Ligand-binding Pocket in the Dengue Virus Envelope Glycoprotein

Y. Modis,[1] S. Ogata,[2] D. Clements,[2] S.C. Harrison[1]
[1]Howard Hughes Medical Institute, Children's Hospital, and Harvard Medical School, Boston, MA, U.S.A.
[2]Hawaii Biotech, Inc., Aiea, HI, U.S.A.

Introduction

Dengue virus, a member of the flavivirus family, imposes one of the largest social and economic burdens of any mosquito-borne viral pathogen [1]. There is no specific treatment for infection, and control of dengue virus by vaccination has proved elusive [2]. Several other flaviviruses are important human pathogens, including yellow fever, West Nile, tick-borne encephalitis (TBE), and Japanese encephalitis viruses (JE) [2].

Three structural proteins (C, M, and E) and a lipid bilayer package the positive-strand RNA genome of flaviviruses [3]. The core nucleocapsid protein, C, assembles with RNA on the cytosolic face of the endoplasmic-reticulum membrane. The assembling core buds through the ER membrane, thereby acquiring an envelope that contains the major envelope glycoprotein, E, and the so-called precursor membrane protein, PrM. The particle passes through the secretory pathway, where a furinlike protease cleaves PrM to M in a late trans-Golgi compartment. The cleavage, which removes most of the ectodomain of PrM, releases a constraint on E and primes the particle for low-pH-triggered membrane fusion. Uncleaved, immature particles are not fusion competent [2, 3].

E, which mediates both receptor binding [4] and fusion [5], is a so-called "class II" viral fusion protein [6, 7]. The more familiar class I fusion proteins, exemplified by the haemagglutinin (HA) of influenza virus and gp120/gp41 of HIV, have a "fusion peptide" at or near the N-terminus of an internal cleavage point [8]. This hydrophobic and glycine-rich segment, buried in the cleaved-primed trimer of the class I fusion protein, emerges when a large-scale conformational rearrangement is triggered by low pH (in the case of HA), receptor binding (in the case of gp120/gp41), or other cell-entry-related signals. The likely sequence of events that follows includes an interaction of the fusion peptide with the target-cell membrane and a refolding of the trimer. The latter step brings together the fusion peptide and viral-membrane anchor, thereby drawing together the cellular and viral membranes and initiating the bilayer fusion process [6]. The class II proteins, found so far in flaviviruses and alphaviruses, have evolved a structurally different but mechanistically related fusion architecture [6, 7]. As in class I proteins, a proteolytic cleavage (of PrM to M in flaviviruses, or of pE2 to E2 in alphaviruses) yields mature virions, with the fusion proteins in a metastable conformation, primed for fusion. The fusion peptide, an internal loop at the tip of an elongated subdomain of the protein [5], is buried at a protein interface and becomes exposed in the conformational change initiated by exposure to low pH [9, 10]. Because only the prefusion structures of one flaviviral and one alphaviral envelope protein have been determined previously, we know rather little about the conformational rearrangements set in motion by exposure to low pH (in the early endosome following viral uptake). The structures do suggest that the conformational changes involve hinge motions about interdomain linkages [9], together with oligomeric rearrangements on the viral surface [11-13]. In the case of the flaviviruses, the E dimers found on the surface of the virion recluster irreversibly into trimers when exposed to pHs of <6.3 [11].

Methods and Materials

E protein from dengue virus type 2 S1 strain was supplied by Hawaii Biotechnology Group Inc. (Aiea, HI). The protein was expressed and purified as described by Ivy et al. [14]. Crystals grow from a 10-g/L solution at 4˚C by hanging drop vapor diffusion in 11% PEG 8000, 1 M sodium formate, 20% glycerol and 0.1 M HEPES (pH 8). The addition of 0.5% n-octyl-β-D-glucoside (β-OG) prior to crystallization significantly improved the abundance and diffraction limit of the crystals. Dimensions of the primitive hexagonal cell were approximately a = b = 81 Å and c = 287 Å, with two molecules per asymmetric unit.

Crystals were derivatized by being soaked in mother liquor containing 0.5 mM K_2PtCl_4, 0.5 mM $Yb_2(SO_4)_3$, 0.5 mM $KAu(CN)_2$, or 10 mM Me_3PbAc for 24 h. Data sets were collected at 100K on beamlines A1 and F1 of the Cornell High Energy Synchrotron Source (CHESS, Cornell University), except for the "Native 1" data set, which was collected on beamline ID-19 at the APS. The data were processed with HKL [15].

The data sets were scaled to the most isomorphous native data set, and isomorphous difference Pattersons were calculated. Two initial heavy atom sites were identified by using the lead derivative. Additional sites were located in the three other derivative data sets by using cross-difference Fourier maps. Initial phases were optimized by refining the heavy atom parameters against maximum likelihood targets. Phases were improved by solvent flattening and twofold noncrystallographic symmetry (NCS) averaging. The

space group was determined as $P3_121$. The atomic coordinates were refined against the best native dataset, "Native 1," first as a rigid body, then by simulated annealing using torsion angle dynamics. Further cycles also included restrained refinement of B-factors for individual atoms and energy minimization against maximum likelihood targets. The atomic model was completed by using $2F_o - F_c$ and $F_o - F_c$ Fourier maps. A total of 137 water molecules were added. The final model also includes two glycans, and it includes one molecule of β-OG per protein molecule.

The structure of dengue E in the absence of β-OG was determined by refining the atomic coordinates against a data set collected from a crystal grown in the absence of β-OG. The kl hairpin (residues 270-279) and residues 165-169 were completely rebuilt.

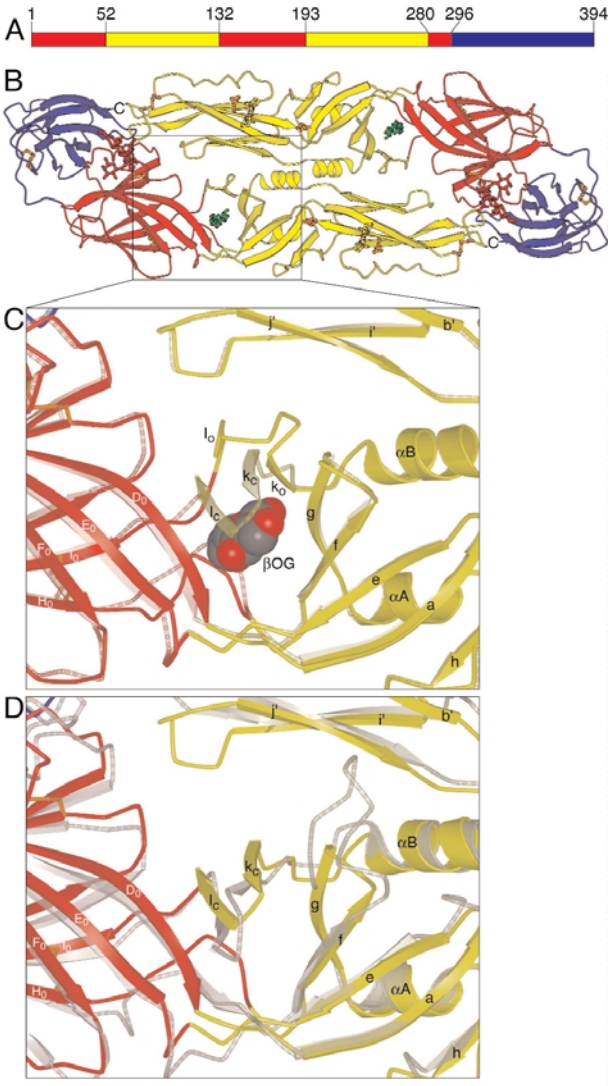

FIG. 1. Dengue E protein and its ligand-binding pocket. A shows domain definition of dengue E. Domain I is red, domain II is yellow, and domain III is blue. B shows the dengue E protein dimer, colored as in A, in complex with n-octyl-β-D-glucoside (β-OG). The β-OG (green) is bound in a hydrophobic pocket under the kl hairpin. A putative receptor-binding loop in domain III (residues 382-385) is marked with a triangle. C is an enlargement of the kl hairpin region, with the structure of dengue E in the absence of β-OG (in grey) superimposed. The strands of the kl hairpin are labeled with "o" or "c" subscripts for the open (β-OG-bound) and closed forms, respectively. D shows a superposition of the structures of dengue E and TBE E (grey), both in the absence of β-OG.

Results and Discussion

Molecular Architecture of Dengue E Dimer

Figure 1B shows the three-domain structure of the dengue virus sE dimer. Domain I, the eight-stranded central β-barrel, organizes the structure. Its β-strands are denoted B_0-I_0, with the addition of a short amino-terminal strand (A_0) parallel to strand C_0 at one edge of the barrel. Insertions between strands D_0 and E_0 and strands H_0 and I_0 form the elongated domain II, which bears the fusion peptide at its tip (Fig. 1B). Domain II contains 12 β-strands, denoted a-l, and two α-helices, αA and αB. Domain III is an IgC-like module, with 10 β-strands (A-G, plus a small extra sheet $A_xC_xD_x$ [9]). In all three domains, β-strands predominate. As expected from the 37% sequence identity between dengue and TBE sE, each domain of dengue sE has the same folded structure as its TBE counterpart, but several loops diverge in conformation. The relative domain orientations are also slightly different, consistent with the notion that the links between them might be flexible.

One consistent difference between E proteins from tick-borne and mosquito-borne flaviviruses is the presence, in the latter, of an additional four residues (382-385) between strands F and G of domain III. In our structure, these residues form a compact solvent-exposed bulge (Figs. 1B and 2A). Their relatively high temperature factors suggest some degree of flexibility. This loop has been implicated in receptor binding in dengue virus [4].

There are two glycosylated asparagines on each dengue E subunit: Asn 153 on domain I, and Asn 67 on domain II. Asn153, conserved in most flavivirus envelope proteins, bears a structure modeled here as a tetrasaccharide, although it contains additional, poorly ordered sugars. The fourth sugar is a mannose, which appears to be important for viral entry [16]. The glycan projects outward from the surface of the protein, and somewhat discontinuous electron-density features suggest that it makes a crystal contact with the Asn 67 glycan of another sE dimer.

Ligand-binding Pocket

The most significant difference between the

structures of dengue sE with and without β-OG is an altered conformation of the kl loop, which shifts toward the dimer contact in the presence of the detergent, forming a salt bridge and a hydrogen bond with i and j of the dimer partner. To effect this movement, strands k and l switch sheets, from $F_0E_0D_0$lk to efgkl (Fig. 1C). The shift closes the "holes" along the dimer contact to either side of the twofold axis and opens a tapering, hydrophobic channel at the interface between domains I and II. This channel accepts a single β-OG molecule. The glucosyl head group of β-OG lies at the channel's mouth, with several hydrogen bonds fixing a well-ordered orientation; the hydrocarbon chain projects well into the channel's cavity. In TBEV sE, which was studied in the absence of β-OG, the kl loop is in the "closed" position, and the hydrophobic residues are buried (Fig. 1D).

Mutations of residues that participate in the domain I/II interface just described alter the threshold pH for fusion (Fig. 2). Most of them involve side chains in the β-OG binding pocket. We take this correlation as a strong indication that domains I and II indeed change orientation during the fusion-promoting conformational change. We propose that the opening of the kl hairpin pries open the hydrophobic interface, causing domain II to hinge away from its dimer partner and to project the fusion peptide at its tip toward the membrane of the target cell. Two crystallographic observations are consistent with such a hinge. In two different crystal forms of dengue sE obtained in the absence of β-OG, domain II shifts by about 5° with respect to domain I. The same is true for two different crystal forms of TBE sE. In both cases, the hinge angle is quite small, because a larger bend would disrupt the dimer contact at the tip of domain II and expose the fusion peptide. Indeed, it is just such a disruption that occurs at low pH.

In the pH-threshold mutations, substitution of longer hydrophobic side chains by shorter ones generally lowers the maximum pH that triggers fusion (Fig. 2). We suggest that shorter side chains may allow a tighter and more stable closed form of the pocket, requiring a greater drop in pH to flip it open. Attenuated viruses with single mutations in the kl hairpin region have been obtained by passage in cell culture [17, 18]. Accumulation of such mutations might result in even stronger attenuation.

Implications for Viral Assembly and Fusion

The outer surfaces of mature flavivirus particles contain 180 subunits each of E and M, in a compactly organized icosahedral array [3]. Any conformational change in E is therefore likely to induce a concerted reorganization across the entire surface of the virion. The E proteins cluster into trimers when they undergo their conformational change induced by low pH [11]. We do not yet know which domains contribute to the trimer contacts. On the basis of image reconstructions from electron cryomicroscopy of fusion-competent TBE recombinant subviral particles, which contain 60 subunits each of E and M [12], we propose that a hinge motion of domain II away from its dimer partner during the low-pH-induced transition could result in the formation of trimer contacts by domain II, with only a modest reorientation of domains I and III within the surface lattice. The resulting trimer would display three fusion-peptide loops at its tip.

FIG. 2. Mutations affecting the pH threshold of fusion in flaviviruses. The mutated residues line the interior of the ligand-binding pocket. For uncovered residues, the residue type in the virus in which the mutation was identified is listed first, followed by the residue type in dengue 2. The coloring code is the same as that for Fig. 1.

In conclusion, we have identified the kl hairpin as a key structural element for initiating the low-pH conformational change that leads to formation of fusion-competent trimers. The opening up of a ligand-binding pocket just at the locus of a likely hinge suggests that compounds inserted at this position might hinder further conformational change and hence inhibit the fusion transition. In the context of the virion surface, their action might resemble that of some of the well-studied anti-picornaviral compounds, which block a concerted structural transition in the icosahedral assembly [19]. Alternatively, small molecules that pry open the kl hairpin on binding in the ligand-binding pocket may inhibit infection by facilitating the low-pH conformational change, causing premature triggering. Our structural observations suggest direct ways to search for such inhibitors.

Acknowledgments

We thank G. Gao, W. Clemons, and staff at beamlines F1 and A1 at CHESS (Cornell University)

and beamline ID-19 (APS). This work was supported by a long-term fellowship to Modis from the Human Frontier Science Program Organization and by National Institutes of Health Grant CA13202 to Harrison. Harrison is a Howard Hughes Medical Institute Investigator. Use of the APS was supported by the DOE Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-ENG-38.

References

[1] D.J. Gubler, Trends Microbiol. **10**, 100-103 (2002).
[2] D.S. Burke and T.P. Monath, in Fields Virology, edited by D.M. Knipe and P.M. Howley (Lippincott Williams and Wilkins, Philadelphia, PA), pp. 1043-1125 (2001).
[3] B.D. Lindenbach and C.M. Rice, in Fields Virology, edited by D.M. Knipe and P.M. Howley (Lippincott Williams and Wilkins, Philadelphia, PA), pp. 991-1041 (2001).
[4] W.D. Crill and J.T. Roehrig, J. Virol. **75**, 7769-7773 (2001).
[5] S.L. Allison, J. Schalich, K. Stiasny, C.W. Mandl, and F.X. Heinz, J. Virol. **75**, 4268-4275 (2001).
[6] W. Weissenhorn, A. Dessen, L.J. Calder, S.C. Harrison, J.J. Skehel, and D.C. Wiley, Mol. Membr. Biol. **16**, 3-9 (1999).
[7] F.X. Heinz and S.L. Allison, Curr. Opin. Microbiol. **4**, 450-455 (2001).
[8] J.J. Skehel and D.C. Wiley, Annu. Rev. Biochem. **69**, 531-569 (2000).
[9] F.A. Rey, F.X. Heinz, C. Mandl, C. Kunz, and S.C. Harrison, Nature **375**, 291-298 (1995).
[10] J. Lescar, A. Roussel, M.W. Wien, J. Navaza, S.D. Fuller, G. Wengler, and F.A. Rey, Cell **105**, 137-148 (2001).
[11] S.L. Allison, J. Schalich, K. Stiasny, C.W. Mandl, C. Kunz, and F.X. Heinz, J. Virol. **69**, 695-700 (1995).
[12] I. Ferlenghi, M. Clarke, T. Ruttan, S.L. Allison, J. Schalich, F.X. Heinz, S.C. Harrison, F.A. Rey, and S.D. Fuller, Mol. Cells 7, 593-602 (2001).
[13] R.J. Kuhn, W. Zhang, M.G. Rossmann, S.V. Pletnev, J. Corver, E. Lenches, C.T. Jones, S. Mukhopadhyay, P.R. Chipman, E.G. Strauss, T.S. Baker, and J.H. Strauss, Cell **108**, 717-725 (2002).
[14] J. Ivy, E. Nakano, and D. Clements, in United States Patent and Trademark Office (Hawaii Biotechnology Group, Inc., Aiea, HI, 1997).
[15] Z. Otwinowski and W. Minor, Methods Enzymol. **276**, 307-326 (1997).
[16] S.L. Hung, P.L. Lee, H.W. Chen, L.K. Chen, C.L. Kao, and C.C. King, Virology **257**, 156-167 (1999).
[17] E. Lee, R.C. Weir, and L. Dalgarno, Virology **232**, 281-290 (1997).
[18] T.P. Monath, J. Arroyo, I. Levenbook, Z.X. Zhang, J. Catalan, K. Draper, and F. Guirakhoo, J. Virol. **76**, 1932-1943 (2002).
[19] T.J. Smith, M.J. Kremer, M. Luo, G. Vriend, E. Arnold, G. Kamer, M.G. Rossmann, M.A. McKinlay, G.D. Diana, and M.J. Otto, Science **233**, 1286-1293 (1986).
[20] Y. Modis, S. Ogata, D. Clements and S.C. Harrison, A ligand-binding pocket in the dengue virus envelope glycoprotein. Proc. Natl. Acad. Sci. U.S.A., **100**, 6986-6991 (2003).