Reprinted from Nature, Vol. 311, No. 5987, pp. 678–680, 18 October 1984
© *Macmillan Journals Ltd., 1984*

#29

Three-dimensional structure of an antigenic mutant of the influenza virus haemagglutinin

M. Knossow*‡, R. S. Daniels†, A. R. Douglas†, J. J. Skehel† & D. C. Wiley*

* Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, 02138, USA
† National Institute of Medical Research, Mill Hill, London NW7 1AA, UK

Antigenic variation in the haemagglutinin (HA) glycoprotein of influenza virus is associated with recurrent epidemics of respiratory disease in man (for review see ref. 1). We have examined the size of structural changes necessary to alter the antigenicity of HA by determining the three-dimensional structure of the HA from an antigenic mutant containing a single amino acid substitution which was selected by growth of virus in the presence of monoclonal antibodies. Here we present evidence that the simple addition of an amino acid side chain which results in only minor local distortions of the structure of the HA is sufficient structural alteration for a virus to escape neutralization by a monoclonal antibody. Our results also demonstrate that single amino acid substitutions can cause only local changes in the HA structure, verifying the assumption made in several studies to locate antigenic sites on the HA[2–5] and other molecules[6,7], and indicate that proposals[8,9] of large conformational changes to account for variations in HA antigenicity are unnecessary in this case. The structure of the variant antigen has independently been successfully predicted (M. Karplus, personal communication).

An antigenic variant, V3a, was selected by growing the parent (X31) virus[10] in the presence of a neutralizing monoclonal antibody, HC3a, as described previously[11,12]. Table 1 shows that while haemagglutination by the parent virus is inhibited by HC3a ascitic fluid at a dilution of 1:12,800, the variant virus V3a is not inhibited at a 1:100 dilution of the fluid. A similar result was obtained in enzyme-linked immunosorbent antibody binding assays, in which parent virus and V3a mutant virus bound antibodies at ascitic fluid dilutions of 1:100,000 and 1:100, respectively.

The amino acid sequence of the entire V3a HA was determined by sequencing the viral RNA using methods described previously[13,17,18]; only one nucleotide change was observed, G to A, at nucleotide 514 which, when translated into amino acids, established that V3a has an aspartic acid at position 146 in the HA1 polypeptide instead of the glycine in the parental sequence[19].

Isomorphous crystals of V3a were grown under conditions in which the X31 HA crystallizes[14]. (This has been observed for all of four single amino acid variants tested to date.) We collected 76,840 unique X-ray reflections to 3.0 Å resolution on 1° oscillation photographs recorded at 4 °C by methods described previously[15]. Partially recorded reflections (≥50%) were included in the data set after adjustment by the post-refinement procedure[16]. The symmetry R factor (see Fig. 2 legend) on fully recorded reflections is 0.11 and on all reflections more than half recorded, 0.14. The parent and variant structure factors were scaled together using a Wilson statistics curve (R scale $= (F_{obs\,V3} - F_{obs\,X31})/(F_{obs\,X31}) = 0.161$) and a difference Fourier was calculated to 3 Å resolution using phases from the refined model (see Fig. 2 legend for details).

The difference Fourier averaged about the molecular 3-fold axis shows nine peaks above the 5σ level. The largest peak is 12σ and is at the location expected for an aspartic acid side chain if it were added to Gly 146 (Fig. 2*b*); the next highest peak is $<6\sigma$. Similar results with fewer noise peaks were observed on a difference Fourier calculated with a shell of data from 10 to 6 Å resolution.

The coordinates of the aspartic acid were obtained from model building on a $2F_0 - F_c$ map after refining native coordinates against V3a data. The O^{δ} of the carboxylate of Asp 146 appears within 3.0 Å of the N^{ε} of Arg 141, forming an apparent hydrogen bond. No other distortion of the external loop (amino acids 139–147) or of other features of the protein are observed, with the exception of a small (less than 1 Å) adjustment in the position of the side chain of arginine 141, presumably to accommodate the hydrogen bonding geometry required by the aspartic acid.

The structure of the variant HA suggests two possible mechanisms for the decreased affinity of the monoclonal antibody for the mutant HA: (1) The difference in the properties (shape, polarity and/or charge) of the aspartic acid and glycine side chains may directly affect the binding affinity. (2) The aspartic acid 146–arginine 141 interaction may stabilize the local structure of the site to which the monoclonal antibody binds, which would decrease the binding affinity if the antibody needed to locally distort or denature the site for tight binding. The former possibility corresponds to the classical 'static' view of an antibody–antigen complex, while the latter would apply if the free energy decrease on binding to a distorted site were greater than that required to distort the site. Present knowledge about the state of an antigen when 'seen' by immunogenic lymphocytes during the generation of an immune response and about the structure of an antigen–antibody complex leaves both possibilities feasible. In the case of V3a, the electrostatic charge of the Asp side chain would lead to unfavourable charge interactions (the burying of an unpaired charge) in complexes with antibodies against the parent (Gly).

Although multiple amino acid substitutions in the HA appear to be necessary for a virus to acquire the potential to cause an epidemic[2], single amino acid substitutions in the HA are sufficient for a virus to escape neutralization by monoclonal antibodies[11]. The three-dimensional structure of the HA of an antigenic variant presented here indicates that the substitution of an aspartic acid for a glycine, without any significant changes in the positions of the atoms in the rest of the HA molecule, is a sufficient alteration to reduce by three orders of magnitude the binding of a monoclonal antibody. The absence of other structural changes in the variant HA, in turn, proves that the monoclonal antibody used to select this variant must bind directly to the region of the HA surface designated site A by Wiley *et al.*[2].

Since submitting this manuscript we have found that H. H.-L. Shih, J. Brady and M. Karplus (unpublished) have independently and with no prior knowledge of our results successfully predicted the structural changes in the mutant V3a, relative to the parent protein. Using a perturbation approach making use of the empirical energy function program CHARMM[23], they calculated a conformational potential energy surface[24] for the

Table 1 Haemagglutination-inhibition reactions of X31 virus and antigenic mutant V3a

Virus	Monoclonal antibody HC3a	HC31	HC101	HC83	HC263
X31	12,800	6,400	12,800	6,400	25,600
V3a	>100	6,400	12,800	6,400	25,600

Haemagglutinin-inhibition reactions were done by standard procedures. The values shown are the reciprocals of the highest dilutions of the antibodies at which haemagglutination was inhibited. The specificities of the different monoclonal antibodies are defined by the amino acid substitutions observed in antigenic variants of X31 virus that the individual antibodies select. Thus, antibody HC3a selects mutants with amino acid substitutions at residue 146; HC31 at residue 198; HC101 at residue 63; HC83 at residue 193; and HC263 at residue 158. All monoclonal antibodies tested, other than HC3a which was used to select V3a, react with X31 and V3a equally.

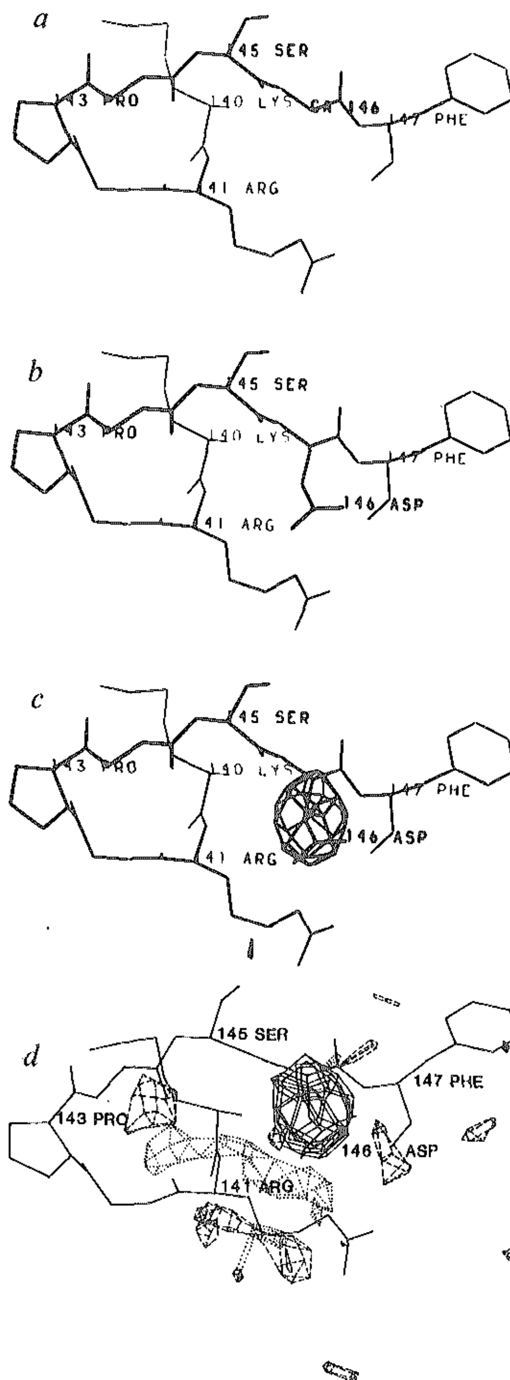
‡ Present address: Laboratoire de Physique, Centre Pharmaceutique, 92290 Chatenay-Malabry, France.

Fig. 1 A schematic diagram of a monomer of the haemagglutinin from the 1968 Hong Kong influenza virus[15], showing the location of the single mutation, Gly to Asp 146.

Fig. 2 *a*, Structure of the amino acid 140–147 loop which contains a glycine at position 146 in the parent HA. Atomic coordinates for the parent structure[15] were obtained by model building (I. A. Wilson and D.C.W.) on a version of the BILDER program[20] modified by Robert Ladner. These were refined to an R factor of 0.26 after six cycles of restrained least-squares refinement[21] where the contribution of the structure factor to the least-squares equation was evaluated by the gradient-curvature method on a difference map (ref. 22 and M. Lewis and D. Rees, personal communication). ($R = \Sigma|siI_{hi} - \bar{I}_h|/\Sigma\bar{I}_h$) where si = scale factor (including film temperature factor) for crystal i, and I_{hi} = intensity of reflection h of crystal i.) Exact 3-fold symmetry was imposed on the HA trimer in the crystallographic asymmetric unit by refining only a monomer's coordinates from 3-fold averaged difference maps. An overall B factor of 14 was used. Details of the refinement will be published elsewhere. *b*, The coordinates of the 140–147 loop of the antigenic variant V3a showing the location of the Asp (for Gly) substitution. Note that no conformational changes are observed between *a* and *b*. The V3a loop coordinates were obtained from model building on a $2F_0 - F_c$ map after refining the X31 coordinates (minus the loop) against the V3a data. *c*, The V3a difference map (F_{obs} V3a $- F_{obs\ X31}$) showing the 12σ peak at Asp 146. The only significant difference is at the position expected by the addition of the Asp 146 side chain instead of Gly 146 in the wild type. The difference Fourier was calculated using observed parent and variant structure factors with the refined phases described for *a*. *d*, The same as *c* with additional positive (dashed) and negative (dotted) contours at the noise level, 3σ. The interpretable 3σ peaks are a ridge of difference density indicating a small shift in the position of the side chain of Arg 141, which is within hydrogen bonding distance of Asp 146. (This figure was produced with an Evans and Sutherland PS-300 on the version of FRODO written by J. W. Pflugrath, M. A. Saper, B. Bush and A. Jones.)

aspartic acid side chain in the presence of the rest of the protein. Energy refinement[23] of the low energy conformations on this surface yielded side chain torsion angles for Asp 146 with ±20° of the X-ray results. Further, the prediction of an adjustment in the position of Arg 141 to improve the hydrogen bond to the carboxyl group of Asp 146 led to our finding its displacement in the experimental difference map. Details of the method and results will be given separately.

We thank David Stevens and Rose Gonsalves for assistance, Eugene Brown of Genetic Institute for synthesizing one of the oligonucleotide primers, and acknowledge support from NIH AI 13654 (D.C.W.). M.K. was the recipient of fellowship DRG-542 from the Damon Runyon–Walter Winchell Cancer Fund. We also thank Henry H.-L. Shih, John Brady and Martin Karplus for allowing us to quote from their unpublished results predicting the structure of the single-site mutant



Received 25 April; accepted 14 June 1984.

1. Webster, R. G. & Laver, W. G. in *Influenza Viruses and Influenza* (ed. Kilbourne, E. D.) 270–314 (Academic, New York, 1975).
2. Wiley, D. C., Wilson, I. A. & Skehel, J. J. *Nature* **289**, 373–378 (1981).
3. Caton, J. J., Brownlee, C. C., Yewdell, C. A. & Gerhard, W. *Cell* **31**, 417–427 (1982).
4. Daniels, R. S., Douglas, A. R., Skehel, J. J. & Wiley, D. C. *J. gen. Virol.* **64**, 1657–1662 (1983).
5. Daniels, R. S. *et al.* in *The Origin of Pandemic Influenza Viruses* (eds Laver, G. W. & Chu, C. M.) (Elsevier, New York, in the press).
6. Colman, P. M., Varghese, J. N. & Laver, W. D. *Nature* **303**, 41–44 (1983).
7. Evans, D. M. A., Minor, P. D., Schild, G. S. & Almond, J. W. *Nature* **304**, 459–462 (1983).
8. Laver, W. G., Air, G. M. & Webster, R. G. in *Genetic Variation Among Influenza Viruses* (ed. Nayak, D. P.) 283–295 (Academic, New York, 1981).
9. Webster, R. G. & Laver, W. G. *Virology* **104**, 139–148 (1980).
10. *Bull. Wld Hlth Org.* **45**, 119–124 (1971).
11. Laver, W. G., Air, G. M., Dopheide, T. A. & Ward, C. W. *Nature* **283**, 454–457 (1980).
12. Skehel, J. J. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
13. Rogers, G. N. *et al. Nature* **304**, 76–78 (1983).
14. Wiley, D. C. & Skehel, J. J. *J. molec. Biol.* **112**, 343–347 (1977).
15. Wilson, I. A., Skehel, J. J. & Wiley, D. C. *Nature* **289**, 368–373 (1981).
16. Winkler, F. K., Schutt, C. E. & Harrison, S. C. *Acta crystallogr.* **A35**, 901–911 (1979).
17. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
18. Patel, T. P. *et al. Nucleic Acids Res.* **10**, 5605–5620 (1982).
19. Verhoeyen, M. *et al. Nature* **286**, 771–776 (1980).
20. Diamond, R. in *Biomolecular Structure, Conformation, Function and Evolution* Vol. 1 (ed. Srinivasan, R.) 567–588 (Pergamon, Oxford, 1980).
21. Hendrickson, W. A. & Konnert, J. in *Biomolecular Structure, Function, Conformation, and Evolution* Vol. 1 (ed. Srinivasan, R.) 43–47 (Pergamon, Oxford, 1980).
22. Cochran, W. *Acta crystallogr.* **4**, 408–411 (1951).
23. Brooks, B. *et al. J. comp. Chem.* **4**, 187–217 (1983).
24. Gelin, B. R. & Karplus, M. *Biochemistry* **18**, 1256–1268 (1979).

Printed in Great Britain by Macmillan Production Ltd., Basingstoke, Hampshire