirschner, F. M. Eckenrode, W. G. Tarpley, and C. S. Tomich. 1990. *Escherichia coli* expression, purification, and biological activity of a truncated soluble CD4. *AIDS Res. Hum. Retroviruses*. 6:465–479.