Supplemental data

1. In silico evaluation of the effect of point mutations within the antigen combining site on the affinity of 14G2a towards GD2

Having uncovered the structural determinants of GD2 recognition by 14G2a we attempted to improve the affinity of the antibody by rationally designed mutagenesis. To this end, based on the crystal structure analysis a number of point mutations were proposed which could presumably induce additional interactions between the antibody and the ganglioside sugar. The point mutants were preselected *in silico* to rationalize the number of variants directed to *in vitro* testing. The preselection was based on comparison of calculated free binding energies for the wild type and mutant complexes with GD2.

1.1 Approaches to free energy calculation

Theoretical prediction of absolute binding affinities for protein-ligand complexes is an important and still not fully solved problem of computational biology. Several procedures were developed for calculation of relative binding free energies for molecular systems including: free energy perturbation (FEP) [1], thermodynamic integration (TI) [1], MM-PBSA (Molecular Mechanics-Poisson-Boltzmann Surface Area) [2], and linear interaction energy (LIE) methods [3]. The most accurate estimations are obtained from FEP and similar TI methods. However, obtaining convergent results from the MD simulations in these methods requires extensive conformational sampling between two different states of the studied system which occur in a thermodynamic cycle. This makes those methods very computationally expensive and thus poorly suited for comparing even a limited number of receptor variants. In the MM-PBSA approach the free energies of binding are usually calculated only based on the MD simulation of a bound state using explicit solvent. The binding free energies are estimated using a Poisson-Boltzmann continuum solvent representation together with a surface-area-dependent term and molecular mechanics energies with an assumption that the structure of either the ligand or the receptor does not change during complex formation [3]. However, the last assumption is only rarely met in true systems. The LIE approach, which we utilized in this study, also relies on ensemble averages generated during MD simulations. The sampling is restricted to two physically relevant states. The first is defined by the ligand in solution while the second corresponds to the solvated receptor-ligand complex. This method assumes that the absolute binding free energy of a ligand is composed of a polar and a nonpolar contribution. The electrostatic contribution to the binding free energy is derived using the electrostatic linear response approximation. The nonpolar contribution is estimated by applying an empirically derived coefficient based on intermolecular van der Waals interaction energies observed in the MD simulations. A detailed description of the LIE method and

examples of its application were previously described [3, 4]. The LIE method was selected in this study for binding free energy estimation because it is computationally much less demanding than more rigorous FEP/TI methods while still being able to deliver accurate results for similar systems [5].

1.2 Employed methodology

To reveal the impact of single point mutations on ligand binding we used the LIE method [4]. LIE is a semi-empirical approach to the estimation of changes in the ligand binding free energy associated with the changes within the receptor. The MD simulations providing energy data for the LIE procedure were calculated for the wild type (WT) 14G2a antibody in complex with GD2 ganglioside, based on crystal structure determined in this study (Protein Data Bank ID: 4TUO) and for the set of point mutants of the antibody (Figures 1 and 2). The structures for short protein fragments not defined within the crystal structure were predicted using Modeller software package [6]. For each particular variant evaluated the selected residue in the vicinity of the binding site was mutated using VMD [7] program, using Mutator plugin. The ligand topology was constructed using Automated Topology Builder (ATB) [8] and manually adjusted to obtain agreement with GROMOS96 [9] force field parameter set used during all simulations. Protein-ligand complexes (including ordered water found within the crystal structure) were inserted into simulation box and solvated with additional water molecules. Geometry optimization and relaxation of constructed systems was accomplished using 1000 steps of steepest decent algorithm followed by 1000 ps of positionally restrained MD (harmonic restraint of 1000 kJ/mol nm^2 was imposed on protein backbone atoms). Next, each system was subject to simulated annealing (SA) procedure lasting 500 ps, with the temperature raising from 310 K to 400 K and then slowly decreasing to 310 K. Again, restraints were set on protein backbone atoms. Finally, the production run of unrestrained MD was performed for 1000 ps. Simulations were performed at the temperature of 310 K and pressure of 1013 hPa. The SPC water model [10] was used. The Particle Mesh Ewald method [11] was applied for treatment of the long-range electrostatic interactions, the simulation step was 1 fs and trajectory frames were recorded every 1 ps. In order to obtain appropriate energy terms for LIE calculations, collected MD trajectories were post-processed Coulombic interaction were evaluated with plain cutoffs. All MD simulations and data analysis were performed using GROMACS v.4.6.5 software package [12]. The mutants were ranked according to the difference in the binding free energy calculated for the ligand and the WT protein complex, and the ligand and the point mutant complex according to the following formula: $\Delta\Delta G = \Delta G_{mutant\_complex} - \Delta G_{WT\_complex.}$

1.3 Mutant selection for *in vitro* evaluation

The effect of 62 point mutations (28 in the light chain and 34 in the heavy chain) on free energy of the interaction between 14G2a antibody and GD2 ganglioside was evaluated *in silico* using LIE method. The number of evaluated point mutations covers relatively well the available point mutation space which we have defined such that only mutants of the residues not already directly involved in

antibody:ligand interaction are evaluated to possibly obtain additional favorable contacts without affecting the contacts already observed within the crystal structure. The results of *in silico* evaluation of mutants are shown in Figure S1 for mutants within the light chain and in Figure S2 for mutants within the heavy chain. This molecular modeling clearly suggests that within the light chain substituting $_{L}$His54 with positively charged residues (arginine or lysine) should increase the mutant affinity towards GD2 compared to the wild type. Further, substitution of $_{L}$Val99 with arginine, $_{L}$Leu51 with lysine and $_{L}$His31 with asparagine should favorably affect the ligand binding. In each case the expected increase in affinity is associated with presumable formation of additional hydrogen bonds between the antibody and the ligand. Following mutants were selected for *in vitro* evaluation: $_{L}$His31Asn, $_{L}$His54Arg and $_{L}$Val99Arg.

Within the heavy chain, molecular modeling suggests that substituting $_{H}$Ala50 with long, positively charged sidechains of lysine or arginine should have a favorable effect on ligand binding. Similar effect should supposedly be obtained by substituting $_{H}$Glu101 with a positively charged sidechain of lysine. It is further predicted that substituting $_{H}$Asp52 for histidine or $_{H}$Ser59 for lysine should increase the affinity of the mutant towards GD2 compared to the wild type. Again, as for the light chain mutants in each case the expected increase in affinity is associated with predicted formation of additional hydrogen bonds between the antibody and the ligand. Based on this results following mutants were selected for *in vitro* evaluation: $_{H}$Ala50Lys, $_{H}$Ser59Lys and $_{H}$Glu101Lys.

Five out of six mutants selected for *in vitro* screening were obtained in the form of soluble proteins. $_{L}$Val99Arg could not have been obtained in our expression system. The results of *in vitro* evaluation of the affinity the mutants towards GD2 are described and discussed in the manuscript.

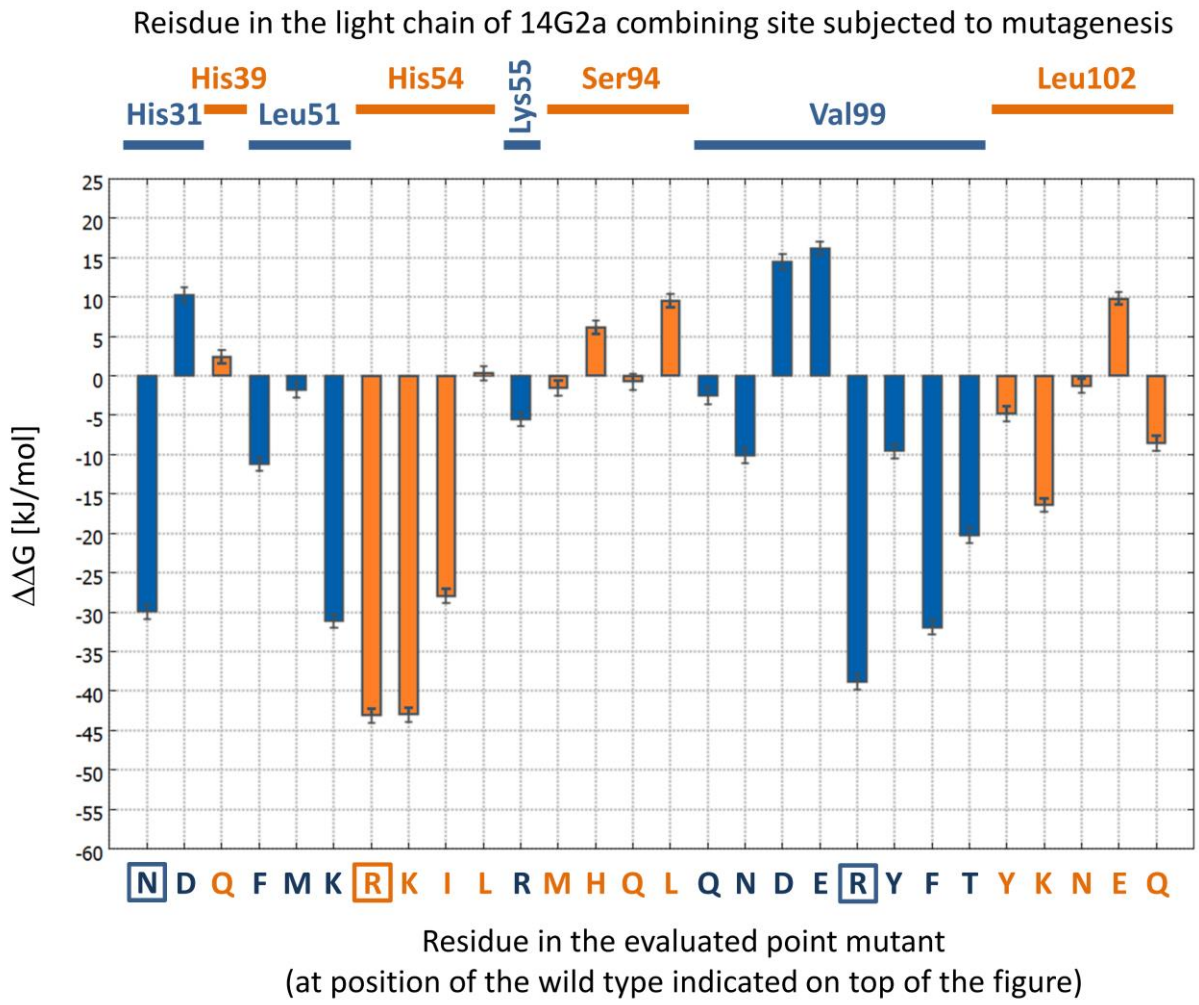

Fig. S1. *In silico* estimated ΔΔG for 14G2a point mutants within the light chain compared to the wild type. Bars represent estimated ΔΔG ($\Delta\Delta G = \Delta G_{mutant\_complex} - \Delta G_{WT\_complex}$) values for point mutations at the antigen combining site and within the light chain. The residue mutated within the wild type is indicated on top of the figure (colors of the bars and descriptions are switched with every residue for clarity). The substituting residues within the mutant are indicated at the bottom of the figure. Standard error of the mean (SEM) is shown. The mutants selected for *in vitro* evaluation are highlighted by a box around the mutant residue at the bottom of the figure.

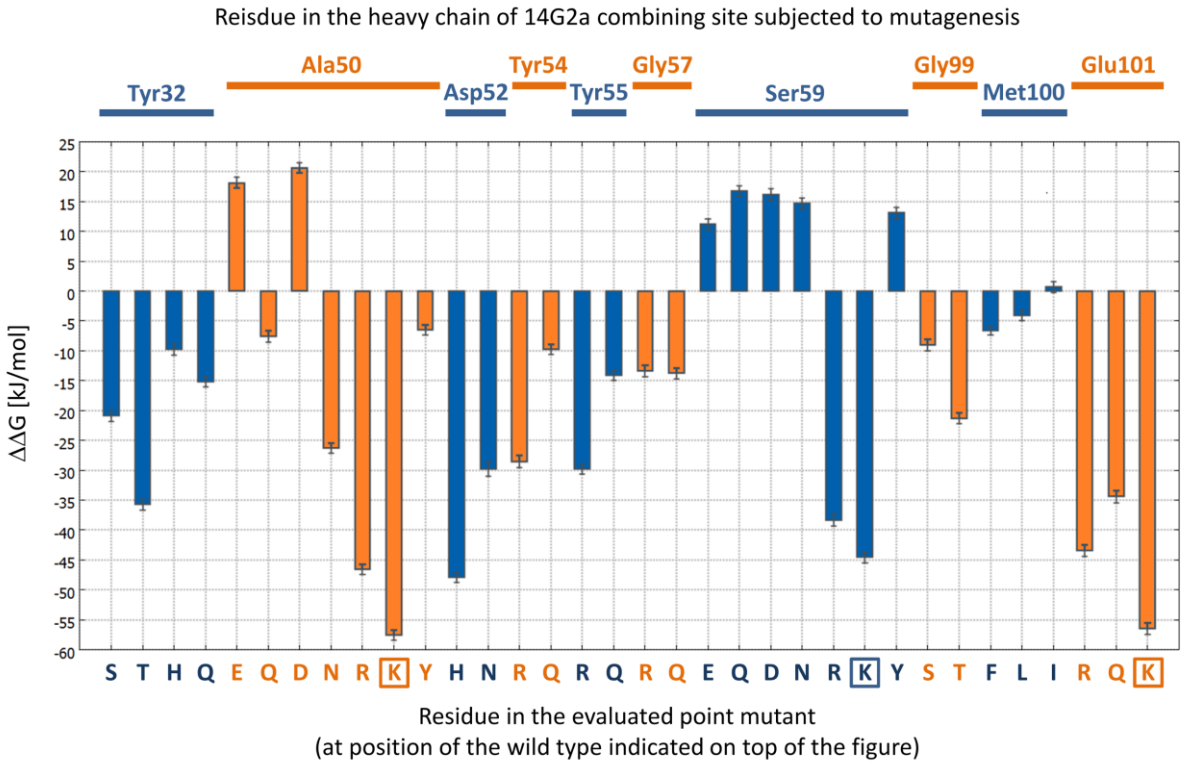

Fig. S2. *In silico* estimated ΔΔG for 14G2a point mutants within the heavy chain compared to the wild type. Bars represent estimated ΔΔG ($\Delta\Delta G = \Delta G_{mutant\_complex} - \Delta G_{WT\_complex}$) values for point mutations at the antigen combining site and within the heavy chain. The residue mutated within the wild type is indicated on top of the figure (colors of the bars and descriptions are switched with every residue for clarity). The substituting residue within the mutant is indicated at the bottom of the figure. Standard error of the mean (SEM) is shown. The mutants selected for *in vitro* evaluation are highlighted by a box around the mutant residue at the bottom of the figure.

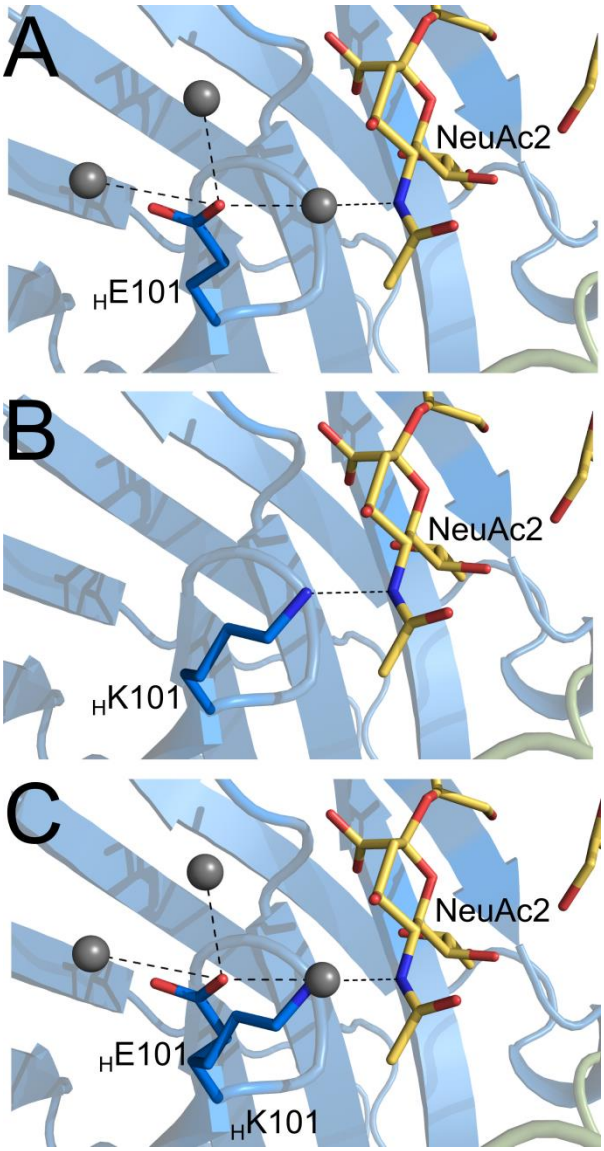

Fig. S3 Hypothetical effects of glutamic acid replacement with lysine at position 101 in the 14G2a heavy chain. (A) Glutamic acid contacts the NeuAc2 residue of GD2 *via* a water molecule (B) A structural model of $_{H}$Lys101 mutant shows possible mechanism of water replacement by the lysine side chain (C) Superposition of crystal structure of the GD2 sugar-bound 14G2a fragment with the model of $_{H}$E101K 14G2a mutant.

2. Comparison of the binding modes of GD2, peptide 1 and peptide 2 at the antigen combining site of 14G2a

The comparison of the binding mode of GD2 and mimicking peptides (1 and 2) at the antigen combining site of 14G2a demonstrates that the observed mimicry is best described by a functional rather than a structural model. This means that the majority of functional groups of neither of the peptides mimic the structural arrangement and interactions of those of the carbohydrate, but rather both peptides sample the antigen combining site of 14g2a in a manner which is largely incomparable to that of the carbohydrate. This conclusion is exemplified in Table S2 which summarizes the interactions between 14G2a and GD2 and between 14G2a and the mimetic peptides (peptide 1 and peptide 2).

Table S1. Comparison of the major interactions between 14G2a and GD2 ganglioside sugar with those between 14G2a and mimicking peptides 1 and 2.

Antibody Residue	Group	GD2 Residue	Type of interaction	Peptide 1	Type of interaction	Peptide 2	Type of interaction
Light Chain							
His31	sidechain	GalNAc	Hydrogen bond	Pro8	Pocket formation	Ala14	Hydrogen bond
Arg32	sidechain	GalNAc	Hydrogen bond	-	-	-	-
		Gal	Hydrogen bond				
		Gal / GalNAc	Hydrogen bond				
		GalNAc	Water mediated				
		Glc	Water mediated				
		Gal	Water mediated				
Asn33	sidechain	Gal	Hydrogen bond	Pro8	Pocket formation	Ala14	Hydrogen bond
Tyr37	sidechain	-	-	Pro8	Pocket formation; CH-π interaction	-	-
				Pro4	Water mediated	-	-
His39	sidechain	NeuAc2	Hydrogen bond	Met6	Hydrogen bond	Leu5	Hydrogen bond
His54	sidechain	-	-	Asn5	Hydrogen bond	-	-
Lys55	sidechain	NeuAc2	Hydrogen bond	Pro4	Hydrogen bond	-	-
Ser 96	sidechain	NeuAc2	Hydrogen bond	Met6	Hydrogen bond	Thr6	Hydrogen bond
	mainchain	GalNAc	Hydrogen bond	-	-	-	-
Thr97	mainchain	GalNAc	Hydrogen bond	-	-	-	
Val99	mainchain	GalNAc	Hydrogen bond	-	-	-	-
Heavy Chain							
Gly31	sidechain	-	-	Arg1	Hydrogen bond	-	-
	mainchain	-	-	-	-	Asn3	Hydrogen bond
Asn33	mainchain	NeuAc2	Hydrogen bond	Arg1	Hydrogen bond	Asn3	Hydrogen bond
	sidechain	NeuAc1	Hydrogen bond	-	-	Val1	Hydrogen bond
	sidechain	-	-	-	-	Thr6	Water mediated
Asn35	sidechain	NeuAc2	Hydrogen bond	-	-	Val1	Hydrogen bond
Asp52	sidechain	NeuAc1	Hydrogen bond	Arg1	Hydrogen bond	-	-
Gly99	mainchain	NeuAc2	Hydrogen bond	Asn3	Hydrogen bond	-	-
Glu101	sidechain	-	-	Asn3	Hydrogen bond	-	-

3. X-ray structure of 14G2a – peptide 1 complex explains the prior results of peptide mutagenesis

A systematic study evaluating the effect of residue by residue mutagenesis within peptide 94# on its affinity towards 14G2a was previously reported [13]. Peptide 94# (RCNPNMEPPRCF) is identical in its antibody binding core with peptide 1 (RCNPNMEPPRCWAAEGD**)** evaluated in this study. 94# was used for mutagenesis due to its shorter length while peptide 1 was used for crystallography due to slightly increased affinity compared to 94# owned to C-terminal extension found in peptide 1 compared to peptide 94# (the detailed description of the origin and properties of both peptides is provided in [13]). In the prior study 27 mutants of peptide 94# were evaluated in a setup where the tested peptide competed with GD2 present on neuroblastoma cells for 14G2a binding. The extent of competition was evaluated by flow cytometry. The study allowed to ascertain the influence of particular residues on the overall affinity of the peptide, but in the lack of structural information, it did not allow to distinguish which mutations affect the overall peptide structure and which the peptide-antibody interactions. Our structure of 14G2a in complex with peptide 1 allows to understand the underlying molecular interactions explaining the effects previously observed for particular mutants. At the same time, the agreement between previous mutagenesis data and the current structural information validates the binding mode observed within the X-ray crystal structure. A residue by residue analysis of the observed effects of mutagenesis on peptide affinity in the light of current structural data is provided in table S1.

Table S2. Insights into structure – activity relationship in recognition of peptide 1 by 14G2a based on X-ray crystallography and mutagenesis studies.

Amino acid	Correspondence of *structural (S)* and mutagenesis (M) data
Arg1	*S: N-terminal amine provides hydrogen bonds with $_H$Asp52 and $_H$Gly31 in one molecule contained in the ASU and $_H$Asn33 in the second molecule in ASU* M: Consistently, deletion of the first residue abolishes binding *S: The sidechain of Arg1 is not involved in peptide:antibody interaction, the sidechain is not defined by electron density* M: Arg1Ala substitution reduces binding, this reduction is not explained by structural data.
Cys2	*S: Cys2-Cys11 disulphide provides a crucial scaffolding element by cyclizing peptide 1* M: Consistently Cys2Ala or Cys11Ala substitutions abolish binding by disrupting the cyclic structure of the peptide. Further in consistence the peptide fragments flanked by cysteine residues (NPNMEP, NPNMEPPR) have no affinity towards 14G2a demonstrating that cysteine mediated cyclization is a prerequisite of interaction.

Asn3	*S: The sidechain of Asn3 is involved in hydrogen bonds with $_H$Gly99 and $_H$Glu101 and water mediated interaction with $_H$Tyr32 and Asn5 (intermolecular).* M: Substitution of Asn3 with either Ala, Gln or Asp abolishes binding since none or not all the described hydrogen bonds are supported
Pro4	*S: The sidechain of Pro4 provides a scaffolding element within the peptide by intermolecular CH-CH interaction with Pro9 and allows a tight turn* M: Substitution of Pro4 with Ala significantly weakens the CH-CH interaction and does not allow such a tight turn
Asn5	*S: The sidechain of Asn5 points away from the antigen combining site, but nevertheless is involved in water mediated interaction with $_H$Tyr32 and $_H$Glu101* M: Asn5Ala impairs binding since the water mediated interactions are not supported. Binding is not completely abolish since the above interactions do not contribute significantly to the overall binding.
Met6	*S: The sidechain of Met6 provides a hydrophobic anchor which inserts deep into the antigen binding pocket* M: Met6Leu substitution negatively affects binding but the interaction is not completely abolished since the sidechain of leucine has comparable properties to that of methionine. In turn, the binding of Met6Ala nad Met6Phe is abolished since the small sidechain of the former residue may not provide the required anchor while the latter is too large to fit into the pocket.
Glu7	*S: The sidechain of Glu7 provides crucial direct and water mediated interactions within the water filled cavity. Tight restraints of the cavity do not allow any modifications of this residue* M: All tested substitutions at position 7 of the peptide (Glu7Ala; Gln; Asp; and Asn) have abolished binding by disrupting the water mediated contacts with Fab.
Pro8	*S: Pro8 is involved in CH-π interaction with the sidechain of $_L$Tyr37.* M: Pro8Ala substitution results in weakening of antibody:mutant peptide interaction by weakening the CH-π interaction.
Pro9	*S: The sidechain of Pro9 provides a scaffolding element within the peptide by intermolecular CH-CH interaction with Pro4* M: Substitution of Pro9 with Ala decreases peptide affinity by weakening the CH-CH interaction.
Arg10	*S: The sidechain of Arg10 is not involved in antibody:peptide interaction, is solvent exposed and not defined by electron density* M: Consistently, substitution of Arg10 with Ala does not affect binding. Even complete removal of Arg10 (by ligation of Pro9 and Cys11) only weekly affects

	binding
Cys11	Same as Cys2
Phe12	*S: The sidechain of Trp12 within peptide 1 provides a significant apolar interface for the interaction with the antibody* M: Peptide #94 differs from peptide 1 by having phenylalanine residue at position 12. We assume that analogically to Trp12 in peptide 1, Phe12 in peptide 94# provides apolar interaction with the antibody, but since the surface provided by Phe12 is smaller than that provided by tryptophane Phe12Trp substitution in 94# increases binding. Phe12Ala substitution abolishes binding since the apolar interaction surface is not provided. Phe12Tyr substitution decreases binding most probably because of steric constraints of the pocket.
Ala13, Ala14 and beyond	S*: Ala13 fills a small opening within the ligand itself. Ala14 is involved in intermolecular interactions but does not interact the antibody. Residues beyond Ala14 are not defined by electron density.* M: Variant of 94# (RCNPNMEPPRCF) having a C-terminal extension (ie. RCNPNMEPPRCFAAEGD) binds better than 94# most probably because Ala13 and Ala14 are present and provide the interactions found within the crystal structure. This suggests that removal of three last residues (EGD) should not affect binding, however such peptides were not tested.

Supplementary references

1. Brandsdal BO, Osterberg F, Almlof M, Feierberg I, Luzhkov VB, Aqvist J: **Free energy calculations and ligand binding**. *Adv Protein Chem* 2003, **66**:123-158.
2. Wang W, Donini O, Reyes CM, Kollman PA: **Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions**. *Annu Rev Biophys Biomol Struct* 2001, **30**:211-243.
3. Aqvist J, Luzhkov VB, Brandsdal BO: **Ligand binding affinities from MD simulations**. *Acc Chem Res* 2002, **35**(6):358-365.
4. Aqvist J, Medina C, Samuelsson JE: **A new method for predicting binding affinity in computer-aided drug design**. *Protein Eng* 1994, **7**(3):385-391.
5. Almlof M, Aqvist J, Smalas AO, Brandsdal BO: **Probing the effect of point mutations at protein-protein interfaces with free energy calculations**. *Biophys J* 2006, **90**(2):433-442.
6. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U, Sali A: **Comparative protein structure modeling using Modeller**. *Curr Protoc Bioinformatics* 2006, **Chapter 5**:Unit 5 6.
7. Humphrey W, Dalke A, Schulten K: **VMD: visual molecular dynamics**. *J Mol Graph* 1996, **14**(1):33-38, 27-38.
8. Koziara KB, Stroet M, Malde AK, Mark AE: **Testing and validation of the Automated Topology Builder (ATB) version 2.0: prediction of hydration free enthalpies**. *J Comput Aided Mol Des* 2014, **28**(3):221-233.
9. Lin Z, van Gunsteren WF: **Refinement of the application of the GROMOS 54A7 force field to beta-peptides**. *J Comput Chem* 2013, **34**(32):2796-2805.

10. van der Spoel D, van Maaren P, Berendsen H: **A systematic study of water models for molecular simulation: Derivation of water models optimized for use with a reaction field**. *J Chem Phys* 1998, **108**:10220-10230.
11. Darden T, York D, Pedersen L: **Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems.** . *J Chem Phys* 1993, **98**:10089-10092.
12. Pronk S, Pall S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D *et al*: **GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit**. *Bioinformatics* 2013, **29**(7):845-854.
13. Horwacik I, Kurcinski M, Bzowska M, Kowalczyk AK, Czaplicki D, Kolinski A, Rokita H: **Analysis and optimization of interactions between peptides mimicking the GD2 ganglioside and the monoclonal antibody 14G2a**. *International journal of molecular medicine* 2011, **28**(1):47-57.