Conformation and Function of the N-linked Glycan in the Adhesion Domain of Human CD2

Daniel F. Wyss, Johnathan S. Choi, Jing Li, Maria H. Knoppers, Kevin J. Willis, Antonio R. N. Arulanandam, Alex Smolyar, Ellis L. Reinherz, Gerhard Wagner*

The adhesion domain of human CD2 bears a single N-linked carbohydrate. The solution structure of a fragment of CD2 containing the covalently bound high-mannose N-glycan [-(*N*-acetylglucosamine)$_2$-(mannose)$_{5-8}$] was solved by nuclear magnetic resonance. The stem and two of three branches of the carbohydrate structure are well defined and the mobility of proximal glycan residues is restricted. Mutagenesis of all residues in the vicinity of the glycan suggests that the glycan is not a component of the CD2-CD58 interface; rather, the carbohydrate stabilizes the protein fold by counterbalancing an unfavorable clustering of five positive charges centered about lysine-61 of CD2.

Human CD2 is a cell surface glycoprotein present on T lymphocytes and natural killer cells and is important in mediating both cellular adhesion and signal transduction through interactions with its counterreceptor CD58 (*1*). Nuclear magnetic resonance (NMR) solution studies of the NH_2-terminal domain of rat (*2*) and human CD2 (*3*), and x-ray crystallographic studies of the complete extracellular region of rat (*4*) and human CD2 (*5*) have shown that the extracellular segment consists of two immunoglobulin superfamily (IgSF) domains: a nine-stranded NH_2-terminal V-set domain lacking the first half of strand A (Fig. 1A), and a seven-stranded membrane-proximal C2-set domain. The NH_2-terminal membrane-distal domain mediates the adhesion function of the molecule (*6*) by binding to its counterreceptor CD58 in humans (*7*) and CD48 in rodents (*8*). CD2, CD58, and CD48 are thought to have evolved from a common precursor involved in homophilic interactions (*9*). Mutational screening of the human CD2 surface indicated that the interface with CD58 is located at the GFCC′C″ face of the protein (Fig. 1A) and is a highly charged surface area (*10*, *11*). Interestingly, the same surface area identified as the receptor interface forms a homodimeric contact in the crystal structures of rat and human CD2 (*4*, *5*). Based on mutations of human CD58 (*12*) and complementary mutations of rat CD2 and rat CD48 (*13*), models for the interactions of human CD2 with human CD58 and rat CD2 with rat CD48 have recently been proposed, suggesting that in both cases the receptor-ligand interaction involves the major β sheet surface of each adhesion domain.

The NH_2-terminal adhesion domain in human CD2 contains a single consensus N-glycosylation site (Fig. 1A) that is not conserved in rat CD2. Thus, the significance of the N-glycan for human CD2 function has been uncertain. Recny *et al.* (*14*) reported that after treatment with peptide:N-glycosidase F (PNGase F), the resulting deglycosylated adhesion domain was no longer recognized by monoclonal antibodies (mAbs) to CD2 and failed to bind to its natural ligand, CD58, on sheep red blood cells (SRBCs), suggesting that N-

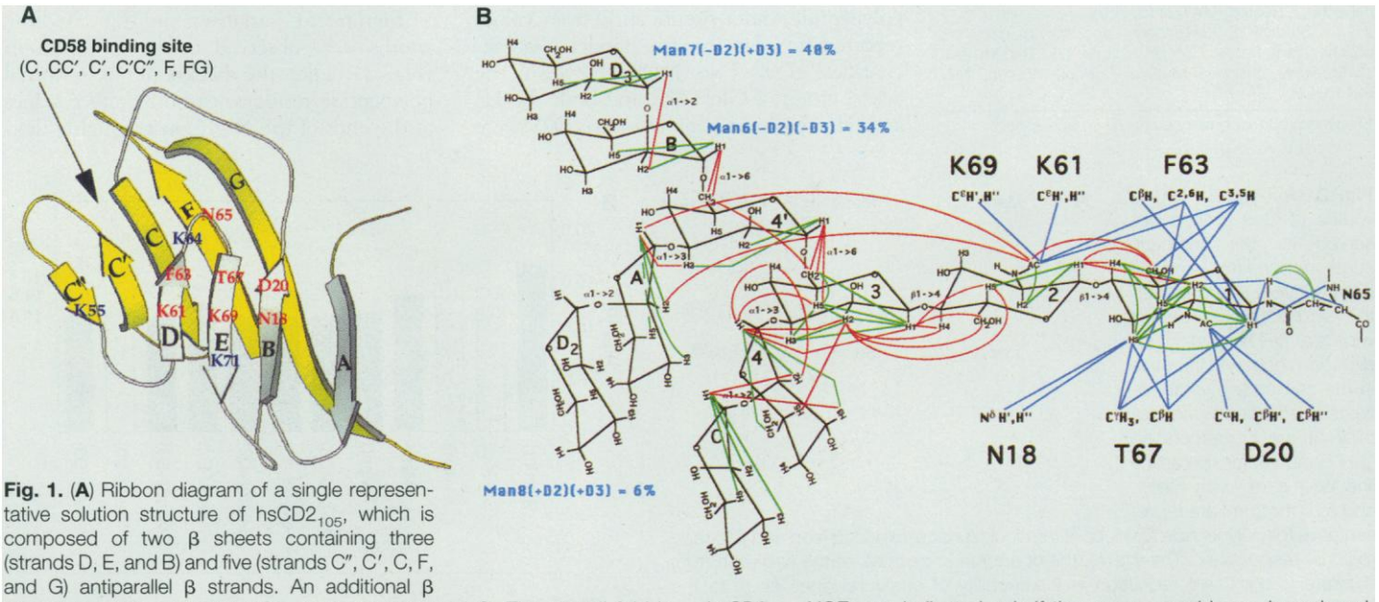

Fig. 1. (**A**) Ribbon diagram of a single representative solution structure of hsCD2$_{105}$, which is composed of two β sheets containing three (strands D, E, and B) and five (strands C″, C′, C, F, and G) antiparallel β strands. An additional β strand, A, makes parallel NOE contacts to strand G. The overall fold is characteristic of a V-set IgSF domain, but lacks the first half of strand A. A single high-mannose N-glycan is attached to the protein at Asn^{65} at the tip of the DE loop and is oriented toward the β sheet containing the strands D, E, and B. Polypeptide residues that contact the N-glycan are highlighted in red. The CD58 binding site, which includes residues in the C, C′, and F strands as well as in the CC″, C′C″ and FG′ loops, is located on the opposite face of the glycoprotein. Lys^{55}, Lys^{64}, and Lys^{71} are highlighted in purple, and together with Lys^{69} form a clustering of positively charged residues around Lys^{61}. The view is with the DEB face in front and the C″C′CFG face in the back [figure prepared with the MOLSCRIPT program (*29*)]. (**B**) Summary of the NOE data obtained for the high-mannose N-glycan of hsCD2$_{105}$. NOEs are indicated only if they were unambiguously assigned. More NOEs to carbohydrate protons are expected, but they were either not resolved in the NOESY spectra or could not be assigned because many of the H^2, H^3, H^4, H^5, and H^6 sugar resonances exhibit similar chemical shifts (*15*). Carbohydrate intraresidue and interresidue NOEs are shown by green and red lines, respectively, and carbohydrate-protein contacts are indicated by blue lines. Carbohydrate residues are labeled as in (*30*), and amino acid residues are abbreviated with the single-letter code (D, Asp; F, Phe; K, Lys; N, Asn; and T, Thr) and residue numbers. The composition and abundance of the glycomers present in our sample as determined by NMR and electrospray mass spectroscopy methods are also indicated. The exact composition of the residual 20% [-(GlcNAc)$_2$-(Man)$_5$] glycomers was not obtained by either method.

glycosylation is required for human CD2 adhesion function. Moreover, transmembrane CD2 variants with mutations in the consensus N-glycosylation sequence Asn^{65}-Gly^{66}-Thr^{67} [(N65Q (*14*) or T67A (*10*)] that preclude attachment of the high-mannose N-glycan at Asn^{65} could be normally expressed on cell surfaces, but showed neither antibody- nor ligand-binding activity. These data suggest that the N-linked adhesion domain glycan on human CD2 plays an important role in maintaining native receptor structure.

Analysis of a soluble version of the CD2 extracellular segment expressed in Chinese hamster ovary (CHO) cells (*3*, *14*) showed that the N-glycan in the adhesion domain of human CD2 ($hsCD2_{105}$) is heterogeneously occupied with high-mannose oligosaccharides [-$(GlcNAc)_2$-$(Man)_{5-8}$] (Fig. 1B). In addition to NMR assignments of the polypeptide (*3*), we recently achieved ^1H- and ^{13}C-NMR assignments for all the residues of the N-glycan in $hsCD2_{105}$ on the intact glycoprotein (*15*). In conjunction with electrospray mass spectroscopy, we determined the exact composition of the glycomers present in the heterogeneous N-glycan where [-$(GlcNAc)_2$-$(Man)_6$] and [-$(GlcNAc)_2$-$(Man)_7$] are the dominant forms of the glycan representing 34 and 40% of the molecules, respectively. We also identified carbohydrate-polypeptide nuclear Overhauser enhancements (Fig. 1B), which clearly define the carbohydrate's orientation with respect to the polypeptide. Here, we present the solution structure of glycosylated $hsCD2_{105}$ and describe the interaction of the N-glycan with the polypeptide. On the basis of this information, we used site-directed mutagenesis and additional biochemical data to obtain a better picture of the significance of the carbohydrate in the adhesion domain of human CD2.

We have previously determined the polypeptide structure of $hsCD2_{105}$ using predominantly conventional two-dimensional (2D) homonuclear experiments (*3*). We have refined the polypeptide structure of $hsCD2_{105}$ on the basis of heteronuclear 3D experiments on a ^{15}N-labeled sample and high quality homonuclear 2D nuclear Overhauser spectroscopy (NOESY) data obtained on a Varian UNITY*plus* 750 spectrometer (*16*). Thus, we identified many nuclear Overhauser effects (NOEs) within the carbohydrate and between the first two *N*-acetylglucosamine (GlcNAc) residues and the polypeptide residues Asn^{18}, Asp^{20}, Lys^{61}, Phe^{63}, Asn^{65}, Thr^{67}, and Lys^{69} (Fig. 1B). Many of the carbohydrate resonances, in particular those in contact with the polypeptide, differ significantly from values reported for identical free model oligosaccharides (*15*). Two NOEs between the acetyl group of GlcNAc2 and both the H^1 and H^2 resonances of ManA(−D2) were also observed, indicating that for a substantial time the arm with the Man4′ residue is folded toward the trisaccharide core. These NOEs were not detected for the ManA(+D2) residue, either because of the low concentration of this glycan form or because of a different conformation. Our findings are in good agreement with those based on a detailed comparison of spin-spin coupling and chemical shift data for a series of high-mannose oligosaccharides (*17*) and molecular dynamics simulations for a model of ribonuclease B (RNase B) containing a [-$(GlcNAc)_2$-$(Man)_5$] glycoform (*18*).

In solving the glycoprotein structure of $hsCD2_{105}$, we addressed the question of whether parts of the carbohydrate are tightly folded similar to the polypeptide. The NOEs we observed between the N-glycan and the protein could have resulted from transient encounters of the otherwise flexible carbohydrate with the polypeptide. Therefore, we measured the carbon line widths of C-H groups in the sugar and the polypeptide units in a highly resolved 2D ^1H-^{13}C spectrum recorded on a Varian UNITY*plus* 750 spectrometer (*19*). The line widths are directly related to mobility and provide a qualitative comparison of the relative flexibility of carbohydrate and polypeptide groups. Here, larger and smaller line widths indicate lower and higher internal mobility, respectively (*19*).

Significant variations in the ^{13}C line widths were observed in the glycoprotein (Fig. 2). Like the backbone of terminal polypeptide residues, carbohydrate residues at the ends of the N-glycan are highly flex-

D. F. Wyss, M. H. Knoppers, K. J. Willis, A. R. N. Arulanandam, Procept, Cambridge, MA 02139, USA.
J. S. Choi and G. Wagner, Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA.
J. Li, A. Smolyar, E. L. Reinherz, Laboratory of Immunobiology, Dana-Farber Cancer Institute, and Department of Medicine, Harvard Medical School, Boston, MA 02115, USA.
*To whom correspondence should be addressed.

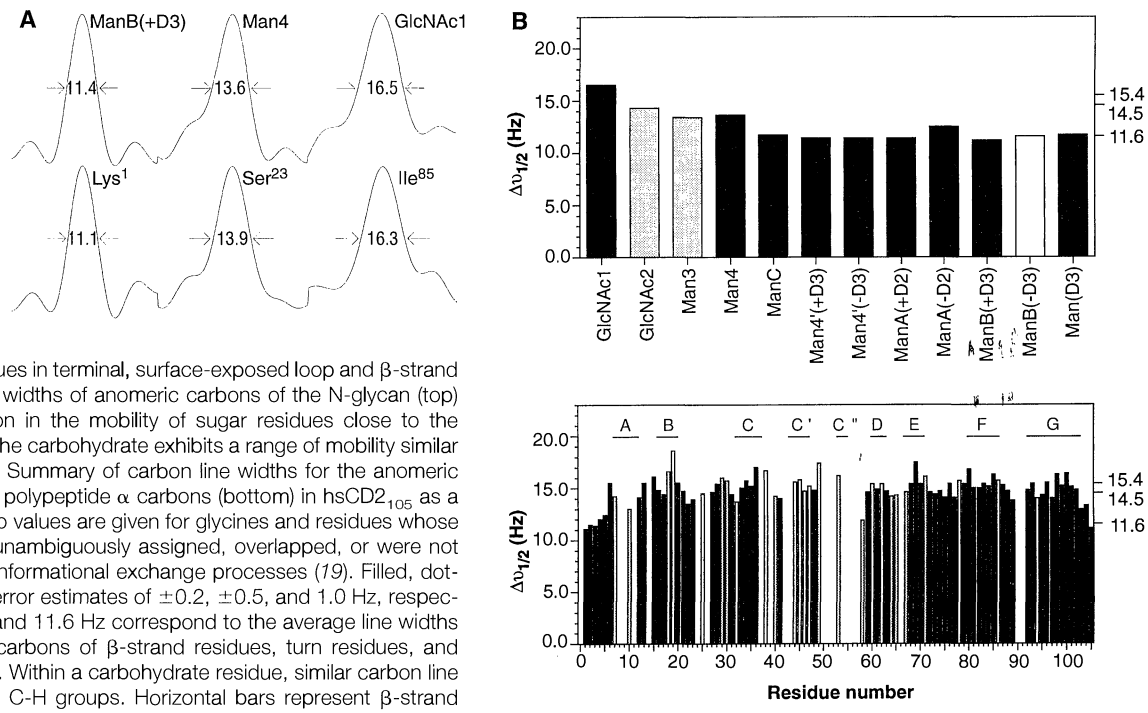

Fig. 2. (**A**) Typical ^{13}C line widths of C-H groups observed for the anomeric sugar carbons (top) and the polypeptide α carbons (bottom) in $hsCD2_{105}$. The line widths at half heights are indicated in hertz. An increase in the carbon line width reflects a reduction in the mobility of the corresponding C-H bond vector. α-carbon line widths of Lys^1, Ser^{23}, and Ile^{85} (bottom) are representative for polypeptide residues in terminal, surface-exposed loop and β-strand regions, respectively. The line widths of anomeric carbons of the N-glycan (top) illustrate a significant reduction in the mobility of sugar residues close to the N-glycosylation site. Notably, the carbohydrate exhibits a range of mobility similar to that of the polypeptide. (**B**) Summary of carbon line widths for the anomeric sugar carbons (top) and most polypeptide α carbons (bottom) in $hsCD2_{105}$ as a function of residue number. No values are given for glycines and residues whose cross peaks were either not unambiguously assigned, overlapped, or were not observed because of slow conformational exchange processes (*19*). Filled, dotted, and empty bars indicate error estimates of ±0.2, ±0.5, and 1.0 Hz, respectively. The values 15.4, 14.5, and 11.6 Hz correspond to the average line widths measured for polypeptide α carbons of β-strand residues, turn residues, and terminal residues, respectively. Within a carbohydrate residue, similar carbon line widths were observed for the C-H groups. Horizontal bars represent β-strand residues. β strands are labeled according to an IgSF V-set domain (Fig. 1A).

ible [Fig. 2B, compare residues 1 to 5 (bottom) with ManC and ManD3 (top)]. However, GlcNAc1, which shows many NOEs to the polypeptide, has a highly restricted mobility similar to that of α carbons of the protein core, whereas GlcNAc2 exhibits a mobility identical to the average mobility of α-carbons of the polypeptide. Man3 and Man4 have a mobility comparable to that of the backbone of surface-exposed loops of the polypeptide. ManA(−D2), which shows NOEs to GlcNAc2 (Fig. 1B) and hence is folded toward the trisaccharide core [-GlcNAc1-GlcNAc2-Man3-], has a slightly lower mobility than ManA(+D2), for which such NOEs were not observed.

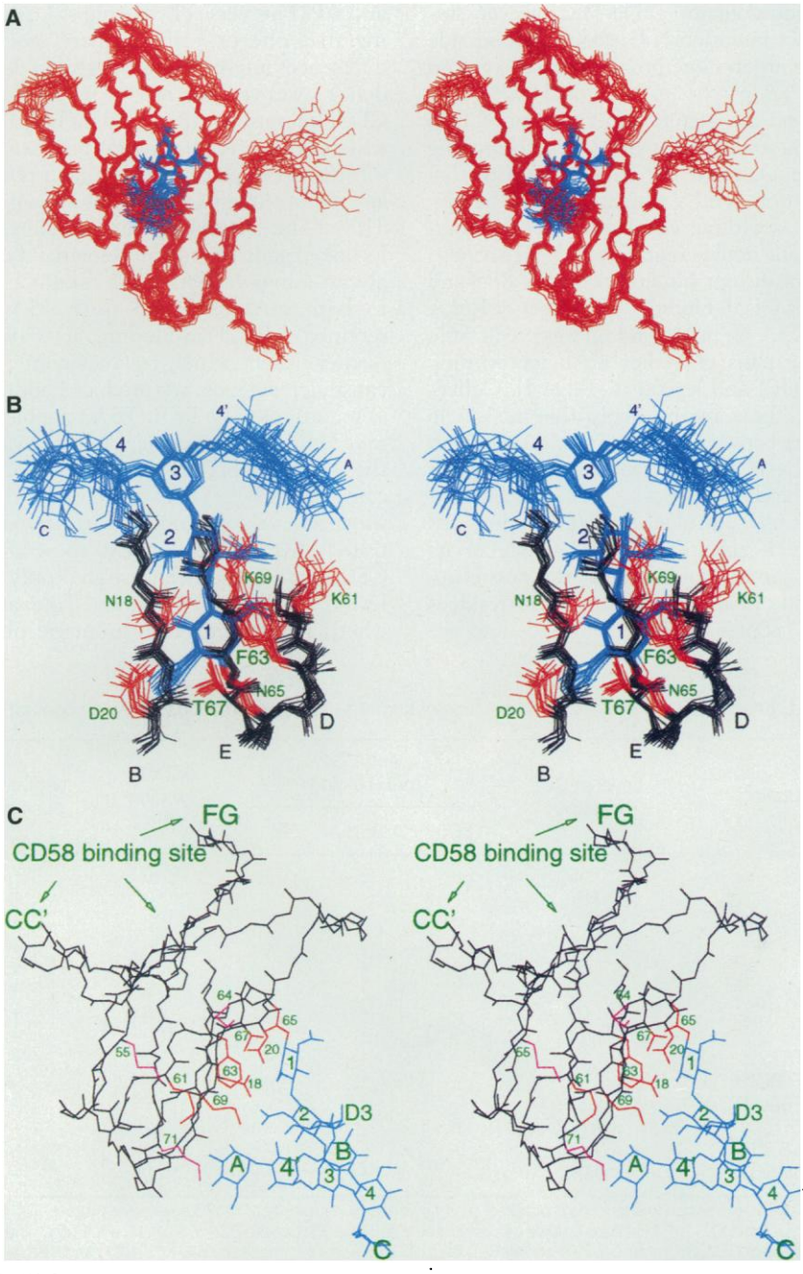

Fig. 3. (**A**) Stereo diagram of backbone atoms of the polypeptide (red) and heavy atoms of the trisaccharide core [-GlcNAc1-GlcNAc2-Man3-] (blue) for 16 NMR structures of $hsCD2_{105}$ superimposed on the NMR mean coordinates. The orientation is the same as in Fig. 1A. (**B**) DEB sheet and part of the high-mannose N-glycan of 16 model structures based on the superposition of backbone heavy atoms of all residues (16 to 20, 60 to 71) in the DEB sheet and heavy atoms of the trisaccharide core [-GlcNAc1-GlcNAc2-Man3-]. The backbone, the side chains for the contact residues, and the high-mannose N-glycan are shown in black, red, and blue, respectively. Glycoprotein structures were calculated for the [-$(GlcNAc)_2$-$(Man)_7$] glycomer (the glycomer shown in Fig. 1B minus ManD2), but for clarity the more disordered arm [-ManB-ManD3] is not shown. The sugar residues are annotated as in Fig. 1B. (**C**) Representative glycosylated model structure of $hsCD2_{105}$ showing the polypeptide backbone in black, the side chains of the polypeptide residues that contact the glycan in red, the side chains of the polypeptide residues that together with Lys^{69} form a clustering of positively charged residues around Lys^{61} in purple, and the high-mannose N-glycan {[-$(GlcNAc)_2$-$(Man)_7$] glycomer} in blue. This carbohydrate is located on the opposite site of the CD58 binding site, and a conformation in which its [4′-A] arm is oriented toward the trisaccharide core [-GlcNAc1-GlcNAc2-Man3-] is significantly populated.

In general, the side chain carbons of the polypeptide are more mobile than the α carbons (*19*). In particular, the γ carbons of leucines, which are further away from the polypeptide backbone than the β carbons of isoleucines and valines, exhibit significantly higher mobility than the α carbons of the same residue (*20*). All measured side chain carbons in loop regions of the polypeptide have a mobility similar to that of terminal-polypeptide α carbons and anomeric carbons at the ends of the N-glycan. Moreover, most of the side chain carbons in β strands showed line widths similar to that of the anomeric carbons of GlcNAc2, Man3, and Man4. Thus, the glycan exhibits a range of mobility similar to that of the polypeptide.

Structures of the intact glycoprotein composed of the [-$(GlcNAc)_2$-$(Man)_7$] glycomer (the glycomer shown in Fig. 1B minus ManD2) were computed by addition of intracarbohydrate and carbohydrate-polypeptide NOE distance restraints to the intrapolypeptide restraints and forcing the sugar residues into their lowest energy chair conformations (*21*). Figure 3A shows a stereo diagram of the polypeptide backbone atoms and the heavy atoms of the well-defined trisaccharide core for the 16 best model structures of $hsCD2_{105}$ superimposed on the NMR mean coordinates. The DEB sheet and the structured parts of the high-mannose N-glycan are shown superimposed in Fig. 3B. The degree of disorder observed in the calculated model structures is in good agreement with the mobility data (Fig. 2A). As a result of the interactions with the polypeptide, the carbohydrate core residues [-GlcNAc1-GlcNAc2-Man3-] and Man4 are well defined. In addition, ManA(−D2), which is folded toward the carbohydrate core, shows a lower degree of disorder than the outer branches of the carbohydrate. In particular, the [−ManB-ManD3] arm is largely disordered.

A representative model structure (Fig. 3C) illustrates the orientation of the glycan with respect to the protein. The solution structure reveals specific interactions between part of the N-glycan and several residues on the DEB surface opposite to the CD58 binding site, suggesting that the carbohydrate increases the rigidity and stability of the molecule rather than being directly involved in the contact with CD58. This structure shows that there is no direct in-

teraction between the carbohydrate and the proposed CD58 binding site on CD2 because the glycan is oriented toward the BED sheet of $hsCD2_{105}$, opposite to this site. Moreover, the [−Man4′-ManA(−D2)] arm of the N-glycan is folded toward the trisaccharide core [-GlcNAc1-GlcNAc2-Man3-]. This conformation might play an important role in N-linked oligosaccharide biosynthesis (*17*), processing, and transport (*22*).

To probe the contribution of the N-glycan to the stability of the glycoprotein, we studied both enzymatically treated CD2 (*23*) [endoglycosidase H (endo H)–treated $hsCD2_{105}$ (only GlcNAc1 retained) and PNGase F–treated $hsCD2_{105}$ (fully deglycosylated)] and a series of mutants of full-length human CD2 (Table 1) (*24*). Consistent with plasmon resonance analyses on endo H–treated two-domain human CD2 (hsCD2) (*25*), endo H–treated $hsCD2_{105}$ was functional, although anti-CD2 immunoreactivity and CD58-binding activity were present at a slightly reduced level (*20*). Moreover, the CD spectrum of endo H–treated $hsCD2_{105}$ showed only minor differences to the $hsCD2_{105}$ spectrum (Fig. 4). In contrast, fully deglycosylated $hsCD2_{105}$ bound neither antibodies to CD2 (anti-CD2) nor CD58 (*14*, *20*) and showed a CD spectrum markedly different from that of fully glycosylated $hsCD2_{105}$ (Fig. 4). It also showed a strong tendency to aggregate, consistent with observations by Davis *et al.* for the two-domain PNGase F–treated hsCD2 (*25*). These data suggest that complete deglycosylation of the adhesion domain of human CD2 results in a significant destabilization of both $hsCD2_{105}$ and hsCD2.

The analysis of CD2 mutants created by site-directed methods was in agreement with the biochemical data. Elimination of the N-linked glycan on the human CD2 adhesion domain by a single N65Q (*14*) or T67A (*10*) mutation resulted in a loss of CD58-binding activity as well as immunoreactivity of each of three mAbs tested directed at distinct epitopes on the native adhesion domain (Table 1). However, single Ala mutations of residues whose side chains are in close proximity to the glycan (Asn^{18}, Asp^{20}, Lys^{61}, Phe^{63}, and Lys^{69}) (Fig. 3B) had no significant effect on CD58 binding and did not disrupt mAb binding (Table 1). Similarly, removal of a methyl group from Thr^{67}, which contacts the glycan, was without effect in the T67S mutation. The double mutation F63A/T67S also did not disrupt binding of mAbs 8B5 and 35.1 or CD58 binding. The effect of K61A and F63A on mAb binding was predictable on the basis of earlier mAb footprinting data [(*10*) and legend to Table 1]. Collectively, these results imply that no single contact between the amino acid side chain and the glycan is crucial for CD58- and mAb-binding activity.

Rat CD2, unlike human CD2, does not require a glycan for maintenance of its stable protein fold. To test the possibility that one or more rat amino acid residues might confer glycan independence, we compared sequences of rat, human, mouse, and horse CD2 in the region surrounding the N-linked glycosylation site (*26*). Human CD2 differs from CD2 of all the other species at positions 61 (Lys versus Glu) and 63 (Phe versus Leu) (Fig. 5). Assuming that one or both of these sequence differences might determine glycan dependence, we constructed a triple human CD2 mutation (K61E/F63L/T67D) in which the corresponding rat residues were substituted in place of the human residues at each of these two positions as well as at Thr^{67}. The latter substitution deletes the N-linked addition signal. Remarkably, this glycan-minus human CD2 variant bound to both mAb 8B5 and mAb 35.1 and retained full CD58-binding activity. To further dissect which residues confer glycan independence, we produced additional mutations. The F63L/T67D double mutant highlights the critical role of Lys^{61} in the K61E/F63L/T67D triple mutant because the F63L/T67D molecule did not show a level of binding with any mAb tested that was equivalent to wild-type CD2 binding and demonstrated only low-level (<25%) CD58 binding. To examine whether exposure of an aromatic phenyl

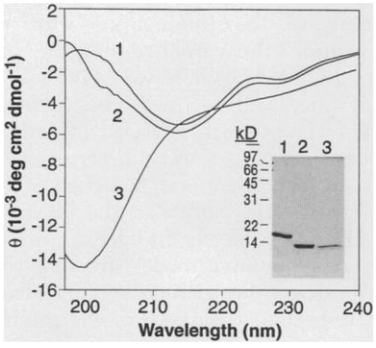

Fig. 4. Comparison of the CD spectra of $hsCD2_{105}$ (1) with endo H (2)– and PNGase F (3)–digested $hsCD2_{105}$. Samples were 10 μM in 10 mM NaP (pH 7), and the pathlength was 2 mm. (**Inset**) SDS–polyacrylamide gel electrophoresis gel (15%) stained with Coomassie. Lane 1, $hsCD2_{105}$; lane 2, endo H–digested $hsCD2_{105}$; and lane 3, PNGase F–digested $hsCD2_{105}$.

Table 1. Effects on CD2 structure and function of human CD2 point mutations at glycan contact sites.

Mutation	Level of CD2 expression*	Anti-CD2 mAb binding†			CD58 binding (%)‡	N-glycan at position 65
		8B5	35.1	$T11_2$		
N65Q	0.49	−	−	−	0	−
T67A	0.60	−	−	−	0	−
N18A§	1.22	+	+	+	103	+
D20A§	0.98	+	+	+	106	+
K61A§	1.30	+	+	−	135	+
F63A	1.20	+	+	−	75	+
T67S	1.02	+	+	+	88	+
K69A§	1.14	+	+	+	109	+
F63A/T67S	1.37	+	+	−	75	+
K61E/F63L/T67D	1.42	+	±	−	95	−
K61E/F63L/T67A	1.12	+	±	−	137	−
F63L/T67D	1.73	±	±	−	22	−
F63A/N65Q	0.44	−	−	−	0	−
K64A/N65Q	0.92	−	−	−	0	−

*Level of CD2 expression is shown as the ratio of mean fluorescence intensity of M32B staining of mutant CD2 to that of wild-type (wt) CD2. †Previous studies demonstrated that 3T4-8B5 maps to K34, R48, K91, and E95 on the CD58 binding face of CD2; 35.1 maps to K64 on the BED face of CD2; and $T11_2$ maps to F54, K55, K61, and K64, a region orthogonal to the GFCC′C″ face and extended to the BED surface (*10*). ‡CD58 binding is expressed as the percentage of the total SRBC rosette number for each mutant relative to that for wt CD2. The binding of each anti-CD2 mAb on mutant CD2 molecules was compared as the mutant CD2:wt CD2 ratio of the mean fluorescence intensities of mAb binding to M32B binding (*10*). Ratio <0.2 (−), 0.2 to 0.5 (±), and >0.5 (+). §Mutations are from (*10*).

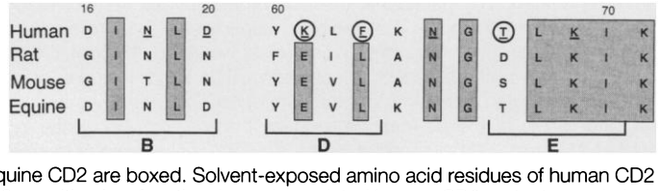

Fig. 5. Comparison of CD2 adhesion domain sequences of different species in the region of the N-linked glycan attachment site in human CD2. Residues conserved in rat, mouse, and equine CD2 are boxed. Solvent-exposed amino acid residues of human CD2 that were mutated to rat CD2, resulting in a functional nonglycosylated human CD2 mutant, are circled. Amino acid residues that contact the N-glycan are underlined.

side chain at Phe[63] might confer structural instability in the absence of the glycan, we generated the F63A/N65Q mutation. This mutant failed to bind CD58 or any mAb on the panel, thereby excluding the possibility that shielding from solvent of the side chain of Phe[63] by the glycan is important for protein stability. Another explanation for the stabilization effect of K61E/F63L/T67D might be that the incorporated T67D change could result in formation of a hydrogen bond to Asn[65], thereby stabilizing a particular DE loop conformation. However, the K61E/F63L/T67A triple mutant shows that this is not the case because its functional properties are virtually identical to those of K61E/F63L/T67D. A Lys at position 64 in the DE loop is present in human and equine CD2, whereas an Ala is present in the corresponding position of mouse and rat CD2. Nevertheless, Lys[64] alone has no direct role in stabilization because the K64A/N65Q mutation does not rescue CD2 structure or function.

Collectively, our results suggest that the N-glycan at Asn[65] in the adhesion domain of human CD2 is required for adhesion functions by stabilizing the polypeptide conformation. Trimming the N-glycan to a single GlcNAc residue reduces the stability of the glycoprotein but does not crucially destabilize the protein. In contrast, complete removal of the sugar leads to partial or full unfolding of the protein possibly followed by aggregation. However, substitution of Lys[61] with a glutamic acid eliminates the need for the carbohydrate to stabilize the protein. This results in a consistent picture for the role of the carbohydrate. Human CD2 contains an accumulation of five surface-exposed lysines on the DEB face of the protein. Lys[61] is situated in the center of this patch surrounded by Lys[69], Lys[71], Lys[64], and Lys[55] (Fig. 3C). This grouping of positive charges is energetically unfavorable. The glycan can stabilize the protein conformation through hydrogen bonds, van der Waals contacts, and possibly entropic contributions and thus counterbalances the clustering of the positive charges. If the central Lys[61] is replaced with a negatively charged glutamic acid, it may form a salt bridge with Lys[69] analogous to an interaction of the corresponding side chains that occurs naturally in rat CD2 (*2*, *4*) where the protein is stable even without the carbohydrate. It thus appears that the unfavorable charge at position 61 of human CD2 is responsible for its glycan requirement. These data further show the marked functional differences among highly structurally related CD2 molecules of different species that result from sequence changes during evolution. Given the distinctive effects of point mutations on protein stability, at least certain Ig domains such as CD2 do not appear to be rigid protein scaffolds, but rather appear to maintain their stable conformation as a result of dynamic interactions between the polypeptide and its attached glycan. On the basis of hydrogen exchange NMR measurements, the presence of a carbohydrate in RNase B has been observed to globally stabilize the polypeptide compared to the glycan-free RNase A (*27*).

We have solved the solution structure of a glycoprotein for which both the conformations of the carbohydrate and the polypeptide were defined. For comparison, only a few x-ray structures of fully glycosylated glycoproteins were obtained. In these structures, usually only the first few sugar residues are defined by proper electron density, presumably because of heterogeneity and increased disorder of sugar residues in the outer branches. Two notable exceptions (*28*) are the longer fragments of carbohydrates resolved in the x-ray structures of (i) an IgG CH2 domain structure (2.8 Å resolution), where the complex oligosaccharides on each heavy chain occupy the space between the domains and by their interaction stabilize a particular hinge conformation; and (ii) the structure of a legume lectin (2.0 Å resolution) in which the N-linked carbohydrate connects two proteins in the crystal.

Our findings indicate that through the interaction with the polypeptide, large parts of the attached glycan can be conformationally restricted, as compared with free oligosaccharides with similar or identical composition. The structural detail herein should prove helpful for further analysis of other glycoproteins, including those where the glycan moiety, unlike that of CD2, is part of a recognition site.

REFERENCES AND NOTES

1. P. Moingeon *et al.*, *Immunol. Rev.* **111**, 111 (1989); W. C. Hahn, E. Menu, B. E. Bierer, in *Lymphocyte Adhesion Molecules*, Y. Shimizu, Ed. (Landes, Austin, TX, 1993), pp. 105–134; S. C. Meuer *et al.*, *Cell* **36**, 897 (1984); R. F. Siliciano, J. C. Pratt, R. E. Schmidt, J. Ritz, E. L. Reinherz *Nature* **317**, 428 (1985); S. Y. Yang, S. Chouaib, B. Dupont, *J. Immunol.* **137**, 1097 (1986); T. Hünig, G. Tiefenthaler, K.-H. Meyer zum Büschenfelde, S. C. Meuer, *Nature* **326**, 298 (1987); J. A. Gollob, J, Li, E. L. Reinherz, J. Ritz, *J. Exp. Med.*, in press; L. Qin, K. D. Chaving, J. Lin, H. Yagita, J. S. Bromberg, *ibid.* **179**, 341 (1994).
2. P. C. Driscoll, J. G. Cyster, I. D. Campbell, A. F. Williams, *Nature* **353**, 762 (1991).
3. D. F. Wyss *et al.*, *Biochem.* **32**, 10995 (1993); J. M. Withka *et al.*, *Structure* **1**, 69 (1993).
4. E. Y. Jones, S. J. Davis, A. F. Williams, K. Harlos, D. I. Stuart *Nature* **360**, 232 (1992).
5. D. L. Bodian, E. Y. Jones, K. Harlos, D. I. Stuart, S. J. Davis, *Structure* **2**, 755 (1994).
6. M. A. Recny, E. A. Neidhardt, P. H. Sayre, T. L. Ciardelli, E. L. Reinherz, *J. Biol. Chem.* **265**, 8542 (1990).
7. S. Shaw *et al.*, *Nature* **323**, 262 (1986); P. Selvaraj *et al.*, *ibid.* **326**, 400 (1987).
8. K. Kato *et al.*, *J. Exp. Med.* **176**, 1241 (1992); P. A. van der Merwe *et al.*, *Eur. J. Immunol.* **23**, 1373 (1993).
9. Y. W. Wong, A. F. Williams, S. F. Kingsmore, M. F. Seldin, *J. Exp. Med.* **171**, 2115 (1990).
10. A. R. N. Arulanandam *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 11613 (1993).
11. C. Somoza, P. C. Driscoll, J. G. Cyster, A. F. Williams, *J. Exp. Med.* **178**, 549 (1993).
12. A. R. N. Arulanandam *et al.*, *ibid.* **180**, 1861 (1994).
13. P. A. van der Merwe, P. N. McNamee, E. A. Davies, A. N. Barclay, S. J. Davis, *Curr. Biol.* **5**, 74 (1995).
14. M. A. Recny *et al.*, *J. Biol. Chem.* **267**, 22428 (1992).
15. D. F. Wyss, J. S. Choi, G. Wagner, *Biochemistry* **34**, 1622 (1995).
16. Experimental restraints derived from mainly homonuclear NMR data previously used to calculate solution structures of the polypeptide (*3*) were modified through a series of additional heteronuclear 3D and homonuclear 2D experiments. The following 3D NMR experiments were done at 11.7 or 14.1 T at 25°C with 0.7 mM uniformly ^{15}N-labeled hsCD2$_{105}$ in a 9:1 mix of H_2O and D_2O in 20 mM deuterated acetate (pH 4.5): ^{15}N-NOESY-HSQC (mixing time of 75 ms), HNHA [G. W. Vuister and A. Bax, *J. Am. Chem. Soc.* **115**, 7772 (1993)], HNHB [S. J. Archer, M. Ikura, D. A. Torchia, A. Bax, *J. Magn. Reson.* **95**, 636 (1991); J. C. Madsen, O. W. Sørensen, P. Sørensen, F. M. Poulsen, *J. Biomol. NMR* **3**, 239 (1993)], and ^{15}N-TOCSY-HSQC (mixing time of 30 ms). NOESY spectra were recorded at 17.6 T at 25°C and 37°C with 0.7 mM hsCD2$_{105}$ in 100% D_2O in 20 mM deuterated acetate (pH 4.5) with a mixing time of 75 ms. Ninety-two accurate ϕ-angle restraints were derived from the HNHA experiment [D. S. Garrett *et al.*, *J. Magn. Reson. Ser. B* **104**, 99 (1994)]. Nineteen χ1-angle restraints and 13 stereospecific assignments of nonequivalent β-methylene protons were obtained from the HNHB and the TOCSY-HSQC spectra [G. M. Clore, A. Bax, A. M. Gronenborn, *J. Biomol. NMR* **1**, 13 (1991)]. Intrapolypeptide (860 interresidue), 46 intracarbohydrate, and 39 polypeptide-glycan distance restraints were derived from the NOESY spectra. Details of the polypeptide structure refinement will be published elsewhere (D. F. Wyss, J. S. Choi, G. Wagner, in preparation).
17. E. W. Wooten *et al.*, *Eur. Biophys. J.* **18**, 139 (1990).
18. R. A. Dwek, personal communications.
19. Carbon line widths of hsCD2$_{105}$ C-H groups were obtained from a ^{1}H-^{13}C HSQC spectrum, highly resolved in the t_1 dimension, at conditions identical to those for the NOESY spectrum recorded at 37°C (*16*) with a slightly modified pulse sequence [L. E. Kay, P. Keifer, T. Saarinen, *J. Am. Chem. Soc.* **114**, 10663 (1992). Line widths were measured with the line-fitting module of the Felix program in the Insight II package [Felix, version 2.3.5; Biosym Technologies, San Diego (1994)]. ^{13}Cα assignments of essential amino acid residues were obtained largely from a HCCH-TOCSY spectrum [A. Bax, G. M. Clore, A. M. Gronenborn, *J. Magn. Reson.* **88**, 425 (1990); L. E. Kay, G. Xu, A. U. Singer, D. R. Muhandiram, J. D. Forman-Kay, *J. Magn. Reson. Ser. B* **101**, 333 (1993)] on a submillimolar ^{15}N and ^{13}C partially enriched sample (D. F. Wyss, J. S. Choi, G. Wagner, unpublished results). Most other ^{13}Cα assignments were derived from the Hα/Cα_i cross peaks of the HSQC spectrum based on the Hα chemical shifts of the NOESY spectrum recorded at 37°C (*16*). Although large and small line widths indicate little or significant mobility, respectively, on the nanosecond time scale, occasional very large line widths could indicate slow conformational exchange processes on the microsecond time scale. This latter aspect has not been investigated here. A detailed description of the experimental parameters to record and process the HSQC spectrum, as well as the procedure used to obtain the carbon line widths, will be published elsewhere (D. F. Wyss and G. Wagner, in preparation).
20. D. F. Wyss *et al.*, data not shown.
21. One hundred and twenty NMR structures were computed by distance geography (DG) methods [T. F. Havel, *Prog. Biophys. Mol. Biol.* **56**, 43

(1991)] with DGII in the INSIGHT II package [NMRchitect, version 2.3; Biosym Technologies, San Diego (1993)] based on the experimental restraints described in (*16*). In addition, the sugar residues were forced into their lowest energy chair conformations with 68 dihedral angle restraints. Embedded structures were optimized with a simulated annealing protocol [M. Nilges, M. G. Clore, A. M. Gronenborn, *FEBS Lett.* **229**, 317 (1988)] and subsequently energy minimized with experimental restraints according to (*3*). The average root-mean-square deviation from the mean structure was 0.78 Å for all backbone heavy atoms of residues 7 to 103 and the heavy atoms of the trisaccharide core with an SD of 0.07 Å. The 16 structures had on average 16 violations of distance restraints above 0.1 Å with no violation greater than 0.3 Å and on average 21 violated dihedral angle restraints with no violation greater than 4°. Atomic coordinates have been deposited in the Brookhaven protein data bank (access code 1GYA).

22. D. H. van den Eijnden, W. M. Blanken, A. van Vliet, *Carbohydr. Res.* **151**, 329 (1986); W. A. Lubas and R. G. Spiro, *J. Biol. Chem.* **263**, 3990 (1988).
23. Digestion of $hsCD2_{105}$ with endo H and PNGase F was as follows: $HsCD2_{105}$ was digested with PNGase F and purified as described (*14*). Endo H digestions were done in 50 mM citrate (pH 5.5) at a protein concentration of 400 μg/ml. Endo H was added at a concentration of 20 mU and the sample was incubated overnight at 37°C. Digested $hsCD2_{105}$ was then purified with a TSK 2000SW gel filtration column equilibrated in 20 mM sodium acetate, 200 mM $NaSO_4$ (pH 5).
24. For site-directed mutagenesis, the full-length CD2 sequence in M13mp18 vector was used as the parental plasmid for generating the CD2 mutant (*10*). Mutagenesis was done according to the thionucleotide method (Amersham), with sense oligonucleotides and antisense template. The mutants were confirmed by sequencing the mutated single-stranded M13-CD2. Mutant CD2 inserts were subcloned to CDM8 expression vector at the Xba I site (*10*). The entire mutant CD2 adhesion domain in CDM8 was resequenced to ensure the correct mutation in each construct. For transfection, SRBC rosetting, and immunofluorescence assays, COS-7 cells were cultured in Dulbecco's modified Eagle's medium (Bio*whittaker) containing 10% fetal calf serum (FCS), 2 mM glutamine, and 1% penicillin and streptomycin. Cells were seeded in six-well dishes at 5×10^4 per well and subsequently transfected the next day with CDM8-CD2 as described [N. R. Landau and D. R. Littman, *J. Virol.* **66**, 5110 (1992)]. For immunofluorescence, transfected cells were released after 48 hours by incubation in 0.02% EDTA–phosphate-buffered saline (PBS) for 5 to 10 min. The released cells were analyzed by indirect immunofluorescence, with anti-CD2 mAbs 3T48B5 and 35.1 (1:100 dilution of ascites), anti-$T11_2$ (10 μg/ml), and M32B (*10*) rabbit anti-human CD2 heteroantiserum (1:300 dilution). COS-7 cells transfected with the CDM8 vector alone were used as a negative control. The SRBC rosetting assay was done as described (*10*), with 1 ml of 1% SRBC in 5% FCS-RPMI added to each well 48 hours after transfection. After a 1-hour incubation at 37°C followed by four PBS washes, 200 COS-7 cells were examined for rosette formation. A rosette-positive cell was defined as one COS 7 cell binding to ≥ five SRBCs. Vector alone–transfected COS-7 cells served as the negative control.
25. S. J. Davis *et al.*, *J. Biol. Chem.* **270**, 369 (1995).
26. A. S. Tavernor *et al.*, *Eur. J. Biochem.* **219**, 969 (1994).
27. H. C. Joao, I. G. Scragg, R. A. Dwek, *FEBS Lett.* **307**, 343 (1992).
28. J. Deisenhofer, *Biochemistry* **20**, 2361 (1981); D. Freymann *et al.*, *J. Mol. Biol.* **216**, 141 (1990); W. Bode, E. Meyer, J. C. Powers, *Biochemistry* **28**, 1952 (1989); B. Shaanan, H. Lis, N. Sharon, *Science* **254**, 862 (1991).
29. P. J. Kraulis, *J. Appl. Crystallogr.* **24**, 946 (1991).
30. J. F. G. Vliegenthart, L. Dorland, H. van Halbeek, *Adv. Carbohydr. Chem. Biochem.* **41**, 209 (1983).
31. We thank K. Chandrasekhar, K. T. Dayie, and G. Heffron for technical assistance. Supported by the National Institutes of Health, the National Science Foundation, the Keck Foundation, and Sandoz Pharma.

1 May 1995; accepted 26 June 1995

Commitment of CNS Progenitors Along the Dorsoventral Axis of *Drosophila* Neuroectoderm

Gerald Udolph,* Karin Lüer,* Torsten Bossing, Gerhard M. Technau†

In the *Drosophila* embryo, the central nervous system (CNS) develops from a population of neural stem cells (neuroblasts) and midline progenitor cells. Here, the fate and extent of determination of CNS progenitors along the dorsoventral axis was assayed. Dorsal neuroectodermal cells transplanted into the ventral neuroectoderm or into the midline produced CNS lineages consistent with their new position. However, ventral neuroectodermal cells and midline cells transplanted to dorsal sites of the neuroectoderm migrated ventrally and produced CNS lineages consistent with their origin. Thus, inductive signals at the ventral midline and adjacent neuroectoderm may confer ventral identities to CNS progenitors as well as the ability to assume and maintain characteristic positions in the developing CNS. Furthermore, ectopic transplantations of wild-type midline cells into *single minded* (*sim*) mutant embryos suggest that the ventral midline is required for correct positioning of the cells.

The development of the CNS in insects begins shortly after gastrulation, when about 30 neuroblasts (NBs) per hemisegment delaminate from the neuroectoderm according to a stereotypical spatial and temporal pattern (*1*–*3*). Each of these NBs gives rise to a characteristic cell lineage (*4*–*6*). Although the genetic network that regulates NB formation is well understood [reviewed in (*7*)], the mechanisms leading to the specification of the individual NBs are unknown. Candidate genes for NB specification include segmentation genes (*8*, *9*), homeotic genes (*10*), and dorsoventral patterning genes (*11*, *12*). Here, we investigated to what extent the fate of progenitors is determined before their delamination from the ectoderm.

By labeling individual early gastrula cells with the lipophilic fluorescent tracer DiI (*6*) or by isotopic transplantation of single cells labeled with horseradish peroxidase (HRP) (*13*) at the same stage, we have clarified the embryonic lineages of the mesectodermal midline progenitors (*6*) and of several NBs. The type of lineage formed by the various NBs is clearly identifiable by the morphology, number, and position of the neuronal or glial progeny (or of both) and is dependent on their dorsoventral site of origin from the neuroectoderm (Fig. 1, A through E) (Table 1) (*5*, *14*).

To examine the degree to which neuroectodermal progenitor cells (nPCs) and midline cells of the early gastrula [stage 7; stages are described according to (*15*)] are specified to produce specific neural lineages, we performed heterotopic transplantations (Fig. 2 and Table 1). The experiments were designed to test determination as a function of dorsoventral but not of anteroposterior position in the ectoderm. In one experiment (I in Fig. 2), single nPCs were transplanted from dorsal sites [30 to 50% ventrodorsal perimeter (% VD)] of HRP-labeled donors to ventral sites (0 to 20% VD) of unlabeled hosts that were allowed to develop until late embryonic stages (stage 17). From these transplantations, 129 CNS clones were obtained; 81 of these lineages corresponded to ventral NBs (*16*), and 38 clones corresponded to midline lineages (*17*) (Fig. 1, A, B, and E). In 10 cases, identification of the lineage was not possible. The switch to ventral fates of transplanted dorsal nPCs is also corroborated by the expression of a molecular marker. The enhancer-trap line X55 specifically expresses the *lacZ* reporter gene in ventral midline cells (*18*). Dorsal nPCs (30 to 50% VD), which normally do not express the marker, were taken from HRP-labeled X55 donors and transplanted into ventral midline positions of wild-type hosts. Midline clones derived from these transplants stained positive for *lacZ* expression (Fig. 1G) [for *single minded* (*sim*)–*lacZ* expression, see (*19*)]. Thus, dorsal nPCs of the early gastrula are not firmly specified, inductive signals exist at ventral sites to confer ventral neural identities, and dorsal nPCs are receptive to

Institut für Genetik, Universität Mainz, Saarstrasse 21, D-55122 Mainz, Germany.

*These authors contributed equally to this work.
†To whom correspondence should be addressed.