

WHAT IS CLAIMED IS:

1. A protein comprising:
 - (a) a first antigen-binding site comprising an Fab fragment that binds NKG2D;
 - (b) a second antigen-binding site comprising a single-chain variable fragment (scFv) that binds CD33 comprising a heavy chain variable domain and a light chain variable domain; and
 - (c) an antibody Fc domain or a portion thereof sufficient to bind CD16, or a third antigen-binding site that binds CD16, wherein the scFv is linked to the antibody Fc domain or a portion thereof sufficient to bind CD16, or the third antigen-binding site that binds CD16, via a hinge comprising Ala-Ser or Gly-Ala-Ser.
2. The protein of claim 1, wherein the scFv is linked to the antibody Fc domain.
3. The protein of claim 1 or 2, wherein the heavy chain variable domain of the scFv forms a disulfide bridge with the light chain variable domain of the scFv.
4. The protein of claim 3, wherein the disulfide bridge is formed between C44 from the heavy chain variable domain and C100 from the light chain variable domain.
5. The protein of any one of claims 1-4, wherein the scFv is linked to the antibody Fc domain, wherein the light chain variable domain of the scFv is positioned at the N-terminus of the heavy chain variable domain of the scFv, and is linked to the heavy chain variable domain of the scFv via a flexible linker, and the Fab is linked to the antibody Fc domain, and wherein the flexible linker amino acid sequence is the amino acid sequence of SEQ ID NO:186.
6. The protein according to any one of claims 2-4, wherein the heavy chain variable domain of the scFv is linked to the light chain variable domain of the scFv via a flexible linker.
7. The protein of claim 6, wherein the flexible linker comprises the amino acid sequence of SEQ ID NO:186.
8. The protein according to any one of claims 2-4 and 6-7, wherein the heavy chain variable domain of the scFv is positioned at the N-terminus or the C-terminus of the light chain variable domain of the scFv.

9. The protein of claim 8, wherein the light chain variable domain of the scFv is positioned at the N-terminus of the heavy chain variable domain of the scFv.
10. The protein according to any one of claims 1-9, wherein the Fab fragment is linked to the antibody Fc domain or a portion thereof sufficient to bind CD16 or the third antigen-binding site that binds CD16.
11. The protein of claim 10, wherein the heavy chain portion of the Fab fragment comprises a heavy chain variable domain and a CH1 domain, and wherein the heavy chain variable domain is linked to the CH1 domain.
12. The protein according to claim 10 or 11, wherein the Fab fragment is linked to the antibody Fc domain.
13. A protein comprising:
- (a) a first antigen-binding site that binds NKG2D;
 - (b) a second antigen-binding site that binds CD33, comprising a heavy chain variable domain comprising an amino acid sequence at least 95% identical to a heavy chain variable domain listed in Table 1, and a light chain variable domain comprising an amino acid sequence at least 95% identical to a light chain variable domain listed in Table 1; and
 - (c) an antibody Fc domain or a portion thereof sufficient to bind CD16, or a third antigen-binding site that binds CD16.
14. The protein according to claim 13, wherein the heavy chain variable domain and the light chain variable domain of the second antigen-binding site are present on the same polypeptide.
15. The protein according any one of claims 1-14, wherein the first antigen-binding site comprises:
- (1) a heavy chain complementarity-determining region 1 (CDR1) sequence comprising the amino acid sequence of SEQ ID NO:192, a heavy chain complementarity-determining region 2 (CDR2) sequence comprising the amino acid sequence of SEQ ID NO:112, and a heavy chain complementarity-determining region 3 (CDR3) sequence comprising the amino acid sequence of SEQ ID NO:193; and a light chain CDR1 sequence comprising the amino acid sequence of SEQ ID NO:114, a light chain CDR2 sequence comprising the amino acid sequence of SEQ ID

NO:115, and a light chain CDR3 sequence comprising the amino acid sequences of SEQ ID NO:116;

(2) a heavy chain CDR1 sequence comprising the amino acid sequence of SEQ ID NO:93, a heavy chain CDR2 sequence comprising the amino acid sequence of SEQ ID NO:94, and a heavy chain CDR3 sequence comprising the amino acid sequence of SEQ ID NO:95; and a light chain CDR1 sequence comprising the amino acid sequence of SEQ ID NO:96, a light chain CDR2 sequence comprising the amino acid sequence of SEQ ID NO:97, and a light chain CDR3 sequence comprising the amino acid sequences of SEQ ID NO:98;

(3) a heavy chain CDR1 sequence comprising the amino acid sequence of SEQ ID NO:99, a heavy chain CDR2 sequence comprising the amino acid sequence of SEQ ID NO:100, and a heavy chain CDR3 sequence comprising the amino acid sequence of SEQ ID NO:101; and a light chain CDR1 sequence comprising the amino acid sequence of SEQ ID NO:102, a light chain CDR2 sequence comprising the amino acid sequence of SEQ ID NO:103, and a light chain CDR3 sequence comprising the amino acid sequences of SEQ ID NO:104;

(4) a heavy chain CDR1 sequence comprising the amino acid sequence of SEQ ID NO:105, a heavy chain CDR2 sequence comprising the amino acid sequence of SEQ ID NO:106, and a heavy chain CDR3 sequence comprising the amino acid sequence of SEQ ID NO:107; and a light chain CDR1 sequence comprising the amino acid sequence of SEQ ID NO:108, a light chain CDR2 sequence comprising the amino acid sequence of SEQ ID NO:109, and a light chain CDR3 sequence comprising the amino acid sequences of SEQ ID NO:110;

(5) a heavy chain CDR1 sequence comprising the amino acid sequence of SEQ ID NO:192, a heavy chain CDR2 sequence comprising the amino acid sequence of SEQ ID NO:112, and a heavy chain CDR3 sequence comprising the amino acid sequence of SEQ ID NO:195; and a light chain CDR1 sequence comprising the amino acid sequence of SEQ ID NO:114, a light chain CDR2 sequence comprising the amino acid sequence of SEQ ID NO:115, and a light chain CDR3 sequence comprising the amino acid sequences of SEQ ID NO:116;

(6) a heavy chain CDR1 sequence comprising the amino acid sequence of SEQ ID NO:117, a heavy chain CDR2 sequence comprising the amino acid sequence of SEQ ID NO:118, and a heavy chain CDR3 sequence comprising the amino acid sequence of SEQ ID NO:119; and a light chain CDR1 sequence comprising the amino acid sequence of SEQ ID NO:120, a light

chain CDR2 sequence comprising the amino acid sequence of SEQ ID NO:115, and a light chain CDR3 sequence comprising the amino acid sequences of SEQ ID NO:116.

16. The protein according any one of claims 1-15, wherein the first antigen-binding site comprises:

- (1) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:87 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:88;
- (2) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:81 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:82;
- (3) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:83 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:84;
- (4) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:85 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:86;
- (5) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:191 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:88;
- (6) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:89 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:90;
- (7) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:586 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:88;
- (8) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:589 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:88;
- (9) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:592 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:88;

(10) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:595 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:88; or
(11) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:91 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:88.

17. The protein according any one of claims 1-16, wherein the second antigen-binding site comprises a heavy chain variable domain comprising heavy chain CDR1 (CDRH1), heavy chain CDR2 (CDRH2), and heavy chain CDR3 (CDRH3), and a light chain variable domain comprising light chain CDR1 (CDRL1), light chain CDR2 (CDRL2), and light chain CDR3 (CDRL3), wherein the amino acid sequences of CDRH1, CDRH2, CDRH3, CDRL1, CDRL2, and CDRL3 are the amino acid sequences of SEQ ID NOs: 181, 46, 182, 48, 49, and 50; 434, 22, 435, 24, 25, and 26; 181, 28, 436, 30, 31, and 32; 183, 34, 184, 36, 185, and 38; 437, 40, 438, 42, 43, and 44; 181, 52, 439, 54, 55, and 56; 440, 58, 441, 60, 61, and 62; 442, 64, 443, 66, 67, and 68; 181, 70, 444, 72, 73, and 74; 445, 76, 446, 78, 79, and 80; 304, 305, 306, 307, 308, and 309; 310, 311, 312, 313, 314, and 315; 316, 317, 318, 319, 320, and 321; 322, 323, 324, 325, 326, and 327; 328, 329, 330, 331, 332, and 333; 334, 335, 336, 337, 338, and 339; 340, 341, 342, 343, 344, and 345; 346, 347, 348, 349, 350, and 351; 352, 353, 354, 355, 356, and 357; 358, 359, 360, 361, 362, and 363; 364, 365, 366, 367, 368, and 369; 370, 371, 372, 373, 374, and 375; 376, 377, 378, 379, 380, and 381; 382, 383, 384, 385, 386, and 387; 388, 389, 390, 391, 392, and 393; 394, 395, 396, 397, 398, and 399; 400, 401, 402, 403, 404, and 405; 406, 407, 408, 409, 410, and 411; 412, 413, 414, 415, 416, and 417; 422, 423, 424, 425, 426, and 427; or 428, 429, 430, 431, 432, and 433, respectively.

18. The protein according to any one of claims 15-17, wherein the heavy chain variable domain and the light chain variable domain of first antigen-binding site are present on the same polypeptide.

19. The protein according any one of claims 1-12 and 14-18, wherein the second antigen-binding site comprises:

- (a) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:9 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:10;
- (b) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:1 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:2;
- (c) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:3 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:4;
- (d) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:5 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:6;
- (e) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:7 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:8;
- (f) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:11 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:12;
- (g) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:13 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:14;
- (h) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:15 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:16;
- (i) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:17 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:18;
- (j) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:19 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:20;

- [illegible]

- (u) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:286 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:287;
- (v) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:288 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:289;
- (w) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:290 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:291;
- (x) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:292 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:293;
- (y) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:294 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:295;
- (z) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:296 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:297;
- (aa) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:298 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:299;
- (bb) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:300 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:301;
- (cc) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:302 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:303;
- (dd) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:418 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:419; or

(ee) a heavy chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:420 and a light chain variable domain comprising an amino acid sequence at least 95% identical to the amino acid sequences of SEQ ID NO:421.

20. The protein according to any one of claims 1-19 comprising an amino acid sequence selected from the amino acid sequences of SEQ ID NOs:188, 198, 206-223, and 447-488.

21. The protein according to any one of claims 1-19 comprising an scFv linked to an antibody Fc domain, wherein the scFv linked to the antibody Fc domain is the amino acid sequence of SEQ ID NO:187, SEQ ID NO:197, SEQ ID NO:224, SEQ ID NO:225, SEQ ID NO:226, SEQ ID NO:227, SEQ ID NO:228, SEQ ID NO:229, SEQ ID NO:230, SEQ ID NO:231, SEQ ID NO:232, SEQ ID NO:233, SEQ ID NO:234, SEQ ID NO:235, SEQ ID NO:236, SEQ ID NO:237, SEQ ID NO:238, SEQ ID NO:239, SEQ ID NO:240, SEQ ID NO:241, SEQ ID NO:242, or SEQ ID NO:243.

22. The protein according to any one of claims 1-19 comprising the amino acid sequence of SEQ ID NO:189, SEQ ID NO:196, SEQ ID NO:244, or SEQ ID NO:245.

23. The protein according to any one of claims 1-19 comprising an amino acid sequence at least 95% or at least 99% identical to an amino acid sequence selected from the amino acid sequences of SEQ ID NOs:188, 198, 206-223, or 447-488.

24. The protein according to any one of claims 1-19 comprising an amino acid sequence at least 95% or at least 99% identical to the amino acid sequence of SEQ ID NO:187, SEQ ID NO:197, SEQ ID NO:224, SEQ ID NO:225, SEQ ID NO:226, SEQ ID NO:227, SEQ ID NO:228, SEQ ID NO:229, SEQ ID NO:230, SEQ ID NO:231, SEQ ID NO:232, SEQ ID NO:233, SEQ ID NO:234, SEQ ID NO:235, SEQ ID NO:236, SEQ ID NO:237, SEQ ID NO:238, SEQ ID NO:239, SEQ ID NO:240, SEQ ID NO:241, SEQ ID NO:242, or SEQ ID NO:243.

25. The protein according to any one of claims 1-19, wherein the protein comprises

(a) a first polypeptide comprising the amino acid sequence at least 95% or at least 99% identical to the amino acid sequence of SEQ ID NO:187;

- (b) a second polypeptide comprising the amino acid sequence at least 95% or at least 99% identical to the amino acid sequence of SEQ ID NO:189; and
- (c) a third polypeptide comprising the amino acid sequence at least 95% or at least 99% identical to the amino acid sequence of SEQ ID NO:190.

26. The protein of claim 25, wherein the first polypeptide is linked to the second polypeptide via heterodimerization and at least one disulfide bond, and wherein the second polypeptide is linked to the third polypeptide via a disulfide bond.

27. The protein according to any one of claims 1-26, wherein the protein binds to NKG2D with a K_D of 10 nM or weaker affinity, optionally wherein the K_D is measured by Biacore surface plasmon resonance (SPR) analysis.

28. A formulation comprising a protein of any one of claims 1-27, and a pharmaceutically acceptable carrier.

29. Use of the protein according to any one of claims 1-27 or the formulation according to claim 28 in the manufacture of a medicament for treating a CD33-expressing cancer

30. The use of claim 29, wherein the cancer is selected from the group consisting of acute myeloid leukemia (AML), myelodysplastic syndrome (MDS), chronic lymphocytic leukemia (CLL), myeloid blast crisis of CML, acute lymphoblastic leukemia (ALL), acute lymphoblastic lymphoma, myeloproliferative neoplasms (MPNs), lymphoma, non-Hodgkin lymphomas, and classical Hodgkin lymphoma.

31. The use of claim 30, wherein the AML is selected from undifferentiated acute myeloblastic leukemia, acute myeloblastic leukemia with minimal maturation, acute myeloblastic leukemia with maturation, acute promyelocytic leukemia (APL), acute myelomonocytic leukemia, acute myelomonocytic leukemia with eosinophilia, acute monocytic leukemia, acute erythroid leukemia, acute megakaryoblastic leukemia (AMKL), acute basophilic leukemia, acute panmyelosis with fibrosis, and blastic plasmacytoid dendritic cell neoplasm (BPDCN).

32. The use of claim 30 or 31, wherein the AML is characterized by expression of CLL-1 on the AML leukemia stem cells (LSCs).
33. The use of claim 32, wherein the LSCs further express a membrane marker selected from CD34, CD38, CD123, TIM3, CD25, CD32, and CD96.
34. The use of any one of claims 28-30, wherein the AML is a minimal residual disease (MRD).
35. The use of claim 34, wherein the MRD is characterized by the presence or absence of a mutation selected from *FLT3-ITD* ((Fms-like tyrosine kinase 3)-internal tandem duplications (ITD)), *NPM1* (Nucleophosmin 1), *DNMT3A* (DNA methyltransferase gene *DNMT3A*), and *IDH* (Isocitrate dehydrogenase 1 and 2 (IDH1 and IDH2)).
36. The use of claim 30, wherein the MDS is selected from MDS with multilineage dysplasia (MDS-MLD), MDS with single lineage dysplasia (MDS-SLD), MDS with ring sideroblasts (MDS-RS), MDS with excess blasts (MDS-EB), MDS with isolated del(5q), and MDS, unclassified (MDS-U).
37. The use of claim 30 or 36, wherein the MDS is a primary MDS or a secondary MDS.
38. The use of claim 30, wherein the ALL is selected from B-cell acute lymphoblastic leukemia (B-ALL) and T-cell acute lymphoblastic leukemia (T-ALL).
39. The use of claim 30, wherein the MPN is selected from polycythaemia vera, essential thrombocythemia (ET), and myelofibrosis.
40. The use of claim 30, wherein the non-Hodgkin lymphoma is selected from B-cell lymphoma and T-cell lymphoma.
41. The use of claim 30, wherein the lymphoma is selected from chronic lymphocytic leukemia (CLL), lymphoblastic lymphoma (LPL), diffuse large B-cell lymphoma (DLBCL), Burkitt lymphoma (BL), primary mediastinal large B-cell lymphoma (PMBL), follicular lymphoma, mantle cell lymphoma, hairy cell leukemia, plasma cell myeloma (PCM) or multiple myeloma (MM), mature T/NK neoplasms, and histiocytic neoplasms.

42. The protein according to any one of claims 1-27 or the formulation according to claim 28 in the manufacture of a medicament for treating a CD33-expressing cancer for use in therapy.