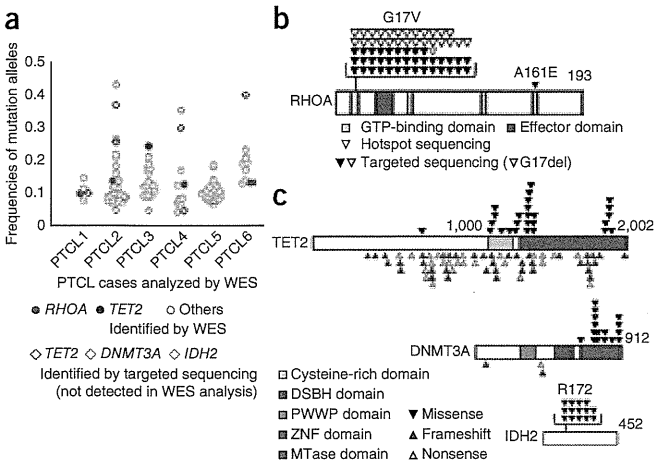

Figure 1 Discovery of a *RHOA* mutation encoding p.Gly17Val in PTCL by whole-exome sequencing. (**a**) Somatic mutations identified in three AITL and three PTCL-NOS samples are shown with the frequencies of mutation alleles plotted. Red and blue filled circles indicate the *RHOA* mutation encoding p.Gly17Val and *TET2* mutations, respectively. Mutations of *TET2*, *IDH2* and *DNMT3A* that were not found by whole-exome sequencing (WES) but were identified later by targeted deep sequencing are also depicted by open diamonds: blue, *TET2*; orange, *DNMT3A*; purple, *IDH2*. (**b**) Positions of RHOA alterations. Nucleotide-binding domains are represented by yellow boxes. The effector domain is represented by a red box. (**c**) Positions of alterations in the TET2, DNMT3A and IDH2 proteins. Black, red and yellow arrowheads indicate missense, frameshift and nonsense mutations, respectively. The cysteine-rich and double-strand β-helix (DSBH) domains of TET2 are represented by a yellow and a red box, respectively. proline-tryptophan-tryptophan-proline (PWWP), zinc-finger (ZNF) and methyltransferase (MTase) domains of DNMT3A are shown by light blue, blue and purple boxes, respectively.

deep sequencing of all coding sequences ($n = 79$) or the mutational hotspot (c.50G>T; p.Gly17Val) ($n = 80$) of *RHOA* (**Supplementary Fig. 1** and **Supplementary Table 3**). *RHOA* mutations were found in 66 of the 159 specimens, with a much higher frequency in AITL (51/72; 70.8%) than PCTL-NOS (15/87; 17.2%) (**Fig. 1b**, **Table 1** and **Supplementary Table 4**). We identified no *RHOA* mutations other than the c.50G>T (p.Gly17Val) mutation except for an in-frame deletion (c.49_51delGGA) resulting in a p.Gly17del (PTCL33) alteration and a missense SNV (c.482C>A) resulting in a p.Ala161Glu (PTCL59) alteration in cases negative for the p.Gly17Val alteration (**Fig. 1b** and **Supplementary Table 4**). We validated all low-frequency mutant *RHOA* alleles (frequency of 0.02–0.05) using an independent deep sequencing platform (Online Methods). No *RHOA* mutations encoding p.Gly17Val were found in other hematologic malignancies, including in myeloid neoplasms ($n = 142$), mature B cell neoplasms ($n = 91$) and mature T cell neoplasms other than AITL and PTCL-NOS ($n = 11$) (**Table 1**), suggesting that the *RHOA* mutation encoding p.Gly17Val is highly specific to AITL and PTCL-NOS among hematologic malignancies.

According to the pathologic definition in the Online Methods[5,9], we classified 21 of 59 immunohistochemically characterized PTCL-NOS cases as T_{FH}-like PTCL-NOS cases. Thirteen of the 21 T_{FH}-like PTCL-NOS cases (61.9%) had the *RHOA* mutation encoding p.Gly17Val, whereas none of the remaining 38 PTCL-NOS cases had this mutation ($P < 0.001$) (**Supplementary Table 5**). Given that almost all AITL cases showed T_{FH}-like features, these findings implied a strong correlation between the *RHOA* mutation encoding p.Gly17Val and the T_{FH}-like phenotype of PTCL, similar to the correlation previously shown between *TET2* mutations and the T_{FH}-like phenotype of PTCL[5]. No clinical parameters were significantly different in the mutation-positive and mutation-negative cases (**Supplementary Fig. 5** and **Supplementary Table 6**).

To investigate the correlation between mutations in *RHOA* and other genes, we also resequenced *TET2*, *IDH1*, *IDH2* and *DNMT3A* in addition to *RHOA* in the subcohort of 79 PTCL (AITL, 46; PTCL-NOS, 33) cases (**Supplementary Figs. 1** and **6**). A total of 97 *TET2* mutations were identified in 54 of the 79 PTCL specimens (68.4%) (AITL, 38 (82.6%); PTCL-NOS, 16 (48.5%)). Similarly, we found *DNMT3A* mutations in 21 PTCL specimens (26.6%) (AITL, 12 (26.0%); PTCL-NOS, 9 (27.3%)). We identified *IDH2* mutations affecting Arg172 (p.Arg172Met, p.Arg172Thr, p.Arg172Ser, p.Arg172Lys and p.Arg172Gly) in 14 cases (17.7%) (AITL, 14 (30.4%); PTCL-NOS, 0 (0%)) (**Figs. 1c** and **2a**, **Supplementary Tables 7** and **8**, and **Supplementary Note**). No *IDH1* mutations were identified. Several mutations in *TET2*, *IDH2* and *DNMT3A*, which had escaped detection in the whole-exome sequencing analysis, were newly identified in the same whole-exome sequencing cohort by this targeted resequencing. Our inability to detect these mutations using whole-exome sequencing might be explained by their low allelic mutational burdens and/or by low sequencing coverage in whole-exome sequencing (**Fig. 1a**). Unexpectedly, however, *TET2* and *DNMT3A* mutations with high-frequency alleles were also newly found in three and two cases, respectively (**Fig. 1a**). The cause of our inability to identify *TET2* and *DNMT3A* mutations by whole-exome sequencing might be the presence of substantial numbers of mutant reads in the reference bone marrow samples (**Supplementary Fig. 7**, **Supplementary Tables 9** and **10**, and **Supplementary Note**).

Remarkably, mutations in *RHOA*, *TET2* and *IDH2* showed strong correlations; all *RHOA*-mutated cases also had *TET2* mutations ($P < 0.001$), and all but one of the *IDH2* mutations were confined to tumors also having *RHOA* and *TET2* mutations ($P < 0.001$) (**Fig. 2a** and **Supplementary Note**). The predominant *TET2* alleles showed significantly higher allelic burden than mutant *RHOA* and *IDH2* alleles in most cases (*TET2* versus *RHOA*, $P < 0.001$; *TET2* versus *IDH2*, $P = 0.001$; **Fig. 2b,c**), whereas *RHOA* and *IDH2* mutations had similar allele frequencies (**Fig. 2d**). Skewed distributions of relative allele frequencies among these mutations strongly suggested that *TET2* mutations predated *RHOA* and/or *IDH2* mutations in most cases.

Table 1 ***RHOA*** **mutation encoding p.Gly17Val in various hematologic malignancies**

Disease	Number of mutated cases (%)
T cell malignancies	$n = 170$
AITL[a]	51/72 (70.8)
PTCL-NOS	15/87 (17.2)
with AITL features	13/21 (61.9)
without AITL features	0/38 (0)
ND[b]	2/28 (7.1)
Other T cell malignancies	0/11 (0)
B cell malignancies	$n = 91$
DLBCL	0/44 (0)
Follicular lymphoma	0/19 (0)
Other B cell malignancies	0/28 (0)
Myeloid malignancies	$n = 142$
AML	0/89 (0)
MDS	0/36 (0)
MPN	0/14 (0)
MDS/MPN	0/3 (0)

DLBCL, diffuse large B cell lymphoma; AML, acute myeloid leukemia; MDS, myelodysplastic syndrome; MPN, myeloproliferative neoplasm.
[a]Including one case with RHOA p.Gly17del and one case with RHOA p.Ala161Glu.
[b]Not determined.

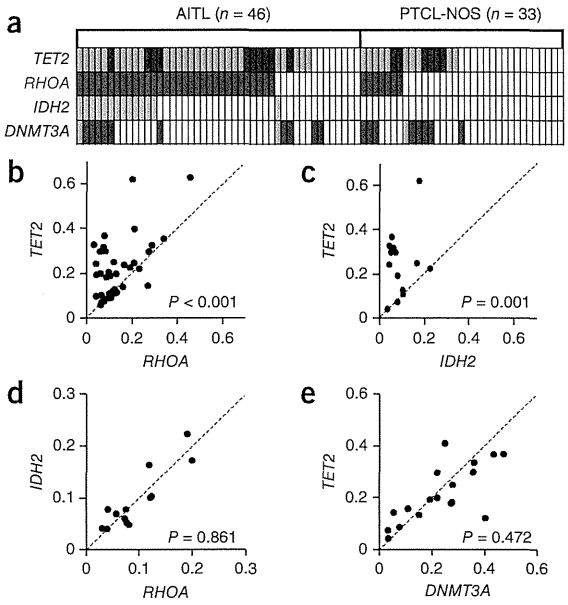

Figure 2 Relationship between *RHOA*, *TET2*, *IDH2* and *DNMT3A* mutations in PTCL. (**a**) Distribution of mutations in *RHOA*, *TET2*, *IDH2* and *DNMT3A* in 79 PTCL (46 AITL and 33 PTCL-NOS) samples that were analyzed by targeted deep sequencing. Two or three distinct *TET2* mutations and two distinct *DNMT3A* mutations were identified in multiple samples. Dark blue and dark green indicate samples having a single *TET2* or *DNMT3A* mutation, respectively, and light blue and light green indicate samples having multiple *TET2* or *DNMT3A* mutations. (**b**–**e**) Comparison of the allele frequencies of two selected mutations in samples harboring mutations in *TET2* and *RHOA* (**b**), *TET2* and *IDH2* (**c**), *RHOA* and *IDH2* (**d**) and *TET2* and *DNMT3A* (**e**). Each axis shows the frequencies of the mutant alleles. When multiple mutations existed in a single gene, the frequencies of major alleles are indicated. Data were analyzed statistically by Wilcoxon rank-sum test.

Mutations in *DNMT3A* largely overlapped and had similar allelic burdens as *TET2* mutations (**Fig. 2e**), but their correlation with *RHOA* or *IDH2* mutations was much less clear (**Fig. 2a**).

To determine the clonal structure of the *RHOA* mutation encoding p.Gly17Val and of other gene mutations, we isolated $CD4^+$ T cells, a fraction enriched for tumor cells and other fractions, from the specimens of two cases (PTCL159 and PTCL160; **Supplementary Figs. 8** and **9**), and we analyzed mutations by targeted resequencing as well as by Sanger sequencing. In PTCL159 (PTCL-NOS in the skin), we found the *RHOA* mutation encoding p.Gly17Val, two *TET2* mutations and a *DNMT3A* mutation (**Supplementary Fig. 8** and **Supplementary Table 7**). Somatic origin of these mutations was confirmed (**Supplementary Fig. 8**). We identified the *RHOA* mutation encoding p.Gly17Val in purified $CD4^+$ cells but not in $CD8^+$ cells. One of the two *TET2* mutations and the *DNMT3A* mutation were identified in both $CD4^+$ and $CD8^+$ cell fractions with apparently similar allelic burdens to each other in the two types of cells, whereas the remaining *TET2* mutation was found only in $CD4^+$ cells and was absent in $CD8^+$ cells (**Supplementary Fig. 8**). These observations suggested that the *RHOA* mutation encoding p.Gly17Val and one of the two *TET2* mutations were confined to $CD4^+$ tumor cells, whereas the other *TET2* mutation and the *DNMT3A* mutation were shared by both $CD4^+$ tumor cells and $CD4^+$ and $CD8^+$ reactive cells (**Supplementary Fig. 8**). In contrast, the *RHOA* mutation encoding p.Gly17Val and two *TET2* mutations identified in PTCL160 (AITL) were all confined to tumor cells (**Supplementary Fig. 9**, **Supplementary Table 7** and **Supplementary Note**). These data indicate that the *RHOA* mutation encoding p.Gly17Val was a specific event in tumor cells. In contrast, *TET2* and *DNMT3A* mutations seemed to have taken place in either $CD4^+$ tumor cells or early progenitor cells such as those that give rise to all hematopoietic cells, as previously described[6,7].

RHOA encodes a small GTPase, which has a highly conserved amino acid structure across species (**Supplementary Fig. 10**). RHOA operates as a molecular switch that regulates a wide variety of biological processes through cycling between an active (GTP-bound) state and an inactive (GDP-bound) state[13,14]. RHOA is activated by specific guanine-exchange factors (GEFs) that catalyze the dissociation of GDP and the rebinding of GTP, and signaling is terminated by hydrolysis of GTP to GDP, a reaction that is stimulated by GTPase-activating proteins (GAPs)[13,14].

Three-dimensional model structures of the Gly17Val RHOA protein suggest compromised binding to GDP and GTP[15,16] (**Supplementary Fig. 11** and **Supplementary Note**). In fact, when we expressed RHOA proteins in NIH3T3 cells, a substantial fraction of wild-type RHOA protein bound GTP or GTPγS in a rhotekin pulldown assay[17], whereas no GTP- or GTPγS-bound form was pulled down for the Gly17Val RHOA mutant (**Fig. 3a**), suggesting severely reduced GTP and GTPγS binding by the Gly17Val mutant.

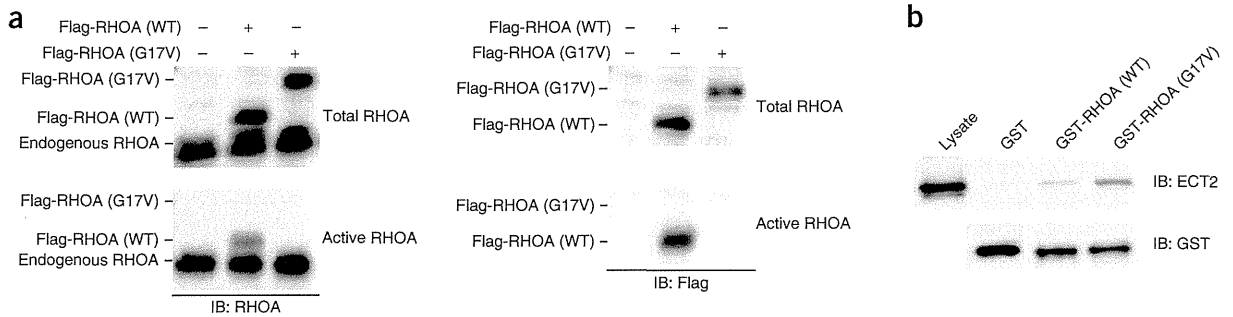

Figure 3 Dominant-negative effect of the Gly17Val RHOA mutant on wild-type RHOA. (**a**) Rhotekin pulldown assay for endogenous RHOA and exogenously expressed wild-type and Gly17Val RHOA in NIH3T3 cells. Extracts from NIH3T3 cells expressing Flag-tagged wild-type or Gly17Val RHOA were pulsed with GTPγS and incubated with glutathione Sepharose beads on which the RHO-binding domain of rhotekin fused to GST was immobilized, and precipitated protein was blotted with antibodies to RHOA (left) and Flag (right) to detect active RHOA specifically. IB, immunoblot; WT, wild type. (**b**) GEF-binding activity of wild-type and Gly17Val RHOA. Lysate from NIH3T3 cells, transiently expressing ECT2 with an N-terminal deletion, was incubated with Sepharose beads on which GST-fused wild-type or Gly17Val RHOA protein was immobilized, and precipitated protein was blotted with antibody to ECT2.

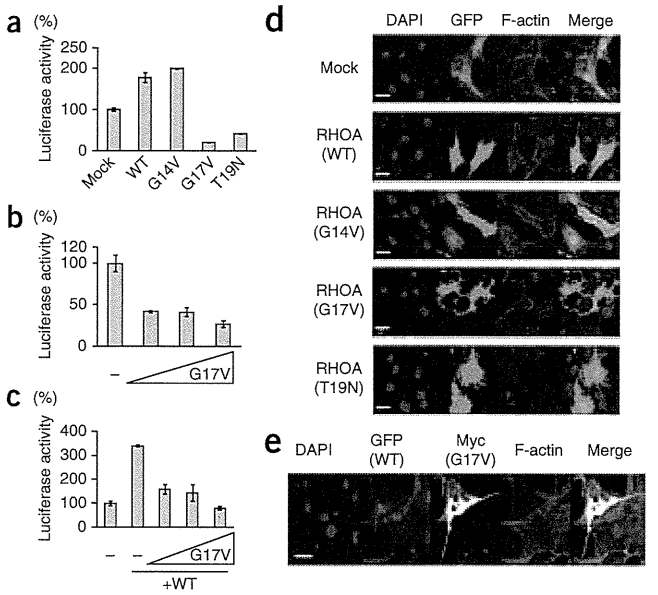

Figure 4 Effects of the Gly17Val RHOA mutant on transcriptional regulation and actin cytoskeleton formation in NIH3T3 cells. (**a**–**c**) Effect of Gly17Val RHOA on the transcriptional activity of the SRF-RE. (**a**) Activity of the SRF-RE reporter in NIH3T3 cells expressing wild-type or mutant (Gly14Val, Gly17Val or Thr19Asn) RHOA protein. (**b**) Effect of increasing amounts (16, 48 or 144 ng/well) of Gly17Val RHOA on SRF-RE reporter activity in NIH3T3 cells. (**c**) Effect of increasing amounts (16, 48 or 144 ng/well) of Gly17Val RHOA on SRF-RE reporter activity enhanced by exogenously expressed wild-type RHOA. In each plot in **a**–**c**, the mean ± s.d. of triplicate experiments is shown. A representative result from three independent experiments is shown. (**d**,**e**) Effect of Gly17Val RHOA on actin cytoskeleton formation. (**d**) F-actin staining with phalloidin (red) in NIH3T3 cells transiently transfected with vector expressing wild-type or mutant (Gly14Val, Gly17Val or Thr19Asn) RHOA. GFP is used as a marker for transduction with each cDNA. (**e**) NIH3T3 cells stably expressing wild-type RHOA were transfected with vector expressing Myc-tagged Gly17Val RHOA. Scale bars in **d**,**e**, 30 μm.

Moreover, the Gly17Val RHOA mutant reduced GTP binding by both the endogenous and exogenous wild-type RHOA proteins in a dose-dependent manner (**Supplementary Figs. 12** and **13**), suggesting a dominant-negative nature for Gly17Val RHOA. This view was further supported by the finding that the Gly17Val RHOA mutant bound ECT2, one of the RhoGEFs, more tightly than wild-type RHOA, as was previously described for Gly17Ala RHOA[18] (**Fig. 3b** and **Supplementary Note**). The Gly17del and Ala161Glu mutants also showed impaired binding capacity for GTP/GTPγS and inhibited GTP binding by wild-type RHOA protein (**Supplementary Fig. 14**). Together, these results support the notion that the RHOA mutants contribute to the pathogenesis of PTCL through the inhibition of wild-type RHOA in a dominant-negative manner, although the amount of mutant RHOA protein seemed to be low in both NIH3T3 cells and primary AITL tumor cells (**Supplementary Fig. 15**, **Supplementary Table 11** and **Supplementary Note**), for an unknown reason.

In accordance with these findings, unlike wild-type RHOA and mutant Gly14Val RHOA, the Gly17Val RHOA mutant did not activate transcription from the serum response factor–responsive element (SRF-RE)[19] (**Fig. 4a**,**b**) and instead repressed transcription from SRF-RE activated by exogenously expressed wild-type RHOA (**Fig. 4c**), as did a known dominant-negative mutant of RHOA (Thr19Asn) (**Fig. 4a** and data not shown). Gly17Val as well as Thr19Asn RHOA also attenuated actin stress fiber formation in NIH3T3 cells, which was markedly induced by wild-type and Gly14Val RHOA[20] (**Fig. 4d**). Furthermore, the Gly17Val RHOA mutant inhibited the assembly of actin stress fibers in NIH3T3 cells stably expressing wild-type RHOA (**Fig. 4e**). All these data suggest that the Gly17Val mutant functions in a dominant-negative manner with respect to wild-type RHOA.

To investigate the effect of wild-type and Gly17Val RHOA on T cells, we established Jurkat cells inducibly expressing wild-type or Gly17Val RHOA (**Fig. 5a**). When wild-type RHOA was expressed, the proliferation of Jurkat cells was significantly decreased (WT Dox (+) versus Mock DOX (+), $P < 0.001$, days 2–4; **Fig. 5b**), and G1-to-S cell cycle progression was suppressed (**Supplementary Fig. 16**). In contrast, inducibly expressed Gly17Val RHOA did not affect the growth or cell cycle progression of Jurkat cells (**Fig. 5b** and **Supplementary Fig. 16**). We further performed mRNA sequencing analysis to examine the effect of the *RHOA* mutation encoding p.Gly17Val on gene expression, using RNA prepared from Jurkat cells inducibly expressing wild-type or Gly17Val RHOA or mock-transfected cells, as well as RNA from NIH3T3 cells transiently expressing wild-type or Gly17Val RHOA or mock-transfected cells. Gene Set Enrichment Analysis (GSEA)[21,22] demonstrated that the serum response factor (SRF) pathway, known to be activated under RHOA signaling[23], was significantly enriched at a false discovery rate (FDR) q value less than 0.25 for cells expressing wild-type RHOA versus mock-transfected cells in both Jurkat and NIH3T3 cells

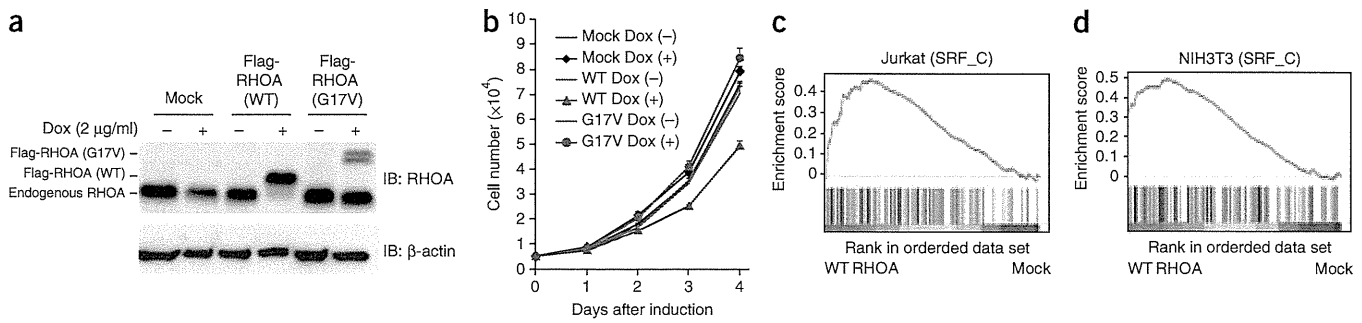

Figure 5 Effect of Gly17Val RHOA on T cells. (**a**) Doxycycline (Dox)-induced expression of wild-type and Gly17Val RHOA in Jurkat cells. A protein blot with antibody to RHOA is shown. β-actin is used as a loading control. (**b**) Proliferation of Jurkat cells inducibly expressing wild-type or Gly17Val RHOA. Absorbance (at 450 nm) was converted to cell number. The mean ± s.d. of quadruplicate experiments is shown. A representative result from three independent experiments is shown. (**c**,**d**) GSEA for Jurkat cells inducibly expressing wild-type or Gly17Val RHOA or mock transfected and NIH3T3 cells transiently expressing wild-type or Gly17Val RHOA or mock transfected ($n = 2$ each). The SRF pathway was differentially enriched in both Jurkat cells (**c**) and NIH3T3 cells (**d**). SRF_C refers to the V$SRF_C gene set.

(**Fig. 5c,d** and **Supplementary Table 12**). The SRF pathway was reported as an essential mediator of T cell development in the thymus[24,25], although we found no clue to its functional relevance in AITL development in the literature. We did not observe enrichment of the SRF pathway in either cell type expressing Gly17Val RHOA compared to mock-transfected cells or cells expressing wild-type RHOA. These findings further support the notion that Gly17Val RHOA is a loss-of-function mutant.

The extremely high frequency and specificity of the *RHOA* mutation encoding p.Gly17Val in AITL and AITL-related PTCL cases unequivocally underscore its major role in the development of these subtypes of PTCL (**Supplementary Fig. 17**). The finding of somatic mutation of *RHOA* in lymphoma, particularly of a mutation with a loss-of-function and/or dominant-negative nature, was rather unexpected because the oncogenic potential of RHOA has been implicated in human cancers[26]. However, several lines of evidence previously suggested a tumor-suppressive role for RHOA in T-lineage cells[26,27]. Moreover, transgenic expression of C3 transferase, an inhibitor of the Rho family of proteins (RHOA, RHOB and RHOC) under the *Lck* promoter has been shown to induce thymic T cell lymphoma in mice[28]. Our observations in Jurkat cells expressing wild-type RHOA are also along these lines. Clearly, further studies are warranted to clarify the molecular pathogenesis mediated by the unique *RHOA* mutation encoding p.Gly17Val in AITL and related PTCL, and such studies might have promising implications for the development of novel diagnostics and therapeutics.

URLs. European Genome-phenome Archive, https://www.ebi.ac.uk/ega/; Genomon-exome, http://genomon.hgc.jp/exome/en/index.html; Picard, http://picard.sourceforge.net/; dbSNP131, http://www.ncbi.nlm.nih.gov/projects/SNP/; 1000 Genomes Project, http://www.1000genomes.org/; MSigDB, http://www.broadinstitute.org/gsea/msigdb.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Genome sequence data are available at the European Genome-phenome Archive under accession EGAS00001000557.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper

ACKNOWLEDGMENTS

We thank Y. Okoshi, N. Obara, Y. Yokoyama, H. Nishikii, N. Kurita and M. Seki for contributing to sample collection and banking. We also thank Y. Sakashita and T. Takahashi for technical assistance, and S. Narumiya and I. Kitabayashi for efficient discussion. This work was supported by Grants-in-Aid for Scientific Research (KAKENHI) from the Ministry of Education, Culture, Sports, Science and Technology of Japan (22134006 to S.O.; 22130002, 24390241, 25112703 and 25670444 to S.C.; 25461407 to M.S.-Y.) and was supported by the Sagawa Foundation for Promotion of Cancer Research, the Naito Foundation, the Mochida Memorial Foundation for Medical and Pharmaceutical Research (M.S.-Y.).

AUTHOR CONTRIBUTIONS

M.S.-Y. prepared DNA samples, sorted the tumor cells, resequenced the samples, and sorted and integrated information. T.E. analyzed the function of wild-type and mutant RHOA. K.Y. resequenced the samples and contributed to the resequencing data analyses. Y. Shiraishi, E.N., K.C., H.T. and S.M. performed bioinformatics analyses of the resequencing data. R.I. and O.N. created the model structure for mutant RHOA. Y.M., H.M., Y.K., R.N.-M., N.B.T., K.S., T.N., Y.H. and M.N. contributed to sample collection and preparation. N.T., S. Sakata, N.N. and K.T. immunostained specimens and performed pathohistological analyses. Y. Okuno and M.S. contributed to the resequencing. A.S.-O. and Yusuke Sato contributed to mRNA sequencing. K.I., Y. Ohta, J.F., S. Shimizu, T.K., Yuji Sato and T.I. collected samples. M.S.-Y., T.E., K.Y., S.O. and S.C. generated figures and tables, and wrote the manuscript. All authors participated in discussions and interpretation of the data and results.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Swerdlow, S.H. *et al. WHO Classification of Tumors of Haematopoietic and Lymphoid Tissues* 4th edn, 306–311 (IARC Press, Lyon, France, 2008).
2. de Leval, L. *et al.* The gene expression profile of nodal peripheral T-cell lymphoma demonstrates a molecular link between angioimmunoblastic T-cell lymphoma (AITL) and follicular helper T (TFH) cells. *Blood* **109**, 4952–4963 (2007).
3. Delhommeau, F. *et al.* Mutation in *TET2* in myeloid cancers. *N. Engl. J. Med.* **360**, 2289–2301 (2009).
4. Mardis, E.R. *et al.* Recurring mutations found by sequencing an acute myeloid leukemia genome. *N. Engl. J. Med.* **361**, 1058–1066 (2009).
5. Lemonnier, F. *et al.* Recurrent *TET2* mutations in peripheral T-cell lymphomas correlate with TFH-like features and adverse clinical parameters. *Blood* **120**, 1466–1469 (2012).
6. Quivoron, C. *et al.* TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. *Cancer Cell* **20**, 25–38 (2011).
7. Couronné, L., Bastard, C. & Bernard, O.A. *TET2* and *DNMT3A* mutations in human T-cell lymphoma. *N. Engl. J. Med.* **366**, 95–96 (2012).
8. Cairns, R.A. *et al. IDH2* mutations are frequent in angioimmunoblastic T-cell lymphoma. *Blood* **119**, 1901–1903 (2012).
9. Rodríguez-Pinilla, S.M. *et al.* Peripheral T-cell lymphoma with follicular T-cell markers. *Am. J. Surg. Pathol.* **32**, 1787–1799 (2008).
10. Yoshida, K. *et al.* Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **478**, 64–69 (2011).
11. Chapman, M.A. *et al.* Initial genome sequencing and analysis of multiple myeloma. *Nature* **471**, 467–472 (2011).
12. Morin, R.D. *et al.* Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* **476**, 298–303 (2011).
13. Bustelo, X.R., Sauzeau, V. & Berenjeno, I.M. GTP-binding proteins of the Rho/Rac family: regulation, effectors and functions *in vivo*. *Bioessays* **29**, 356–370 (2007).
14. Etienne-Manneville, S. & Hall, A. Rho GTPases in cell biology. *Nature* **420**, 629–635 (2002).
15. Ihara, K. *et al.* Crystal structure of human RhoA in a dominantly active form complexed with a GTP analogue. *J. Biol. Chem.* **273**, 9656–9666 (1998).
16. Shimizu, T. *et al.* An open conformation of switch I revealed by the crystal structure of a Mg^{2+}-free form of RHOA complexed with GDP. Implications for the GDP/GTP exchange mechanism. *J. Biol. Chem.* **275**, 18311–18317 (2000).
17. Reid, T. *et al.* Rhotekin, a new putative target for Rho bearing homology to a serine/threonine kinase, PKN, and rhophilin in the rho-binding domain. *J. Biol. Chem.* **271**, 13556–13560 (1996).
18. Arthur, W.T., Ellerbroek, S.M., Der, C.J., Burridge, K. & Wennerberg, K. XPLN, a guanine nucleotide exchange factor for RhoA and RhoB, but not RhoC. *J. Biol. Chem.* **277**, 42964–42972 (2002).
19. Cheng, Z. *et al.* Luciferase reporter assay system for deciphering GPCR pathways. *Curr. Chem. Genomics* **4**, 84–91 (2010).
20. Ridley, A.J. & Hall, A. The small GTP-binding protein rho regulates the assembly of focal adhesions and actin stress fibers in response to growth factors. *Cell* **70**, 389–399 (1992).
21. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550 (2005).
22. Mootha, V.K. *et al.* PGC-1α–responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).
23. Hill, C.S., Wynne, J. & Treisman, R. The Rho family GTPases RhoA, Rac1, and CDC42Hs regulate transcriptional activation by SRF. *Cell* **81**, 1159–1170 (1995).
24. Mylona, A. *et al.* The essential function for serum response factor in T-cell development reflects its specific coupling to extracellular signal–regulated kinase signaling. *Mol. Cell. Biol.* **31**, 267–276 (2011).
25. Fleige, A. *et al.* Serum response factor contributes selectively to lymphocyte development. *J. Biol. Chem.* **282**, 24320–24328 (2007).
26. Karlsson, R., Pedersen, E.D., Wang, Z. & Brakebusch, C. Rho GTPase function in tumorigenesis. *Biochim. Biophys. Acta* **1796**, 91–98 (2009).
27. Hébert, M. *et al.* Rho-ROCK–dependent ezrin-radixin-moesin phosphorylation regulates Fas-mediated apoptosis in Jurkat cells. *J. Immunol.* **181**, 5963–5973 (2008).
28. Cleverley, S.C., Costello, P.S., Henning, S.W. & Cantrell, D.A. Loss of Rho function in the thymus is accompanied by the development of thymic lymphoma. *Oncogene* **19**, 13–20 (2000).

ONLINE METHODS

Subjects and samples. Samples were obtained from individuals with AITL or PTCL-NOS, as well as from individuals with other mature T cell, mature B cell and myeloid neoplasms, and were used after approval was obtained from the local ethics committees at all participating institutes (**Supplementary Tables 1** and **3**). Informed consent was obtained from all living subjects. High-molecular-weight genomic DNA was extracted from archived specimens that were frozen fresh or after fixation. DNA was also extracted from paraffin-embedded, formalin-fixed samples for targeted amplicon sequencing. Constitutional DNA samples were obtained from buccal swabs, mononuclear cells from apparently tumor-free bone marrow aspirates or peripheral blood. Data on clinical outcomes were available for 71 subjects. Samples of a subcohort of PTCL-NOS cases were reviewed by four expert hematopathologists.

Within PTCL-NOS cases, a subgroup without the typical morphology of AITL but having two or more of the following immunostaining features was designated T_{FH}-like PTCL-NOS[5,9]: (i) positive staining for CD10 in tumor cells, (ii) positive staining for PD-1 in tumor cells, (iii) proliferation of CD21-positive follicular dendritic cells and (iv) the presence of EBER-positive B cells.

Sorting of the tumor cell–enriched fraction and other fractions. $CD4^+$ and $CD8^+$ T cell fractions were purified from skin tumors from subject PTCL159, and $CD4^+$ and $CD8^+$ T cell, $CD19^+$ B cell and $CD14^+$ monocyte cell fractions were purified from pleural effusion cells from subject PTCL160.

The skin tumor from subject PTCL159 was processed into single-cell suspension. Cells were stained with fluorescein isothiocyanate (FITC)-conjugated anti-CD4 antibody (BD Biosciences, 555346) and phycoerythrin (PE)-conjugated anti-CD8 antibody (Dako, clone DK25) and were then fractionated on a FACSAria (BD Biosciences).

Mononuclear cells (MNCs) were isolated from the pleural effusion of subject PTLC160 by Ficoll-Paque density-gradient centrifugation. MNCs were stained with FITC-conjugated anti-CD4 antibody and anti-CD14 antibody (BD Biosciences, 555397), PE-conjugated anti-CD8 antibody and PE-conjugated anti-CD19 antibody (Dako, clone HD37) and were fractionated on a FACSAria.

Whole-exome sequencing. Tumor DNA was extracted from subject biopsy samples infiltrated with lymphoma cells. DNA from either buccal mucosa, bone marrow MNCs without apparent lymphoma infiltration or peripheral blood cells was used for the paired normal control. Whole-exome capture was accomplished through the hybridization of sonicated genomic DNA to the bait cDNA library synthesized on magnetic beads (SureSelect Human All Exon 50Mb or V4 kit, Agilent Technologies). Captured targets were subjected to massively parallel sequencing using a HiSeq 2000 (Illumina) according to the standard protocol for 100-bp paired-end reads.

Detection of candidate somatic mutations was performed using our in-house pipeline for whole-exome sequencing[10] with minor modifications. Briefly, sequencing reads were first aligned to the human reference genome (hg19) using Burrows-Wheeler Aligner (BWA)[29] version 0.5.8 with default parameter settings. PCR duplicates were eliminated using Picard. The number of reads containing SNVs and indels in both tumor and germline samples was determined using SAMtools[30], and the null hypothesis of equal allele frequencies in tumor and germline samples was tested using the two-tailed Fisher's exact test. A variant was adopted as a candidate somatic mutation if it had $P < 0.01$, was observed in bidirectional reads (i.e., in both the plus and minus strands of the reference sequence) and its allele frequency was less than 0.1 in the corresponding germline sample. Finally, the list of candidate somatic mutations was generated by excluding synonymous SNVs and other variants registered in either dbSNP131, the 1000 Genomes Project or our in-house SNP database constructed from 180 individual samples. All candidates were validated by deep sequencing.

Validation of whole-exome analysis. Genomic DNA from tumors and paired normal samples was amplified using the REPLI-g mini kit (Qiagen). Regions that included candidate mutations were amplified by genomic PCR using KOD cox neo (TOYOBO) with a NotI linker attached to each primer (**Supplementary Table 13**). Products were combined, and DNA was purified using the QIAquick PCR Purification kit (Qiagen) and digested with NotI. Digested DNA was purified again, and a 1.5-μg aliquot of purified DNA was ligated with T4 DNA ligase for 5 h, sonicated into ~150-bp fragments on average using Covaris and used for the generation of sequencing libraries, according to a modified Illumina paired-end library protocol. Libraries were then subjected to deep sequencing on a MiSeq (Illumina) according to the standard protocol for 150-bp paired-end reads.

Data processing and variant calling were performed with a set of modifications to the method described in a previous publication[10]. Each read was aligned to the set of targeted sequences from PCR amplification, for which BLAT[31], instead of BWA[29], was used with the –fine option. Mapping information in the .psl format was converted to the .sam format with paired-read information using an in house–generated my_psl2sam script. The script was derived from the psl2sam.pl script distributed with SAMtools. Minor changes were applied to the original script to give the paired-end information upon conversion. Of the successfully mapped reads, the following reads were excluded from further analysis: reads that mapped to multiple sites, reads that mapped with more than four mismatched bases and reads that had more than ten soft-clipped bases. Next, the Estimation_CRME script was run to eliminate strand-specific errors and to exclude cycle-dependent errors. A strand-specific mismatch ratio was calculated for each nucleotide variant for both strands using data for those bases between 11 and 50 cycles. To calculate the frequency of each SNV, all reads were mapped to the target reference sequence using BLAT. The number of mapped reads was differentially enumerated for the dichotomic alleles, i.e., mutant and wild-type alleles. For indels, individual reads were first aligned to each of the wild-type and indel sequences and then assigned to the one with which better alignment was obtained in terms of the number of matched bases. Allele frequency was calculated by enumerating each allele according to those assignments. SNVs comprising equal to or more than 2.0% of total reads of the tumor sample rather than the germline sample at each nucleotide position, if it existed, were adopted as somatic mutations.

Targeted sequencing of the *RHOA*, *TET2*, *IDH1*, *IDH2* and *DNMT3A* genes. Targeted sequencing was performed to determine the mutation rate in a large series of PTCL samples for the *RHOA*, *TET2*, *IDH1*, *IDH2* and *DNMT3A* genes. DNA samples from 79 tumors (46 AITL and 33 PTCL-NOS) and 9 paired bone marrow or peripheral blood cell samples were analyzed, including 6 pairs of tumors and controls analyzed by whole-exome sequencing.

DNA samples were prepared as follows: 61 DNA samples were extracted from fresh frozen biopsy specimens, and 18 DNA samples were extracted from paraformaldehyde-lysine-periodate (PLP)-fixed frozen specimens (46 samples were original DNA, and 33 samples were amplified using the REPLI-g mini kit). All exons of the selected genes were captured with the SureSelect target enrichment system (Agilent Technologies), and massively parallel sequencing was then performed on a HiSeq 2000.

For each sample, all sequencing reads were aligned to hg19 using BWA version 0.5.8 with default parameters. After all duplicated reads and low-quality reads and bases were removed, allele frequencies of SNVs and indels were calculated at each genomic position by enumerating the relevant reads using SAMtools. Initially, all variants showing allele frequencies of >0.02 were extracted and annotated with ANNOVAR[32] for further consideration if they were found in >6 reads out of >10 total reads and appeared in both plus- and minus-strand reads. All synonymous variants, known SNPs in public and private databases, including dbSNP131, the 1000 Genomes Project as of 21 May 2012 and our in-house database, were removed. Candidate mutations whose allele frequencies were <5% were validated by PCR-based deep sequencing using Ion Torrent (Life Technologies).

Deep sequencing using Ion Torrent. Fragmented DNA was prepared in the same manner as described above. Libraries were then subjected to deep sequencing on Ion Torrent according to the standard protocol for 300-bp single-end reads. After excluding reads whose length was >200 bases or <50 bases to reduce sequencing errors, the allele frequency was calculated for each SNV or indel as described above.

doi:10.1038/ng.2872

Hotspot sequencing to identify *RHOA* mutations encoding p.Gly17Val. Eighty DNA samples from tumors were extracted from unfixed biopsy specimens (n = 1), PLP-fixed frozen specimens (n = 38) and formalin-fixed, paraffin-embedded specimens (n = 41). All samples were original DNA without amplification, except for one sample amplified using the REPLI-g mini kit. Samples were subjected to genomic PCR with tagged PCR primers (**Supplementary Table 14**) and were subsequently prepared using the NEBNext DNA Library-Prep Reagent Set for Illumina (New England BioLabs). Products underwent massively parallel sequencing on a MiSeq according to the manufacturer's protocol. The SNV representing a G-to-T change comprising equal to or more than 2.0% of total reads at the c.G50 nucleotide position of the *RHOA* gene was adopted as the mutation. Methods of data analysis were the same as described above.

Antibodies. Antibodies used for protein blots or immunostaining were mouse anti-RhoA (1:1,000; Cytoskeleton, ARH03), mouse anti–β-actin (1:2,000; Sigma, A5441), mouse anti–DDDDK tag (1:10,000; MBL, M185-3), mouse anti–Myc tag (1:10,000 for WB, 1:500 for IHC; MBL, M192-3), mouse anti–GST tag (1:2,000; MBL, M071-3), rabbit anti-ECT2 (1:1,000; Millipore, 07-1364), goat anti–mouse IgG conjugated to horseradish peroxidase (HRP) (1:10,000; Dako, P0447), goat anti–rabbit IgG conjugated to HRP (1:10,000; Dako, P0448) and Alexa Fluor 647–conjugated goat anti–mouse IgG (1:1,000; Invitrogen, A-21235).

Cell lines and transfection. NIH3T3 cells (American Type Culture Collection) were cultured at 37 °C in low-glucose DMEM (Sigma) supplemented with 10% heat-inactivated FCS and 1% penicillin-streptomycin. Cells were transfected with plasmids using FuGene6 transfection reagent (Promega) according to the manufacturer's protocol. Jurkat cells (European Collection of Cell Cultures) were cultured at 37 °C in RPMI-1640 (Sigma) supplemented with 10% FCS and 1% penicillin-streptomycin.

Mutagenesis and constructs. Human *RHOA* cDNA was isolated by PCR amplification from peripheral blood MNC-derived cDNA. Mutagenesis to create constructs encoding the Gly14Val, Gly17Val, Gly17del, Thr19Asn and Ala161Glu mutants was carried out with the PrimeStar Mutagenesis Basal kit (TaKaRa) according to the manufacturer's instructions. All cDNA-encoded products were tagged at their N terminus with the Flag and/or c-Myc epitope. These constructs were subcloned into the pEF-neo expression vector, the pGCDN-samIRESGFP retroviral vector and the tetracycline-inducible lentivirus-based expression vector CS-TRE-PRE-Ubc-tTA-I2G7 (ref. 33). cDNA encoding the ECT2-GFP fusion protein was kindly provided by T. Ishizaki (Oita University). An N-terminal deletion mutant (residues 414–882) of ECT2 was generated with the PrimeStar Mutagenesis Basal kit. Constructs encoding wild-type and Gly17Val RHOA were subcloned into the pGEX-2tk vector (GE Healthcare). All cDNA sequences were confirmed by Sanger sequencing.

Retrovirus production and generation of stable cell lines. For retrovirus production, each retroviral vector was transfected into 293gp packaging cells with a vesicular stomatitis virus G (VSV-G) expression plasmid[34]. Retrovirus-containing supernatant was used for the transduction of 293gpg cells to establish stable cell lines capable of producing high titers of VSV-G pseudotyped retroviral particles. To establish cell lines stably expressing wild-type or mutant RHOA, NIH3T3 cells were infected with these retroviruses. Infected cells expressing GFP were isolated using a FACSAria. The purity of sorted cell fractions consistently exceeded 95%.

Rhotekin binding assays. The amount of the GTP-bound form of the RHOA protein was measured using the RhoA Activation Assay kit (Cytoskeleton) according to the manufacturer's instructions. Briefly, cell lysate was incubated at 4 °C for 1 h with a GST fusion protein containing the RHO-binding domain of rhotekin (GST-RBD) immobilized on glutathione Sepharose beads. After washing the beads twice with lysis buffer and once with wash buffer provided by the manufacturer, we fractionated bead-bound proteins by 12% SDS-PAGE and immunoblotted with anti-RHOA and anti-Flag antibodies. Total cell lysate was also blotted with anti-RHOA and anti-Flag antibodies to assess the fractional ratios of rhotekin-bound RHOA proteins.

GEF-binding assays. GST-fused wild-type and Gly17Val RHOA proteins were prepared as previously described with minor modification[35]. Briefly, GST-fused wild-type and Gly17Val RHOA proteins were expressed in BL21 competent *Escherichia coli* cells (TaKaRa), which were lysed in lysis buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 1% Triton X-100, 1 mM dithiothreitol and 1 mM phenylmethylsulfonyl fluoride) and subjected to sonication. Lysate was cleared by centrifugation at 20,000*g* for 15 min at 4 °C, incubated with Glutathione Sepharose 4B beads (GE healthcare) for 45 min at 4 °C and washed twice with lysis buffer.

NIH3T3 cells were transiently transfected with a construct expressing the N-terminal deletion mutant of ECT2 by FuGene6. After 48 h, cells were lysed in lysis buffer, cleared by centrifugation and incubated with GST-fused wild-type or Gly17Val RHOA protein bound to Sepharose beads for 2 h. Beads were washed three times with lysis buffer. Bound material was boiled with Laemmli buffer and blotted with anti-GST and anti-ECT2 antibodies.

SRF-RE reporter assays. For the measurement of activity on SRF-RE, luciferase reporter assays were performed using the pGL4.34 reporter vector (Promega), which contains an SRF-RE and a mutant form of the serum response element lacking the ternary complex factor (TCF)-binding domain. SRF-RE was designed to respond to SRF-dependent and TCF-independent signaling such as the signaling that occurs after RhoA activation[19]. NIH3T3 cells were seeded in 24-well plates and cotransfected with pGL4.34 at 40 ng/well, the expression vector pSRα containing β-galactosidase at 20 ng/well and the expression vector pEF-neo containing various *RHOA* cDNA constructs at the concentrations indicated. Luciferase activity was measured at 48 h after transfection, and values were normalized by β-galactosidase activity.

F-actin staining. NIH3T3 cells were transfected with constructs encoding wild-type or mutant RHOA on glass coverslips. After 48 h, cells were fixed with 4% paraformaldehyde in PBS for 15 min at room temperature and permeabilized with 0.5% Triton X-100 in PBS for 10 min. After washing with PBS, cells were incubated with rhodamine phalloidin (100 nM; Cytoskeleton). For double-staining immunohistochemistry, permeabilized cells were blocked with 3% BSA and 0.1% Triton X-100 in PBS. Then, cells were incubated with mouse anti-Myc antibody (1:500 dilution) followed by Alexa Fluor 647–conjugated goat anti–mouse IgG antibody (1:1,000 dilution) and rhodamine phalloidin (100 nM). Nuclei were stained with DAPI. Images were obtained by confocal laser scanning microscopy (Leica).

Lentivirus production and generation of stable cell lines. For lentivirus production, each lentiviral vector was transfected into HEK293T cells with the psPAX2 packaging plasmid and the pMD2.G envelope plasmid. To establish cell lines inducibly expressing wild-type or Gly17Val RHOA, Jurkat cells were infected with these lentiviruses. Infected cells expressing GFP were sorted on a FACSAria. The purity of sorted cell fractions consistently exceeded 95%.

Cell proliferation assays. For cell growth assays, Jurkat cells transduced with lentiviral vectors were incubated in 96-well culture plates, and the absorbance at 450 nm was measured with Cell Counting Kit-8 (Dojindo) according to the manufacturer's instructions.

Cell cycle analysis. Cell cycle distributions were determined by 5-bromo-2′-deoxyuridine (BrdU) and aminoactinomycin D (AAD) incorporation using the APC BrdU Flow kit according to the manufacturer's protocol (BD Pharmingen). Briefly, Jurkat cells were incubated for 30 min in BrdU (10 μM). Then, cells were fixed, permeabilized, treated with DNase and stained with APC-conjugated anti-BrdU antibody and 7-AAD. Flow cytometry was performed on a FACSCalibur cytometer (BD Biosciences), and data were analyzed with FlowJo software (Tree Star).

mRNA sequencing for Jurkat and NIH3T3 cells. Jurkat cells, inducibly expressing wild-type or Gly17Val RHOA, were described above. Wild-type or Gly17Val RHOA protein expression was induced by the addition of 2 μg/ml doxycycline for 2 d (n = 2 for each). NIH3T3 cells were transiently transfected with pGCDNsamIRESGFP vector encoding wild-type or Gly17Val RHOA (n = 2 for each). After 48 h, GFP-positive cells were sorted by FACSAria.

Total RNA was extracted by RNeasy mini kit (Qiagen) using the RNase-free DNase kit (Qiagen) to reduce contamination from genomic DNA according to the manufacturer's protocol. Libraries for sequencing were prepared using the Illumina TruSeq RNA Sample Preparation kit v2, according to the manufacturer's instructions. Briefly, poly(A)$^+$ RNA was recovered from 1 μg of total RNA using oligo(dT)-coated Sera-Mag magnetic beads. Recovered poly(A)$^+$ RNA was then chemically fragmented. RNA fragments were converted to cDNA using SuperScript II and random primers. The second strand was synthesized using RNase H and DNA polymerase I. cDNA ends were repaired using T4 DNA polymerase, T4 polynucleotide kinase and Klenow DNA polymerase. A single adenosine was added to 3′ ends using Klenow fragment (3′-to-5′ exo minus). Adaptors were attached to cDNA ends using T4 DNA ligase. Fragments were then amplified by ten cycles of PCR using Phusion DNA polymerase. Libraries were validated with an Agilent 2200 TapeStation (Agilent Technologies) and were applied to an Illumina flow cell using the Illumina Cluster Station. Sequencing was performed on a HiSeq 2000 with the paired-end 100-bp read option, according to the manufacturer's instructions.

Reads obtained from RNA sequencing were mapped to the reference transcript and genome using the Genomon-fusion pipeline. For the expression level of each gene, the fragments per kilobase of exon per million mapped reads (FPKM) value was calculated from mapped reads on the gene. GSEA was carried out using GSEA version 2.0. The top ten highest gene sets of normalized enrichment score were listed on the basis of FDR q values (<0.25). Curated gene sets (c2.kegg.version 4.0, c3.tft.version 4.0 and c5.bp.version 4.0) used in this study were obtained from MSigDB collections.

29. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
30. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
31. Kent, W.J. BLAT—the BLAST-like alignment tool. *Genome Res.* **12**, 656–664 (2002).
32. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
33. Yamaguchi, T. *et al.* Development of an all-in-one inducible lentiviral vector for gene specific analysis of reprogramming. *PLoS ONE* **7**, e41007 (2012).
34. Ory, D.S., Neugeboren, B.A. & Mulligan, R.C. A stable human-derived packaging cell line for production of high titer retrovirus/vesicular stomatitis virus G pseudotypes. *Proc. Natl. Acad. Sci. USA* **93**, 11400–11406 (1996).
35. Guilluy, C., Dubash, A.D. & Garcia-Mata, R. Analysis of RhoA and Rho GEF activity in whole cells and the cell nucleus. *Nat. Protoc.* **6**, 2050–2060 (2011).

PLOS ONE

Affinity Improvement of a Therapeutic Antibody by Structure-Based Computational Design: Generation of Electrostatic Interactions in the Transition State Stabilizes the Antibody-Antigen Complex

Masato Kiyoshi[1], Jose M. M. Caaveiro[2,3], Eri Miura[1], Satoru Nagatoishi[2], Makoto Nakakido[3], Shinji Soga[4], Hiroki Shirai[4], Shigeki Kawabata[4], Kouhei Tsumoto[1,2,3,5]*

1 Medical Genome Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Chiba, Japan, **2** Department of Bioengineering, Graduate School of Engineering, The University of Tokyo, Bunkyo-ku, Tokyo, Japan, **3** Laboratory of Medical Proteomics, Institute of Medical Science, The University of Tokyo, Minato-ku, Tokyo, Japan, **4** Molecular Medicine Research Laboratories, Drug Discovery Research, Astellas Pharma Inc., Tsukuba, Ibaraki, Japan, **5** Department of Chemistry and Biotechnology, Graduate School of Engineering, The University of Tokyo, Bunkyo-ku, Tokyo, Japan

Abstract

The optimization of antibodies is a desirable goal towards the development of better therapeutic strategies. The antibody 11K2 was previously developed as a therapeutic tool for inflammatory diseases, and displays very high affinity (4.6 pM) for its antigen the chemokine MCP-1 (monocyte chemo-attractant protein-1). We have employed a virtual library of mutations of 11K2 to identify antibody variants of potentially higher affinity, and to establish benchmarks in the engineering of a mature therapeutic antibody. The most promising candidates identified in the virtual screening were examined by surface plasmon resonance to validate the computational predictions, and to characterize their binding affinity and key thermodynamic properties in detail. Only mutations in the light-chain of the antibody are effective at enhancing its affinity for the antigen *in vitro*, suggesting that the interaction surface of the heavy-chain (dominated by the hot-spot residue Phe101) is not amenable to optimization. The single-mutation with the highest affinity is L-N31R (4.6-fold higher affinity than wild-type antibody). Importantly, all the single-mutations showing increase affinity incorporate a charged residue (Arg, Asp, or Glu). The characterization of the relevant thermodynamic parameters clarifies the energetic mechanism. Essentially, the formation of new electrostatic interactions early in the binding reaction coordinate (transition state or earlier) benefits the durability of the antibody-antigen complex. The combination of *in silico* calculations and thermodynamic analysis is an effective strategy to improve the affinity of a matured therapeutic antibody.

Citation: Kiyoshi M, Caaveiro JMM, Miura E, Nagatoishi S, Nakakido M, et al. (2014) Affinity Improvement of a Therapeutic Antibody by Structure-Based Computational Design: Generation of Electrostatic Interactions in the Transition State Stabilizes the Antibody-Antigen Complex. PLoS ONE 9(1): e87099. doi:10.1371/journal.pone.0087099

Editor: Ashley M. Buckle, Monash University, Australia

Received July 26, 2013; **Accepted** December 19, 2013; **Published** January 27, 2014

Funding: This work was supported by the program for world-leading innovative R&D on science and technology (FIRST) from the Japan Society for the Promotion of Science. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: SS, HS, and SK are full-time employees of Astellas Pharma Inc. KT serves as a consultant for Astellas Pharma Inc. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: tsumoto@bioeng.t.u-tokyo.ac.jp

Introduction

The exquisite specificity and high affinity of antibodies are increasingly exploited for therapeutic and biotechnological purposes, such as in cancer immunotherapy, in diagnosis, and in molecular bio-sensors [1–3]. Biobetter, defined as the biomedicine successfully developed through functional and/or physicochemical improvement of a natural molecule, provides a promising strategy for the next generation of therapeutics [4]. However, the optimization of the physicochemical properties of an antibody without perturbing the affinity and specificity for the cognate molecule is a challenging endeavor. Progress in this field is accelerating, and successful examples of improved antibodies using biobetter strategies have been reported, such as higher affinity for the cognate, increased stability in solution, enhanced pharmacokinetics, diminished immunogenicity, or conjugation to drug delivery systems [5–9]. Nevertheless, to advance further the deployment of biobetter strategies in the design and preparation of improved therapeutics it is necessary to strengthen our understanding of the physicochemical properties of engineered antibodies [10].

The interaction surface of antibody-antigen complexes involves a large number of residues and water molecules establishing multiple non-covalent interactions that are difficult to quantify at the molecular and atomic levels. Although the evaluation of protein-protein binding energy remains a challenging task, progress in computing performance and force-field parametrization are rapidly advancing our predictive capabilities [11,12]. Thus structure-based computational techniques are increasingly employed in the design of biotherapeutic antibodies [13].

Herein, we report the improvement of the binding affinity of the mature antibody 11K2 for its target antigen MCP-1 (monocyte chemotactic protein-1), an important therapeutic target in inflammatory diseases such as arteriosclerosis [14], allergy [15],

and rheumatoid arthritis [16]. The design of an optimized version of the antibody 11K2 with enhanced binding capabilities may improve its therapeutic value in the treatment of inflammatory diseases. Although the binding affinity of 11K2 for its antigen is very high (4.6 pM) [17], the analysis of the crystal structure of the antibody-antigen complex (PDB entry code 2BDN) suggests how to further increase the affinity by optimizing the complementary determining region (CDR) of the variable domain of the light chain (V_L) (Figure 1) [18]. Indeed, the interaction surface between the CDR of the V_L of 11K2 and the antigen is 257 $Å^2$, a value significantly smaller than that of the CDR of the variable domain of the heavy chain (V_H) (511 $Å^2$) [19]. Moreover, the evaluation of the thermodynamic properties of the engineered antibodies is a desirable goal towards increasing our predictive capabilities in future optimization strategies.

In this report we have calculated the binding energy between MCP-1 and a virtual library of systematic mutations at each of the 62 residues of the CDR of the antibody 11K2 by computational methods. The *in silico* calculations suggest that single-mutations carrying charged residues at specific locations favor the binding energy with the antigen. The affinity of the most promising candidates was determined experimentally by the technique of surface plasmon resonance (SPR), which identified several muteins with enhanced affinity for the antigen. The thermodynamic analysis reveals the fundamental mechanism explaining the superior binding capabilities of the optimized antibody. Our study highlights the benefits of combining *in silico* and *in vitro* methodologies for a more effective molecular design by biobetter strategies.

Materials and Methods

Comprehensive mutagenesis *in silico*

Each of the 62 residues belonging to the CDR of 11K2 (as defined by Kabat [20] and Chothia [21]) were subjected to systematic mutagenesis *in silico* with each of the other 19 natural amino acids (62×19 = 1,178 mutations). The initial coordinates of 11K2 in complex with MCP-1 were retrieved from the PDB (entry code 2BDN). For each mutation in the antibody, 100 randomized models were generated using the default parameters of the program MODELLER of the Discovery Studio Suite (Accerlys, San Diego, CA) [22]. Each model of the mutated antibody-antigen complex was optimized by a combination of simulated annealing followed by molecular mechanics minimization until the mean square gradient decreased below 0.01 kcal mol^{-1} $Å^{-2}$. The interaction energy between antigen and antibody in each of the 117,800 models was calculated using the default values in the module Einteract of the package software MOE (Chemical Computing Group, Canada) using the AMBER99 force field after hydrogen atoms were explicitly added and minimized. We used the default dielectric constant (80.0), mimicking the behavior of water. Structural models of muteins lacking non-covalent interactions with the antigen were not considered for further analysis. For each mutation tested, the electrostatic and van der Waals interaction energies were summed up and applied as a surrogate to evaluate the affinity of the mutant for the antigen, as described previously [23]. As a reference we built 1,000 model structures of native antibody by optimizing the x-ray structure using the program MODELLER as explained above. Mutations displaying histograms with more favorable energy of interaction than that of wild type protein were selected for *in vitro* examination (H-L27R, H-L27K, H-N28D, H-N28Q, H-D31E, L-Y30K, L-N31R, L-N31K, L-S53D, L-S53E, L-T56D, L-T56E) (see a representative example in Figure S1).

Expression and purification of MCP-1

The DNA encoding MCP-1 was synthesized by GeneArt (Regensbyrg, Germany) and sequence-optimized for expression in *Escherichia coli* (Table S1). The MCP-1 gene was expressed in a vector pET26b (Novagen) displaying a hexa-histidine tag at the C-terminus. The DNA sequence was confirmed by the dideoxy chain-termination method.

E. coli strain Rosetta2 (DE3) cells (Novagen) transformed with the expression vector of MCP-1 were grown at 28°C in 2× YT broth. Protein expression was induced by addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside when the optical density at 600 nm reached a value of 0.6. Cells were allowed to grow for an additional 16 h at 28°C. The cells were harvested by centrifugation at 8,000× *g* for 8 min and the pellet thus obtained was resuspended in 40 ml of a solution containing 0.5 M NaCl and 50 mM Tris-HCl at pH 8.0 (buffer A). Cells were subsequently lysed by the sonication method with an ultrasonic cell-disruptor instrument (Tommy) for 15 min (Output 7, Duty 50). A compact pellet containing the soluble intracellular components was obtained by centrifugation at 40,000× *g* for 30 min. The soluble fraction was collected and applied onto a Ni-NTA column (Novagen) equilibrated with buffer A. Protein was eluted with stepwise increase of imidazole (10, 20, 30, 50, 100, 200, and 300 mM) in buffer A. The fractions containing MCP-1 were pooled, and subjected to size exclusion chromatography using a HiLoad 26/60 Superdex 75-pg column (GE Healthcare) equilibrated with 50 mM Tris, NaCl 500 mM, EDTA 1 mM at pH 7.4.

Expression and purification of 11K2 scFv

The DNA encoding the single-chain variable fragment (scFv) of 11K2 was chemically synthesized by GeneArt and sequence-optimized for expression in *E. coli* (Table S1). The 11K2 scFv construct was expressed in vector pUTE [24] displaying a hexa-histidine tag at the C-terminus. The DNA sequence was confirmed by the dideoxy chain-termination method.

Cells of *E. coli* strain BL21 (DE3) (Novagen) were transformed with the expression vector of 11K2 scFv and grown at 28°C in LB broth. Protein expression, cell harvesting, and cell lysis were performed as described above for MCP-1. A compact pellet containing the insoluble intracellular components was obtained by centrifugation at 7,500×*g* for 30 min. SDS-PAGE analysis and western blotting were conducted using the insoluble fraction (Figure S2). The soluble fraction was discarded. The insoluble fraction was then solubilized with 6 M guanidine-HCl, 0.5 M NaCl and 50 mM Tris-HCl overnight at 4°C. After solubilization, 11K2 scFv was purified in a Ni-NTA column (Novagen) as described above for MCP-1, except that the equilibration and elution buffers were supplemented with 6 M guanidine-HCl (denaturing conditions).

The purified antibody was refolded by the stepwise dialysis method [25]. Briefly, 11K2 scFv was diluted to 7.5 μM with 6 M Guanidine-HCl in 0.2 M NaCl, 50 mM Tris-HCl, 1 mM EDTA (pH 8.0), followed by stepwise dialysis to remove the denaturant. In order to increase the refolding efficiency, 0.2 M L-Arg-HCl was added to the dialysis solution to minimize protein aggregation when the concentration of guanidine-HCl decreased to 1–0.5 M. The refolded antibody was further purified on a HiLoad 26/60 Superdex 75-pg column equilibrated with a solution containing 0.2 M NaCl, 50 mM Tris-HCl and 1 mM EDTA (pH 8.0). The same protocol was employed for the purification of the antibody muteins.

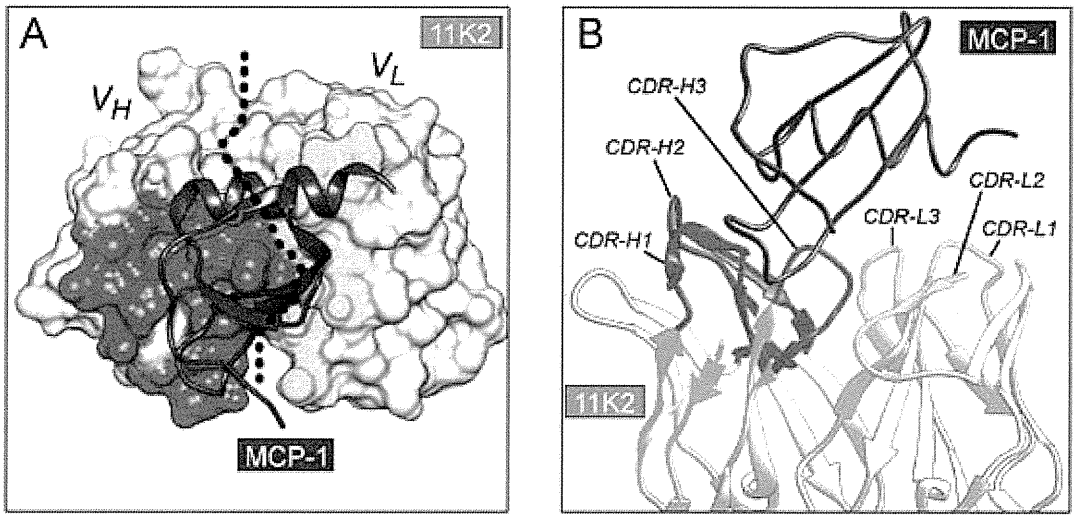

Figure 1. Crystal structure of 11K2 Fab in complex with its antigen MCP-1. The coordinates were retrieved from the PDB (accession code 2BDN). (**A**) Interaction surface of 11K2. The surface in magenta (V_H) and yellow (V_L) correspond to residues of the antibody interacting with the antigen. Other residues are shown in light gray. The dotted line indicates the boundary between V_H and V_L chains. The antigen is depicted in dark green. (**B**) CDR of the antibody. All the residues belonging to the CDR were subjected to virtual systematic mutagenesis. The ribbons in magenta and yellow belong to V_H and V_L chains, respectively. The dark green ribbons depict the antigen.
doi:10.1371/journal.pone.0087099.g001

Binding Assays by SPR

The interaction between MCP-1 and wild-type 11K2 scFv (or muteins) was analyzed by SPR in a Biacore T200 instrument (GE Healthcare). Research grade CM5 Biacore sensor chip (GE Healthcare) was activated by a short treatment with N-hydroxysuccinimide/N-ethyl-N'-(3-dimethylaminopropyl) carbodiimide hydrochloride, followed by immobilization of the antigen MCP-1 at a surface density of ~220 RU. The activated groups on the surface of the sensor were subsequently blocked by injecting 100 µl of a solution containing 1 M ethanolamine. The kinetic data of the binding of 11K2 scFv to the antigen were obtained by injecting increasing concentration of antibody into the sensor chip at a flow rate of 30 µl/min. The measurements were carried out in PBS containing 0.005% (v/v) Tween-20. Contact time and dissociation time were 5 min and 20 min, respectively. Data analysis was performed with the BIAevaluation software (GE Healthcare). Association (k_{on}) and dissociation (k_{off}) rate constants were calculated by a global fitting analysis assuming a Langmuir binding model and a stoichiometry of (1:1). The dissociation constant (K_D) was determined from the ratio of the rate constants [26]:

$$K_D = k_{off}/k_{on}$$

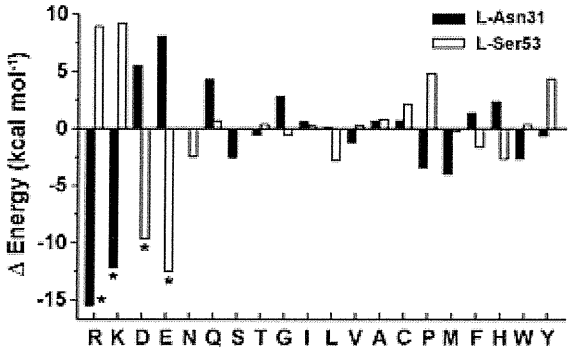

Figure 2. Mutagenesis *in silico*. Average energy values (as the sum of electrostatic and van der Waals energies) of all possible mutations of two different residues (L-Asn31 and L-Ser53) of antibody 11K2 with respect to wild-type. Negative values suggest higher affinity between the mutated protein and the antigen. The mutations indicated by the asterisks were selected for further examination by SPR.
doi:10.1371/journal.pone.0087099.g002

Calculation of thermodynamic parameters

Changes in enthalpy (ΔH°) and entropy (ΔS°) were calculated from the slope and intercept, respectively, of the temperature dependence of the dissociation constant using the van't Hoff approximation [27]:

$$\ln K_D = \Delta H^\circ/RT + \Delta S^\circ/R$$

where R is the gas constant and T is the absolute temperature.

The activation energy parameters were obtained from the temperature dependence of the association rate constant following the Eyring approximation:

$$\ln(k_{on}/T) = -(\Delta H^\ddagger/RT) + (\Delta S^\ddagger/R) + \ln(k_B/h)$$

where k_{on} is the association rate constant, $\Delta H^\ddagger$ is the activation enthalpy, R is the gas constant, T is the absolute temperature, $\Delta S^\ddagger$ is the activation entropy, k_B is the Boltzmann's constant, and h is the Plank's constant.

Results

Computational Selection of Favorable Mutations

To improve the high-affinity of antibody 11K2 for its antigen MCP-1 we selected potentially favorable mutations from a virtual screening (*in silico*) consisting of 1,178 single mutations (19 mutations for each of the 62 residues composing the CDR loops).

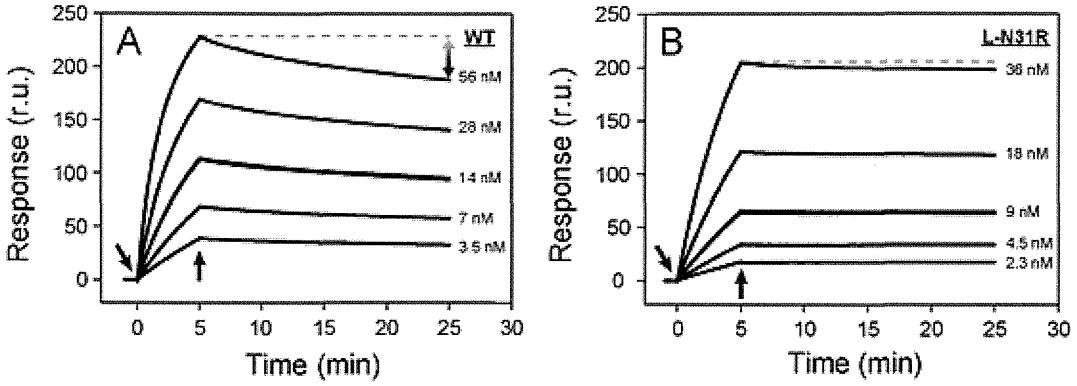

Figure 3. Binding sensorgrams. (**A**) Binding of wild-type 11K2 to its antigen MCP-1. (**B**) Binding of mutein L-N31R to MCP-1. The arrows pointing downward indicate injection of running buffer with 11K2 antibody. The arrows pointing upward correspond to the injection of buffer with no antibody. The response signal is proportional to the amount of scFv 11K2 binding to a chip decorated with antigen MCP-1. The straight dotted line at the top curve in each panel is drawn to appreciate the slower dissociation rate of the mutein with respect to the wild-type protein. The concentration of antibody is given in each panel.
doi:10.1371/journal.pone.0087099.g003

The force field AMBER99 as implemented in the software MOE was used to perform 100 energy minimizations of each mutation (total was 117,800 minimizations), and the corresponding values of energy were subsequently plotted as histograms (Figure S1). The overall shape of the histogram and the median were used to estimate the relative efficacy of each mutation with respect to wild-type antibody. The relative energies of two representative sets of virtual mutations are shown in Figure 2. The residues examined in the example are L-Asp31 and L-Ser53, which are located in the first and second CDR of the V_L chain, respectively. In a large number of virtual mutations, the change of energy is comparatively small (less than ± 3 kcal mol^{-1}). The greatest differences are observed in virtual muteins displaying charged residues (Arg, Lys, Glu or Asp). In some cases, the substitution by a charged residue increases the attractive energy, whereas in other cases the value of energy becomes clearly unfavorable. For example, the calculated energy of mutein L-N31R is clearly more advantageous for binding the antigen than that of wild-type antibody (ΔE^{L-N31R} = -15.4 kcal mol^{-1}), whereas the relative change of energy in mutein L-S53R is clearly destabilizing (ΔE^{L-S53R} = 8.9 kcal mol^{-1}). Mutations displaying very favorable changes of energy in Figure 2 were selected for experimental validation by the technique of SPR (muteins L-D31R, L-D31K, L-S53D, and L-S53E). The following muteins were also selected from the whole virtual screening and examined by SPR: H-L27R, H-L27K, H-N28D, H-N28Q, H-D31E, L-Y30K, L-T56D, L-T56E. We note that 11 muteins from a total of 12 selected mutations involved charged residues (92%).

Table 1. Kinetic parameters of binding of scFv 11K2 to MCP-1 at 25°C.

Protein	k_{on} (M^{-1} s^{-1})	k_{off} (s^{-1})	K_D (nM)	K_D^{WT}/K_D^{mut}	Improved
WT	14×10^4	1.0×10^{-4}	0.80	1	-
H-L27R	14×10^4	3.3×10^{-4}	2.4	0.3	NO
H-L27K	11×10^4	3.9×10^{-4}	3.5	0.2	NO
H-N28D	1.6×10^4	1.6×10^{-4}	1.0	0.8	NO
H-N28Q	22×10^4	19×10^{-4}	8.5	0.09	NO
H-D31E	3.0×10^4	1.5×10^{-4}	5.0	0.16	NO
L-Y30K	13×10^4	3.1×10^{-4}	2.5	0.3	NO
L-N31R	13×10^4	0.22×10^{-4}	0.17	4.6	YES
L-N31K	34×10^4	980×10^{-4}	290	0.003	NO
L-S53D	9.0×10^4	0.55×10^{-4}	0.61	1.3	YES
L-S53E	7.3×10^4	0.14×10^{-4}	0.19	4.2	YES
L-T56D	22×10^4	0.86×10^{-4}	0.39	2.1	YES
L-T56E	10×10^4	0.26×10^{-4}	0.25	3.2	YES

doi:10.1371/journal.pone.0087099.t001

Evaluation of the affinity by SPR

The kinetic rate constants of the binding of wild-type scFv 11K2 (and single-muteins) to immobilized antigen MCP-1 were examined by SPR (Figure 3, Table 1). The injection of the antibody to a surface decorated with antigen produces an increase of the SPR signal that is correlated with the association constant rate (k_{on}) (Figure 3A). The dissociation rate constant (k_{off}) is determined from the signal decay after depleting the solutions from antibody. The values of k_{on} and k_{off} determined at 25°C were 14×10^4 M^{-1} s^{-1} and 1.0×10^{-4} s^{-1}, respectively, corresponding to a dissociation constant (K_D) of 0.8 nM. The small values of k_{off} indicate slow dissociation rates – a clear evidence of tight binding of the antibody to the antigen.

The values of k_{on}, k_{off}, and K_D were also determined for each mutein (Figure 4, Table 1). Significant differences emerge from the comparison of their binding affinities. Whereas a majority of mutations of the CDR of the V_L chain give rise to a robust increase of affinity with respect to the parent antibody (70% of mutations), all mutations belonging to the V_H chain destabilize the antibody-antigen complex. Overall, the mutation with the most favorable effect for the affinity is L-N31R. This mutein binds 4.7-fold stronger to the antigen than the wild-type antibody (Figure 3B; k_{on} = 13×10^4 M^{-1} s^{-1}; k_{off} = 0.22×10^{-4} s^{-1}; K_D = 0.17 nM). Every mutein exhibiting higher affinity for the antigen than that determined for wild-type antibody also displays slower k_{off} values. In contrast, the destabilizing mutations, without exception, accelerate the dissociation of the antibody from the antigen. Thus the simple examination of k_{off} predicts the outcome of the mutation in this particular antibody-antigen system.

Because the mutein L-N31R (displaying the most favorable effect on affinity) incorporates the positively charged residue Arg

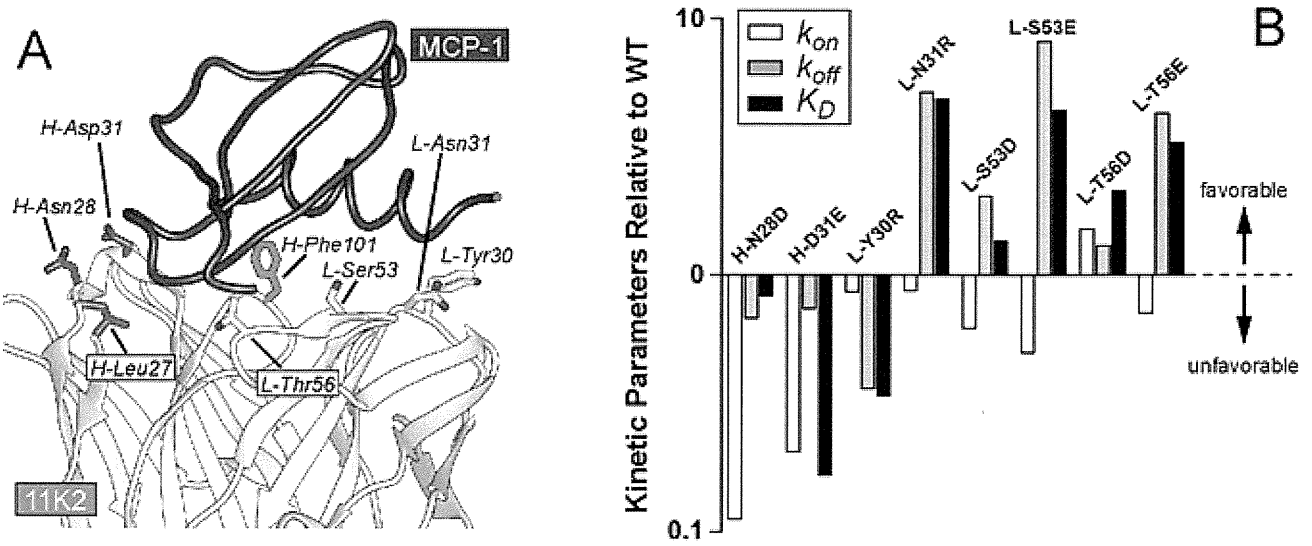

Figure 4. SPR analysis of selected mutations. (**A**) Location of the residues selected in the virtual screening within the crystal structure of the antibody-antigen complex (PDB entry code 2BDN). Mutations belonging to the heavy and light chains are depicted in magenta and yellow, respectively. The antigen is shown in dark green. The hot-spot residue Phe101 is also shown (light green). (**B**) Relative kinetic parameters of the binding of the muteins with respect to wild-type protein. Data is given in Table 1.
doi:10.1371/journal.pone.0087099.g004

we next examined the effect of the ionic strength in three different solutions containing 137, 300 and 500 mM NaCl (Figure 5). No major differences are observed in the kinetic rate constants (k_{on} or k_{off}) or the affinity constant (K_D) of wild-type antibody. Similarly, the kinetic parameters do not change dramatically in mutein L-N31R, although we note that the values of k_{on} decrease slowly but progressively from a value of 17×10^4 M^{-1} s^{-1} in 137 mM NaCl, to a value of 13×10^4 M^{-1} s^{-1} in 500 mM NaCl (25% decrease). Similarly, the affinity also decreases by approximately 25% as manifested by an increase of K_D from a value of 0.36 nM to a value of 0.47 nM. The data indicate a modest role of electrostatic interactions, but only during the association (k_{on}) phase.

Thermodynamic characterization

Thermodynamic parameters for the wild-type antibody and for the optimized muteins were obtained from the temperature dependence of the dissociation constant, as described previously (Figure 6, Table 2) [28,29]._ENREF_28 The van't Hoff enthalpy (ΔH°) and the entropy ($-T\Delta S^\circ$, calculated at 25°C) corresponding to the binding of scFv 11K2 to MCP-1 displayed negative values (ΔH°_{WT} = −7.3 kcal mol^{-1}, $-T\Delta S^\circ_{WT}$ = −5.0 kcal mol^{-1}) indicating favorable contributions from both energetic terms to the free energy of binding (ΔG°_{WT} = −12.3 kcal mol^{-1}). Importantly, the contribution of the enthalpic term increased substantially in the muteins. For example, the value of ΔH° of L-N31R is 3.5-fold more favorable to binding than that of wild-type antibody ($\Delta H^\circ_{L\text{-}N31R}$ = −25.6 kcal mol^{-1}, $\Delta\Delta H^\circ_{L\text{-}N31R}$ = −18.3 kcal mol^{-1}). The change of binding enthalpy of L-N31R is largely (but not completely) compensated by unfavorable changes of entropy ($-T\Delta S^\circ_{L\text{-}N31R}$ = 12.3 kcal mol^{-1}, $-T\Delta\Delta S^\circ_{L\text{-}N31R}$ = 17.3 kcal mol^{-1}, T = 25°C) resulting in a small advantageous change of free energy with respect to the wild-type antibody ($\Delta G^\circ_{L\text{-}N31R}$ = −13.3 kcal mol^{-1}, $\Delta\Delta G^\circ_{L\text{-}N31R}$ = −1.0 kcal mol^{-1}). Similarly, the other muteins exhibit favorable changes of enthalpy not completely compensated by the entropy term. The thermodynamic analysis clearly demonstrates the favorable contribution of the enthalpy to the improved affinity, suggesting that the mutations generate additional non-covalent interactions between the antigen and antibody in agreement with the *in silico* calculations performed above.

Energetic analysis of the transition state

The activation energy of each antibody-antigen complex was determined from the temperature dependence of k_{on} (Table 3,

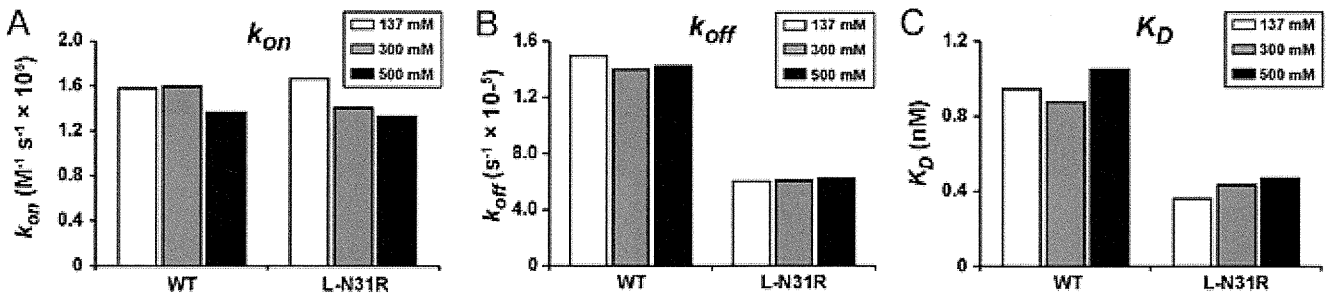

Figure 5. Effect of the ionic strength. (**A**) Association rate constant (k_{on}), (**B**) dissociation rate constant (k_{off}), and (**C**) dissociation constant (K_D). The kinetic parameters of the binding of wild-type 11K2 (or mutein L-N31R) to the antigen MCP-1 were determined in running buffer containing three different concentrations of NaCl (137, 300, or 500 mM) at 25°C.
doi:10.1371/journal.pone.0087099.g005

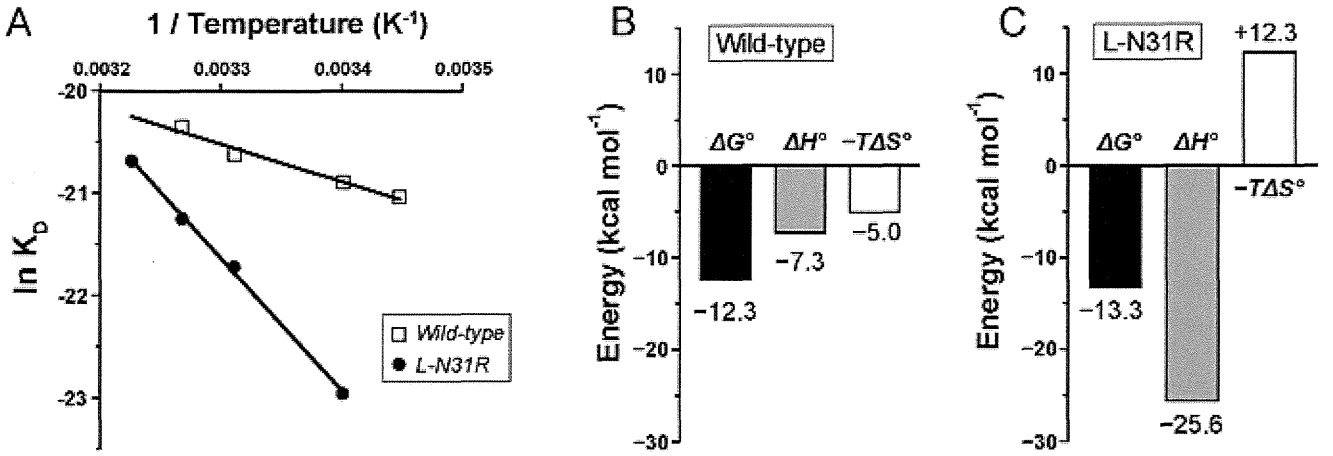

Figure 6. Thermodynamic analysis. (**A**) Regression analysis of the temperature dependence of the dissociation constant K_D yields the van't Hoff enthalpy ($\Delta H°$), entropy ($-T\Delta S°$) and free energy ($\Delta G°$). Empty squares and filled circles correspond to wild-type and L-N31R antibodies, respectively. (**B**) Thermodynamic parameters corresponding to the binding of wild-type antibody to antigen. (**C**) Same parameters obtained for L-N31R.
doi:10.1371/journal.pone.0087099.g006

Figure S3). The activation free energy of wild-type antibody is defined by the unfavorable interactions in the transition state ($\Delta H^{\ddagger} > 0$), reflecting the negative contribution of the dehydration and/or remodeling of protein-protein interactions during the rate determining step ($\Delta G^{\ddagger}_{WT,assoc}$ = 10.6 kcal mol^{-1}, $\Delta H^{\ddagger}_{WT,\mathrm{assoc}}$ = 11.4 kcal mol^{-1}, $-T\Delta S^{\ddagger}_{WT,assoc}$ = −0.8 kcal mol^{-1}). The relative activation free energy of the muteins does not change significantly with respect to the wild-type antibody ($-0.5<\Delta\Delta G^{\ddagger}_{\mathrm{MUT,assoc}}<0.3$). In contrast, the change of enthalpy of the muteins is more advantageous (less unfavorable) than that of wild-type antibody ($-6.9<\Delta\Delta H^{\ddagger}_{MUT,\mathrm{assoc}}<-10.7$ kcal mol^{-1}), suggesting the formation of additional non-covalent interactions between the optimized antibody and the antigen during the transition state (Figure 7A). The values of change of enthalpy in the transition state are correlated with the values of change of enthalpy in equilibrium (Figure 7A). The negative values of enthalpy at equilibrium ($\Delta\Delta H°_{MUT}$) and at the transition state ($\Delta\Delta H^{\ddagger}_{MUT,\mathrm{assoc}}$) demonstrate that the charged residues introduced in the optimized antibody improve the enthalpic contribution to binding. These observations suggest that the charged residues establish electrostatic interactions with the antigen, as depicted in a model of the antibody antigen complex (Figure 7B). The novel interactions are formed early in the complexation reaction, since they play an important role early in the energetic profile of the transition state. In the transition state, these non-covalent interactions are perfectly counterbalanced by unfavorable changes of entropy, reflecting the loss of configurational energy at the rate-limiting step incurred by the approaching proteins ($-T\Delta\Delta S^{\ddagger}_{MUT,assoc}$ = 7.1 ~ 10.7 kcal mol^{-1}, calculated at 25°C) (Figure S4). Although the free energy barrier that the muteins overcome in the transition state is nearly identical within experimental error to that determined for wild-type antibody ($\Delta\Delta G^{\ddagger} \sim 0$), their energetic pathway towards the antibody-antigen complex in equilibrium differs from each other.

Discussion

The affinity of the therapeutic antibody 11K2 for its antigen, the inflammatory cytokine MCP-1, was enhanced up to ~5 fold by employing structure-based computational design. Engineered versions of the antibody were first designed *in silico* and subsequently verified by *in vitro* techniques using a recombinant scFv construct. The detailed thermodynamic characterization revealed the physicochemical principles involved and the operating mechanism.

The selection of suitable candidates of potentially higher affinity among a library of 1,178 virtual mutations was carried out by computational energy minimizations, and the most promising candidates (12 single-mutations) examined experimentally by SPR. Five candidates were mutated at the V_H chain, whereas seven candidates were mutated at the V_L chain. In five cases the affinity of the optimized antibody increased with respect to the wild-type antibody (42% of muteins tested) − a high success rate.

Table 2. Thermodynamic parameters of scFv 11K2.

Protein	$\Delta H°$ (kcal mol^{-1})	$\Delta\Delta H°$ (kcal mol^{-1})	$-T\Delta S°$ (kcal mol^{-1})	$-T\Delta\Delta S°$ (kcal mol^{-1})	$\Delta G°$ (kcal mol^{-1})	$\Delta\Delta G°$ (kcal mol^{-1})
WT	−7.3	0	−5.0	0	−12.3	0
L-N31R	−25.6	−18.3	12.3	17.3	−13.3	−1.0
L- S53D	−13.3	−6.0	0.6	5.6	−12.7	−0.4
L-S53E	−15.1	−7.8	2.4	7.4	−12.7	−0.4
L-T56D	−11.7	−4.4	−1.1	3.9	−12.8	−0.5
L-T56E	−14.5	−7.2	1.9	6.9	−12.6	−0.3

Values of $-T\Delta S°$ and $\Delta G°$ are given at 25°C.
doi:10.1371/journal.pone.0087099.t002

Table 3. Activation energy of association of scFv 11K2 to MCP-1.

Protein	$\Delta H^{\ddagger}$ (kcal mol^{-1})	$\Delta\Delta H^{\ddagger}$ (kcal mol^{-1})	$-T\Delta S^{\ddagger}$ (kcal mol^{-1})	$-T\Delta\Delta S^{\ddagger}$ (kcal mol^{-1})	$\Delta G^{\ddagger}$ (kcal mol^{-1})	$\Delta\Delta G^{\ddagger}$ (kcal mol^{-1})
WT	11.4	0	-0.8	0	10.6	0
L-N31R	3.4	−8.0	6.8	7.6	10.2	−0.4
L-S53D	4.2	−7.2	6.4	7.2	10.6	0
L-S53E	4.5	−6.9	6.4	7.2	10.9	0.3
L-T56D	3.8	−7.6	6.3	7.1	10.1	−0.5
L-T56E	0.7	−10.7	9.9	10.7	10.6	0

Values of $-T\Delta S^{\ddagger}$ and $\Delta G^{\ddagger}$ are given at 25°C.

doi:10.1371/journal.pone.0087099.t003

Importantly, only mutations made in the V_L chain led to higher affinity (five favorable muteins from a total of seven muteins examined, i.e. 71% success rate). Such domain bias contrasts with other computationally-based optimizations, in which favorable muteins are evenly distributed among V_H and V_L chains (reviewed in Kuroda *et al*, 2012) [13]. The deleterious effect of mutations at V_H for the binding of antigen reflects the excellent optimization of this region in the wild type antibody, both in terms of interaction surface (511 Å^2 for V_H; but only 257 Å^2 for V_L) and the presence of a hot-spot residue (Figure 4) [18,30]. In other words, any modification of the V_H chain by site-directed mutagenesis disturbs the carefully orchestrated interaction surface with the antigen and consequently reduces the binding affinity.

The single-mutations selected from the virtual screening incorporate, in 92% of the cases, a charged residue. This observation emphasizes the importance of the electrostatic forces in the computational optimization. In two previous studies the optimized antibodies incorporated multiple mutations in their primary sequences [31,32]. For example, the improvement of the binding affinity of an anti-epidermal growth factor receptor (10-fold) required a triple mutation [31], whereas the optimization of the antibody Y0101 for the antigen VEGF (9-fold) is achieved by incorporating a total of six mutations [32]. Similarly, other examples of antibody engineering not employing electrostatic optimization also required multiple mutations as described in two separate studies (between three and fourteen mutations) [33,34]. In contrast, our best design achieves a considerable increase of affinity for a matured antibody (5-fold) but requiring only a single-mutation, an approach less likely to alter the three-dimensional structure of the antibody. We have not examined two or more simultaneous mutations of 11K2, an approach that could yield an optimized antibody displaying even higher affinity than that of the single-muteins generated herein.

The enhanced affinity of the optimized 11K2 antibodies is correlated with slower dissociation rate constants (k_{off}) rather than faster association rate constants (k_{on}). On the contrary, the destabilizing mutations accelerate the dissociation step, indicating that k_{off} is a valid parameter to predict the effect of the mutation on the affinity of this particular antibody-antigen complex. A previous report suggested that the dissociation step is a first order

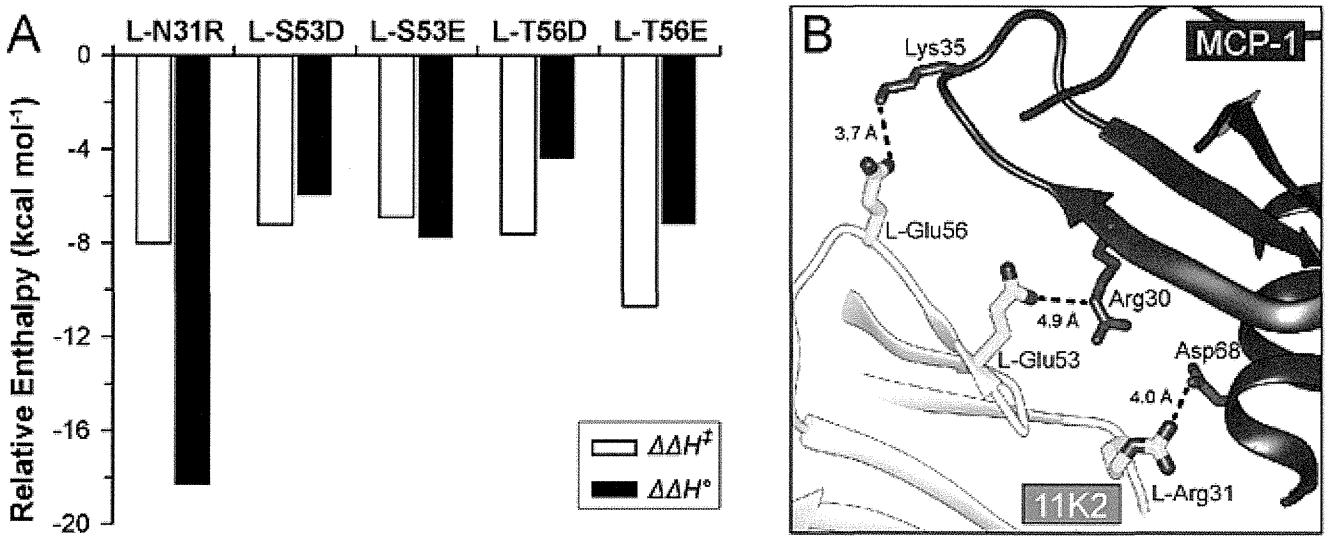

Figure 7. Analysis of the binding enthalpy. (**A**) Favorable changes of binding enthalpy with respect to wild-type antibody at the transition state (empty bar, $\Delta\Delta H^{\ddagger}$) and at equilibrium (filled bar, $\Delta\Delta H^{\circ}$). (**B**) Suggested model of the new interactions formed at the antibody/antigen contact surface upon mutation. Residues depicted in yellow and dark green correspond to 11K2 and MCP-1, respectively. The conformation of the side-chain of the mutated residues was modeled from the Dunbrak library of rotamers [44] as implemented in the program Chimera [45] (the most probably rotamer was always selected, except in L-Arg31, where the second most probable rotamer was chosen). Because Lys35 of MCP-1 is not interacting with a neighboring residue in the crystal structure, the conformation of this residue was also modeled as above. The dotted lines and distances represent putative interactions between the mutated residues and the antigen.

doi:10.1371/journal.pone.0087099.g007

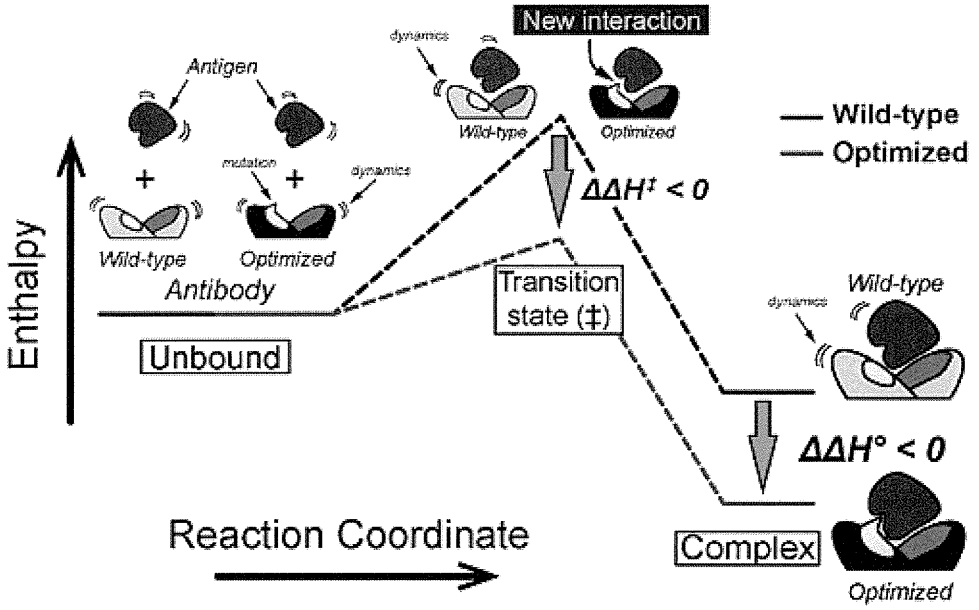

Figure 8. Energetic model of the optimization of an antibody. The diagram represents the enthalpic optimization of an antibody displaying higher affinity for its antigen. The enthalpic advantage acquired during the transition state persists in the complex at equilibrium, although it is largely counterbalanced by the unfavorable contribution of the change of entropy (entropy/enthalpy compensation).
doi:10.1371/journal.pone.0087099.g008

reaction whose rate is dictated by the strength of short range interactions between the proteins (van der Waals forces, hydrogen bond, hydrophobic effect, and salt bridges) [35]. Our results convincingly demonstrate that introducing a charged residue in 11K2 increase the affinity for the antigen and slows down k_{off}.

The examination of the thermodynamic parameters (Table 2) indicates that the optimization of the affinity of the antibody-antigen complex is the result of beneficial contributions of enthalpic nature originating in the transition state ($\Delta\Delta H^{\ddagger}<0$). It is important to understand the underlying mechanism in terms of the reaction coordinate diagram or "interaction pathway" (Figure 8) [36–39]. First, antigen and antibody collide with each other forming the so-called encounter complex (μsec scale) leading to a weakly interacting complex in which the hydration of the protein surface is not altered significantly (not shown) [36,40]._EN-REF_35 The encounter complex progress towards the transition state, although with low efficiency [41]. In the transition state, the antigen-antibody partners lose their hydration layers at the interaction surfaces, and a reconfiguration of intra- and inter-molecular forces takes place. Our results indicate that the beneficial effect of the mutations to the enthalpy in the transition state ($\Delta\Delta H^{\ddagger}<0$) is also preserved in the equilibrium complex ($\Delta\Delta H^{\circ}<0$). In other words, the electrostatic interactions generated in the optimized muteins are actively contributing to binding early in the reaction coordinate, perhaps as early as in the encounter complex, which is a state particularly sensitive to long-range coulombic forces [36]. The new and favorable non-covalent interactions of the optimized muteins are not entirely translated into higher affinity of a similar energetic magnitude, since the entropy/enthalpy compensation effect reduces the influence of the enthalpy advantage in the free energy [42,43]. Our data indicates that ultimately the enthalpy advantage prevails over the entropic disadvantage in the final antigen-antibody complex albeit with low efficiency since only 5% of the favorable enthalpy is *converted* in useful free energy (Figure S4).

In summary, we illustrate the benefits of using computational methods to design an optimized antibody with higher binding affinity for its cognate antigen. In particular, the incorporation of charged residues improve the affinity by a combination of favorable enthalpic contributions early in the transition state (or even earlier in the encounter complex), and slower dissociation rates (k_{off}). We hope our study will encourage further investigations into the fundamental physicochemical basis of protein-protein interactions.

Supporting Information

Figure S1 Representative examples of energy distribution histograms corresponding to the interaction between 11K2 and MCP-1. The figure shows the histograms of the wild type (black dotted), the favorable mutein L-N31R (solid, red), and a neutral mutein L-T52S (solid, blue). Muteins displaying favorable histograms with respect to wild-type antibody (i.e. shifted towards the left in the figure above) were selected for additional analysis.
(TIF)

Figure S2 SDS-PAGE of insoluble fraction of 11K2 scFv.
(TIF)

Figure S3 Decomposition of energy terms along the reaction pathway. The data are given in Table 3 of the manuscript. In each plot, the three energetic levels correspond to the free antibody and antigen, the transition state, and the antibody/antigen complex.
(TIF)

Figure S4 Entropy/enthalpy compensation plot. (**A**) Values at equilibrium. (**B**) Values at the transition state.
(TIF)

Table S1 DNA and primary sequence of MCP-1 and scFv-11K2.
(PDF)

Acknowledgments

We are grateful to Professor Haruki Nakamura for the critical reading of the manuscript and helpful advice. We thank to Shinya Tashiro and Mio Okamoto (The University of Tokyo) for their assistance with data collection (Biacore) and the preparation of the antibody constructs.

Author Contributions

Conceived and designed the experiments: MK MN SS HS SK KT. Performed the experiments: MK EM MN SS. Analyzed the data: MK JMMC SN SS HS SK KT. Contributed reagents/materials/analysis tools: MK EM MN SS HS SK KT. Wrote the paper: MK JMMC SN HS KT.

References

1. Adair JR (1992) Engineering antibodies for therapy. Immunol Rev 130: 5–40.
2. Lawson AD (2012) Antibody-enabled small-molecule drug discovery. Nat Rev Drug Discov 11: 519–525.
3. Muda M, Gross AW, Dawson JP, He C, Kurosawa E, et al. (2011) Therapeutic assessment of SEED: a new engineered antibody platform designed to generate mono- and bispecific antibodies. Protein Eng Des Sel 24: 447–454.
4. Beck A (2011) Biosimilar, biobetter and next generation therapeutic antibodies. MAbs 3: 107–110.
5. Kawa S, Onda M, Ho M, Kreitman RJ, Bera TK, et al. (2011) The improvement of an anti-CD22 immunotoxin: conversion to single-chain and disulfide stabilized form and affinity maturation by alanine scan. MAbs 3: 479–486.
6. Constantinou A, Epenetos AA, Hreczuk-Hirst D, Jain S, Deonarain MP (2008) Modulation of antibody pharmacokinetics by chemical polysialylation. Bioconjug Chem 19: 643–650.
7. Hagihara Y, Saerens D (2012) Improvement of single domain antibody stability by disulfide bond introduction. Methods Mol Biol 911: 399–416.
8. Renaut L, Monnet C, Dubreuil O, Zaki O, Crozet F, et al. (2012) Affinity maturation of antibodies: optimized methods to generate high-quality ScFv libraries and isolate IgG candidates by high-throughput screening. Methods Mol Biol 907: 451–461.
9. Acchione M, Kwon H, Jochheim CM, Atkins WM (2012) Impact of linker and conjugation chemistry on antigen binding, Fc receptor binding and thermal stability of model antibody-drug conjugates. MAbs 4: 362–372.
10. Shiroishi M, Tsumoto K, Tanaka Y, Yokota A, Nakanishi T, et al. (2007) Structural consequences of mutations in interfacial Tyr residues of a protein antigen-antibody complex. The case of HyHEL-10-HEL. J Biol Chem 282: 6783–6791.
11. Barderas R, Desmet J, Timmerman P, Meloen R, Casal JI (2008) Affinity maturation of antibodies assisted by in silico modeling. Proc Natl Acad Sci U S A 105: 9029–9034.
12. Park H, Jeon YH (2011) Free energy perturbation approach for the rational engineering of the antibody for human hepatitis B virus. J Mol Graph Model 29: 643–649.
13. Kuroda D, Shirai H, Jacobson MP, Nakamura H (2012) Computer-aided antibody design. Protein Eng Des Sel 25: 507–521.
14. Kusano KF, Nakamura K, Kusano H, Nishii N, Banba K, et al. (2004) Significance of the level of monocyte chemoattractant protein-1 in human atherosclerosis. Circ J 68: 671–676.
15. Tominaga T, Miyazaki D, Sasaki S, Mihara S, Komatsu N, et al. (2009) Blocking mast cell-mediated type I hypersensitivity in experimental allergic conjunctivitis by monocyte chemoattractant protein-1/CCR2. Invest Ophthalmol Vis Sci 50: 5181–5188.
16. Hayashida K, Nanki T, Girschick H, Yavuz S, Ochi T, et al. (2001) Synovial stromal cells from rheumatoid arthritis patients attract monocytes by producing MCP-1 and IL-8. Arthritis Res 3: 118–126.
17. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, et al. (2000) The Protein Data Bank. Nucleic Acids Res 28: 235–242.
18. Reid C, Rushe M, Jarpe M, van Vlijmen H, Dolinski B, et al. (2006) Structure activity relationships of monocyte chemoattractant proteins in complex with a blocking antibody. Protein Eng Des Sel 19: 317–324.
19. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. J Mol Biol 372: 774–797.
20. Kabat E, Wu TT, Perry H, Gottesman K, Foeller C (1991) Sequences of proteins of immunological interest. Bethesda, National Institutes of Health, 2719 p.
21. Chothia C, Lesk AM (1987) Canonical structures for the hypervariable regions of immunoglobulins. J Mol Biol 196: 901–917.
22. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, et al. (2006) Comparative protein structure modeling using Modeller. Curr Protoc Bioinformatics Chapter 5: Unit 5.6.
23. Nagao C, Izako N, Soga S, Khan SH, Kawabata S, et al. (2012) Computational design, construction, and characterization of a set of specificity determining residues in protein-protein interactions. Proteins 80: 2426–2436.
24. Umetsu M, Tsumoto K, Hara M, Ashish K, Goda S, et al. (2003) How additives influence the refolding of immunoglobulin-folded proteins in a stepwise dialysis system. Spectroscopic evidence for highly efficient refolding of a single-chain Fv fragment. J Biol Chem 278: 8979–8987.
25. Tsumoto K, Shinoki K, Kondo H, Uchikawa M, Juji T, et al. (1998) Highly efficient recovery of functional single-chain Fv fragments from inclusion bodies overexpressed in Escherichia coli by controlled introduction of oxidizing reagent—application to a human single-chain Fv fragment. J Immunol Methods 219: 119–129.
26. Morton TA, Myszka DG (1998) Kinetic analysis of macromolecular interactions using surface plasmon resonance biosensors. Methods Enzymol 295: 268–294.
27. Ross PD, Subramanian S (1981) Thermodynamics of protein association reactions: forces contributing to stability. Biochemistry 20: 3096–3102.
28. Sakamoto S, Caaveiro JM, Sano E, Tanaka Y, Kudou M, et al. (2009) Contributions of interfacial residues of human Interleukin15 to the specificity and affinity for its private alpha-receptor. J Mol Biol 389: 880–894.
29. Cleland WW, Northrop DB (1999) Energetics of substrate binding, catalysis, and product release. Methods Enzymol 308: 3–27.
30. Lutgens E, Faber B, Schapira K, Evelo CT, van Haaften R, et al. (2005) Gene profiling in atherosclerosis reveals a key role for small inducible cytokines: validation using a novel monocyte chemoattractant protein monoclonal antibody. Circulation 111: 3443–3452.
31. Lippow SM, Wittrup KD, Tidor B (2007) Computational design of antibody-affinity improvement beyond in vivo maturation. Nat Biotechnol 25: 1171–1176.
32. Marvin JS, Lowman HB (2003) Redesigning an antibody fragment for faster association with its antigen. Biochemistry 42: 7077–7083.
33. Clark LA, Boriack-Sjodin PA, Eldredge J, Fitch C, Friedman B, et al. (2006) Affinity enhancement of an in vivo matured therapeutic antibody using structure-based computational design. Protein Sci 15: 949–960.
34. Midelfort KS, Hernandez HH, Lippow SM, Tidor B, Drennan CL, et al. (2004) Substantial energetic improvement with minimal structural perturbation in a high affinity mutant antibody. J Mol Biol 343: 685–701.
35. Selzer T, Albeck S, Schreiber G (2000) Rational design of faster associating and tighter binding protein complexes. Nat Struct Biol 7: 537–541.
36. Schreiber G, Haran G, Zhou HX (2009) Fundamental aspects of protein-protein association kinetics. Chem Rev 109: 839–860.
37. Northrup SH, Boles JO, Reynolds JC (1988) Brownian dynamics of cytochrome c and cytochrome c peroxidase association. Science 241: 67–70.
38. Schreiber G (2002) Kinetic studies of protein-protein interactions. Curr Opin Struct Biol 12: 41–47.
39. Volkov AN, Worrall JA, Holtzmann E, Ubbink M (2006) Solution structure and dynamics of the complex between cytochrome c and cytochrome c peroxidase determined by paramagnetic NMR. Proc Natl Acad Sci U S A 103: 18945–18950.
40. James LC, Tawfik DS (2005) Structure and kinetics of a transient antibody binding intermediate reveal a kinetic discrimination mechanism in antigen recognition. Proc Natl Acad Sci U S A 102: 12730–12735.
41. Lipschultz CA, Yee A, Mohan S, Li Y, Smith-Gill SJ (2002) Temperature differentially affects encounter and docking thermodynamics of antibody--antigen association. J Mol Recognit 15: 44–52.
42. Chodera JD, Mobley DL (2013) Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design. Annu Rev Biophys 42: 121–142.
43. Baron R, McCammon JA (2013) Molecular recognition and ligand association. Annu Rev Phys Chem 64: 151–175.
44. Dunbrack RL Jr (2002) Rotamer libraries in the 21st century. Curr Opin Struct Biol 12: 431–440.
45. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, et al. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem 25: 1605–1612.

特集　ATL/HTLV-1研究の最近の進展

miRNAを用いた成人T細胞白血病(ATL)がん幹細胞を標的とした新規治療法開発研究の現状*

渡邉俊樹**

Key Words：cancer stem cel, scFv, miRNA, NF-κB, epigenetics

はじめに

成人T細胞白血病(ATL)は50～60年という長い潜伏期間中にHTLV-1感染末梢血T細胞に複数の遺伝子異常が蓄積することによる多段階発がん機構によって発症する．現在，日本は少なくとも約110万人の感染者が存在し毎年1,000人を超えるATLの発症者/死亡者が報告されている．白血病やウイルスの発見から30年以上になるが，ウイルスによる細胞の不死化や腫瘍化，治療抵抗性などの分子機構は不明な点が多く残されており，有効な治療法はいまだに存在しない．現在の標準的治療法である多剤併用化学療法の成績は，aggressive type ATL(急性型およびリンパ腫型)の生存期間中央値は12か月前後である．血液幹細胞移植療法も臨床研究が続けられているが，ATL患者が高齢であること，移植関連合併症が多いことなどから，骨髄破壊的前処置を行った移植患者の長期生存率は30%台にとどまる．このような現状から，原因ウイルスであるHTLV-1の根絶と，白血病発症予防および新規治療法開発の開発が強く求められており，それを可能にするために，ATL細胞の分子レベルでの病態解明が期待されている[1]．

ATL細胞およびHTLV-1感染細胞の生物学的特徴として，恒常的なNF-κBシグナル経路の活性化があり，これが細胞の異常な増殖および生存の基盤となっている[2][3]．HTLV-1感染細胞ではウイルスタンパク質TaxがNF-κBの定型的(canonical)および非定型的(noncanonical)経路を劇的に活性化するが，Taxの発現が認められないATL腫瘍細胞におけるNF-κBの活性化機構は不明であった．NF-κBシグナルの異常な活性化とそれに伴う腫瘍細胞の生存能の獲得と悪性化は，ATLだけでなく多くの固形がんや悪性リンパ腫，白血病で共通してみられるがん細胞の特徴的な分子病態の1つである[3][4]．

固形がんを含めたがんの治療抵抗性・薬剤耐性の背景として，「がん幹細胞」の存在が注目されている．いくつかの白血病や固形がんにおいては「がん幹細胞」の存在が報告され，その細胞学的な特徴が検討されている[5]．しかし，容易に薬剤耐性を獲得し治療抵抗性を示すATLの「がん幹細胞」の探索などの系統的研究はほとんど行われていない．

われわれは，ATLのmiRNA，mRNA発現解析およびゲノム異常の解析を統合してATLの分子病態の全貌を明らかにすることを目指してきた．その結果，NF-κB恒常的活性化の分子機構として，

* Attempt to develop scFv as DDS for miRNA delivery to ATL cancer stem cell.

** Toshiki WATANABE, M.D., Ph.D.: 東京大学大学院新領域創成科学研究科メディカルゲノム専攻病態医療科学分野〔〒108-8639 東京都港区白金台4-6-1 東京大学医科学研究所2号館3階〕; Laboratory of Tumor Cell Biology, Department of Medical Genome Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo 108-8639, JAPAN

ポリコームPRC2の作用に基づくmiR-31の発現欠損とそれに伴うNIKの過剰発現があることを見出した[6]．また，ATL細胞へmiR-31を補充するとNIKの発現低下とNF-κB活性化の阻害と細胞死の誘導をもたらすことを示した．本稿では，ATL細胞を特徴づけるNF-κB活性化機構に関する知見を整理し，われわれが進めているATLの「がん幹細胞」探索の研究の進捗状況を紹介するとともに，最後に，miR-31を腫瘍細胞特異的に導入することを目指す新たな薬物伝達法(drug delivery system, DDS)として，特異性，低分子性，製造の簡便性を考慮し，単鎖抗体(scFv)を利用した新たな治療法開発の現状を紹介する．

ATL細胞を特徴づける NF-κBの恒常的活性化とその意義

ATL細胞におけるNF-κBシグナルの活性化は，森らによって1999年に最初に報告された[7]．この報告では，HTLV-1感染T細胞株および患者由来ATL細胞で，HTLV-1 Taxの発現の有無にかかわらず，全例で活性化しており，核内のDNA結合活性はp50とp65およびc-Relによるものと報告された．

HTLV-1 TaxはIKKγ(NEMO)との結合を介して強力なNF-κB活性化作用を示す．しかし生体内のATL細胞ではTaxが存在しないため，これとは異なる活性化機構の存在が想定される．実際，山岡らはATL細胞でのTax非依存的IKK活性化を報告した[8]．その後，筆者らは，2005年に，個体内のATL細胞および非腫瘍化HTLV-1感染細胞におけるNF-κBの活性化の評価と新規NF-κB阻害剤DHMEQの作用を解析した[9]．さらに，EMSAによる解析で，ATL由来細胞株における活性化NF-κBはp50/p65およびp52/RelBからなり，新鮮ATL細胞ではp50/p65であることを明らかにした．つまり，ATL細胞株ではclassical pathwayとalternative pathwayの双方の経路が活性化している．この時の解析では新鮮ATL細胞でのp52/RelBの活性化は確認できなかった．その後，山岡らはATL細胞におけるNIKの過剰発現とそれによるNIKの活性とNF-κB経路の活性化を示した[10]．

Dewanらは免疫不全マウス(NOG-mouse)へ移植されたHTLV-1感染細胞株を用いて実験を行い，これらの細胞株が生着することを示すとともに，IKK阻害剤Bay 11-7082が強い抗腫瘍活性を持つことを示した[11]．筆者らは，先に紹介した論文[9]で，NF-κBの核移行阻害剤DHMEQがATL細胞に対して強いNF-κB阻害効果を示し，アポトーシスによる細胞死を誘導すること，免疫不全マウスの移植系でDHMEQが著明な救命効果を示すことを報告した．さらに，キャリア個体内の非腫瘍化HTLV-1感染細胞でもNF-κBの活性化が認められ，DHMEQはこれらに対して選択的に細胞傷害性を示すことを明らかにした．

ATL細胞の大規模統合解析から 明らかになったNF-κB活性化の 新たな分子機構

残された疑問は，ATL細胞におけるNIKの過剰発現の分子機構である．この課題に取り組むため，筆者らはまず，全国的な共同研究組織JSPFAD (http://htlv1.org/)の全面的な協力を得て，その検体バンクを利用してATL細胞のゲノムコピー数異常(CNV)解析，および，遺伝子発現解析(miRNAおよびmRNA)を行い，結果をデータベースに登録した〔accession numbers, GSE31629(miRNA microarray)，GSE33615(gene expression microarray)，and GSE33602(copy number analyses)〕．その中で，ATL 40例，正常$CD4^+$ T細胞22例についてmiRNAマイクロアレイで解析した結果，miR-31が全例において発現抑制(ほとんど欠損)していることが明らかとなった．

miRNAの主な生物学的機能は，標的遺伝子の3′UTRへの結合を介して遺伝子発現を負に制御することである．外来の合成siRNAと異なり，miRNAの配列認識は揺らぎがあり，1つのmiRNAが複数個の遺伝子を制御することができる(Pleiotropic function)．細胞の運命に重大な影響を与える標的の遺伝子の探索には，物理的な抑制効果と同時に標的側の機能や挙動も重要な指標となり，したがって多角的な実験的検証が必須になる．筆者らは，miR-31のT細胞における生物学的意義を明らかにするために，以下の検討を行った．

① miR-31の標的遺伝子の予測を4つのアルゴリズムによって予測．

② ATL細胞のmRNA大規模解析データとの擦り

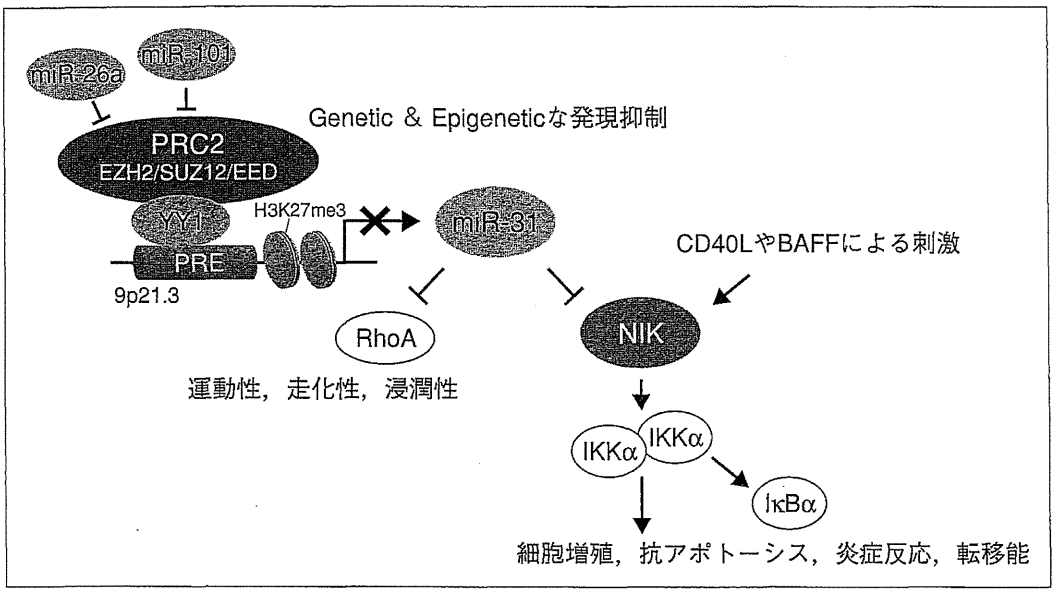

図1　Polycomb-miR-31-NF-κBのリンケージ
Polycomb依存的なmiR-31の発現低下はNIKなどの標的遺伝子を介して細胞の表現型に影響する．この分子間の関係はさまざまな細胞で保存されており，各因子の存在量のバランスによって均衡が保たれている．バランスを崩した細胞は悪性化をたどると考えられる．

合わせによる検証．

③変異を導入したレポーターアッセイ．

④miR-31の増減に対する標的候補遺伝子の定量．

⑤miR-31と標的遺伝子の関係の保存性．

これらの検討の結果，NIKがmiR-31の新規標的遺伝子であることを見出した．NIKはNF-κBの非定型的経路の活性化に必須のリン酸化酵素であり，NIKの発現レベルがNF-κBシグナルの恒常的活性化に直接的に寄与する．ATLでのNIKの過剰発現が報告されていたが[10]，過剰発現の分子機構は不明であった．筆者らはmiR-31発現低下→NIK過剰発現→NF-κB経路活性化という関係をはじめて証明した．さらに，miR-31の発現低下には，PRC2によるH3K27のトリメチル化が関与していることを明らかにした[6]．この知見は世界で初めて，エピジェネティクス(PRC2)－microRNA－シグナル伝達(NF-κB)のリンケージの存在を明らかにしたものである(図1)．さらに，HTLV-1感染T細胞株や，ex vivoのATL細胞にmiR-31を導入すると，NIKの発現低下，NF-κBの活性化の抑制，さらにATL細胞のアポトーシスが誘導された．これらの結果は，miR-31が新規核酸医薬としての可能性を持つ分子であることを示している．

ATL「がん幹細胞」探索研究の現状

いわゆる「がん幹細胞」は，以下の条件を満たす細胞集団として定義される[12]．

①腫瘍原性のある少数の細胞集団．

②免疫不全マウスに連続移植可能．

③腫瘍原性のない細胞とマーカーで区別できる．

④その細胞が分裂・増殖した際に，再び腫瘍原性のあるものとないものからなる集団を生じる．

われわれは，まず免疫不全マウスへの連続移植系を確立することから実験を開始した．感染研の浜口らのグループは，Taxトランスジェニックマウス(Tax-Tg)に発生するATL様の腫瘍で「がん幹細胞」を同定して報告した[13]．筆者らはその解析結果を参考にして，ヒトのATL細胞の移植系を検討した．その結果，広く使用されているNOG SCIDマウスよりも，NOD/SCID/Jak3KO(NOJ)マウス(NOJ SCID)が新鮮ATL細胞の移植実験に有利であることを見出し，このマウスを用いることにした(熊本大学岡田教授より供与)．その結果，ATLマーカーであるCD4とCCR4でソートした細胞集団からNOJ SCIDマウスにおいて連続継代移植可能な細胞集団を得ることができた．

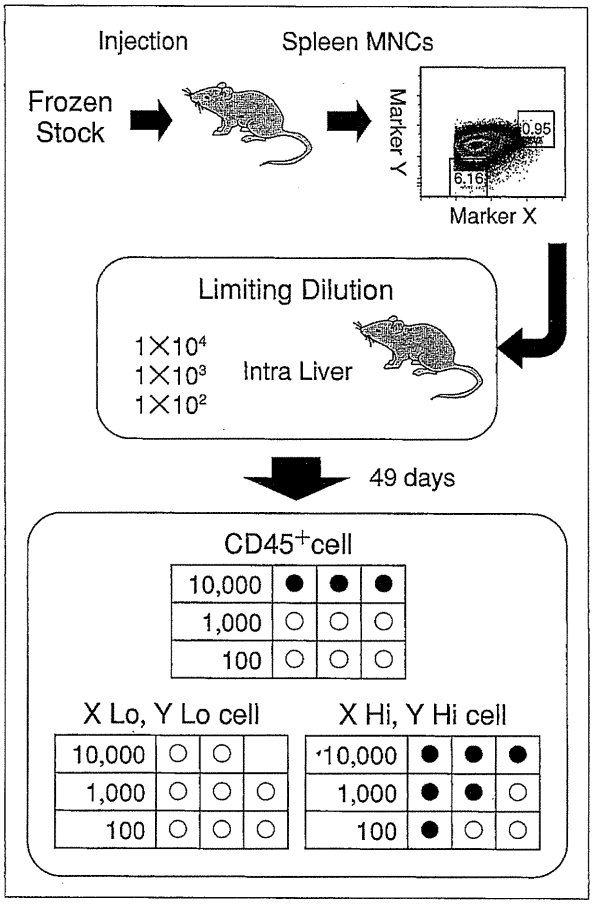

図2　がん幹細胞候補マーカーのスクリーニング
患者ATL検体の移植系と表面マーカーによるソーティング，および限界希釈法(limiting dilution)によるATL「がん幹細胞」の検証結果(予備実験のデータ)．患者検体の移植に際して，CD3，CCR4などのマーカーであらかじめソーティングすることで著明に生着率が向上し，連続継代移植が可能になった．予備実験で班，特定の表面抗原の発現レベルをもとに細胞を分画して移植したところ，少なくとも100個の移植細胞数で腫瘍形成が可能であることが示された．

この系を用いてフローサイトメーターによるATL患者細胞ならびに健常人細胞における細胞表面抗原の網羅的解析を行った．患者末梢血ATL細胞内には免疫不全マウスで造腫瘍性を示し，8代以上にわたって連続移植可能な細胞集団が存在すること，生着した腫瘍細胞の表面マーカー発現レベルを指標に細胞を分画することにより，少数(500個以下)で造腫瘍性を示す分画が存在すること，表面抗原は移植を繰り返すことに伴い変遷を示すこと，したがって，「がん幹細胞」集団を既存のマーカーで容易に定義できないこと，などである[14]．これらの予備的解析結果の一部を図に示す(図2)．これらを踏まえて，筆者らは「がん幹細胞」を同定するマーカーの探索を継続している．

単鎖抗体を利用したmiR-31のATL細胞特異的導入法開発の現状

ATLの特性の1つは，臓器浸潤とリンパ節での増殖である．したがって目的のmiRNAやtoxinをATLに届ける場合には，低分子量の担体の方が好ましい．筆者らが計画する単鎖抗体(scFv)は，その特性として，完全抗体(IgG)と同様の特異性を保持しつつ，より細部まで届くことが知られている．さらにscFvをsiRNAと結合させて導入する実験の結果も報告されている．しかし，ATLに関してはscFvによる治療研究はいまだにない．筆者らの研究計画の概略を模式的に図に示す(図3)．これまでに，モデル実験として，ATL細胞に発現している抗原からOX40およびCD5を選択し，そのモノクローナル抗体を産生するハイブリドーマを入手して，以下の手順で実験を進めてきた．

1．可変領域の配列決定と発現ベクターの作成

単鎖抗体(scFv)作成に必要な重鎖(VH)と軽鎖(VL)の可変領域の塩基配列は5′RACE法を用いてcDNAを増幅して解析した．scFvの簡便かつ大量製造のため，大腸菌発現系によるリコンビナントscFv大量精製から開始した．さらに核酸結合用の融合カチオン性ペプチド候補は，アルギニンやリジン残基を多く含みRNA結合能が知られているペプチド(9 mer Arg, protamine)を採用して発現ベクターを設計した．

2．抗原分子の塩基配列決定と発現ベクター作成

抗体との相互作用解析に用いる細胞表面抗原の塩基配列を正常ヒト末梢血単核球とT細胞株，ATL感染細胞，ATL由来細胞株のtotal RNAを用いたRT-PCR法でcDNAを増幅して決定した．得られた細胞外ドメインの塩基配列を用いて，N末端からヒスチジンタグ，TEVプロテアーゼ認識配列，抗原の順に配置し発現ベクターを作成した．

3．大腸菌発現システムを用いた発現と精製

発現したscFvは，可溶性画分からの精製が可能なものはNiアフィニティーカラムクロマトグ