# Antibody modeling assessment II: Structures and Models

**ARTICLE** *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · AUGUST 2014

Impact Factor: 2.63 · DOI: 10.1002/prot.24554



**9 AUTHORS**, INCLUDING:

Alexey Teplyakov
Johnson & Johnson
**93** PUBLICATIONS **5,706** CITATIONS
SEE PROFILE

Jinquan Luo
Johnson & Johnson
**27** PUBLICATIONS **341** CITATIONS
SEE PROFILE

Juan Carlos Almagro
GlobalBio, Inc.
**46** PUBLICATIONS **1,086** CITATIONS
SEE PROFILE

Gary L. Gilliland
Janssen Research & Development, LLC, Sprin…
**218** PUBLICATIONS **29,072** CITATIONS
SEE PROFILE

PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

# Antibody modeling assessment II. Structures and models

Alexey Teplyakov,[1]* Jinquan Luo,[1] Galina Obmolova,[1] Thomas J. Malia,[1] Raymond Sweet,[1] Robyn L. Stanfield,[2] Sreekumar Kodangattil,[3] Juan Carlos Almagro,[3] and Gary L. Gilliland[1]

[1] Janssen Research & Development, LLC, 1400 McKean Road, Spring House, Pennsylvania, 19477

[2] Department of Integrative Structural and Computational Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California, 92037

[3] Pfizer's Center for Therapeutic Innovation, 3 Blackfan Cir., Boston, Massachusetts, 02115

ABSTRACT

**To assess the state-of-the-art in antibody structure modeling, a blinded study was conducted. Eleven unpublished Fab crystal structures were used as a benchmark to compare Fv models generated by seven structure prediction methodologies. In the first round, each participant submitted three non-ranked complete Fv models for each target. In the second round, CDR-H3 modeling was performed in the context of the correct environment provided by the crystal structures with CDR-H3 removed. In this report we describe the reference structures and present our assessment of the models. Some of the essential sources of errors in the predictions were traced to the selection of the structure template, both in terms of the CDR canonical structures and VL/VH packing. On top of this, the errors present in the Protein Data Bank structures were sometimes propagated in the current models, which emphasized the need for the curated structural database devoid of errors. Modeling non-canonical structures, including CDR-H3, remains the biggest challenge for antibody structure prediction.**

Proteins 2014; 82:1563–1582.


**Key words: structure prediction; crystal structure; canonical structure; VL/VH packing; CDR.**

## INTRODUCTION

The use of monoclonal antibodies in clinical diagnostics and as therapeutics has prompted the development of antibody modeling tools.[1–3] Humanization of non-human antibodies and optimization of antibody properties such as specificity, affinity, and solubility are just a few of the numerous applications of antibody engineering. Rational design of antibodies requires the detailed knowledge of the three-dimensional (3D) structure, which may be obtained experimentally but is often generated by homology modeling. Homology modeling uses 3D structures of antibodies with similar sequences as "templates." Four steps are generally required to build a homology model: template selection, template–target sequence alignment, model building, and model evaluation.[4] There are a number of sequence alignment tools and protein structure databases to fulfill the task including the Protein Data Bank (PDB)[5] with more than 1000 crystal structures of antibody fragments (Fab or Fv).

The antigen-binding sites of antibodies (Ab) are formed by six loops, three from the variable domain of the light chain (VL) and three from the variable domain of the heavy chain (VH).[6] The framework (FW) regions that are outside these loops are highly conserved in both sequence and main-chain conformation, and they can be accurately predicted using homology modeling techniques.[7] Although the VL/VH interface (VL/VH) is relatively conserved, the orientation of VL with respect to VH can vary by as much as 30° from one antibody to another.[8] This variation has been proposed as an additional mechanism to increase the repertoire of antibody specificity.[9] Thus, the correct orientation of the variable domains is a prerequisite for the successful modeling of the antigen-binding site.

The six loops of the antigen-binding site, called "complementarity-determining regions (CDRs)," are hypervariable in sequence.[10] Despite the sequence diversity of the CDRs, five of them assume a limited number

Additional Supporting Information may be found in the online version of this article.

A. Teplyakov and J. Luo contributed equally to this work.

*Correspondence to: Alexey Teplyakov; Biotechnology Center of Excellence, Janssen R&D, 1400 McKean Road, Spring House, PA 19477. E-mail: ateplyak@its.jnj.com

Received 17 December 2013; Revised 23 February 2014; Accepted 6 March 2014
Published online 31 March 2014 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.24554

**Table I**
X-ray Data and Refinement Statistics

| Reference structure | Space gr. (mol/a.u.) | Resolution (Å) | Completeness (%) | Redundancy | Rmerge (%) | $<I/\sigma>$ | Rcryst (%) | Rfree (%) | Protein atoms | Non-protein atoms | PDB code |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Ab01 | C2 (2) | 2.0 (2.1–2.0) | 97.5 (85.3) | 3.3 (3.0) | 5.3 (32.7) | 16.1 (4.1) | 18.0 | 21.1 | 6516 | 393 | 4MA3 |
| Ab02 | $P3_121$ (1) | 2.7 (2.8–2.7) | 99.3 (93.0) | 11.0 (6.4) | 10.3 (47.7) | 20.7 (4.0) | 23.0 | 26.8 | 3321 | 81 | 4KUZ |
| Ab03 | C2 (1) | 1.9 (2.0–1.9) | 91.7 (55.0) | 4.9 (2.7) | 3.5 (15.0) | 32.8 (6.6) | 17.1 | 22.5 | 3251 | 373 | 4KQ3 |
| Ab04 | $P3_1$ (1) | 2.4 (2.5–2.4) | 98.5 (97.9) | 5.7 (5.3) | 5.8 (49.2) | 25.3 (4.3) | 21.0 | 27.9 | 3358 | 64 | 4KQ4 |
| Ab05 | $P2_12_12_1$ (1) | 2.0 (2.1–2.0) | 99.1 (90.1) | 13.3 (5.2) | 8.5 (45.5) | 26.9 (3.7) | 18.4 | 22.1 | 3228 | 256 | 4M6M |
| Ab06 | I23 (1) | 2.8 (2.9–2.8) | 95.7 (61.7) | 3.4 (1.5) | 8.6 (40.5) | 14.5 (1.8) | 18.6 | 25.0 | 3324 | 58 | 4M6O |
| Ab07 | $C222_1$ (1) | 1.9 (2.0–1.9) | 99.3 (94.4) | 13.5 (5.2) | 5.9 (24.6) | 37.9 (6.4) | 18.4 | 21.7 | 3281 | 441 | 4MAU |
| Ab08 | $C222_1$ (1) | 1.9 (2.0–1.9) | 99.4 (95.3) | 9.5 (4.8) | 6.0 (36.5) | 26.3 (4.2) | 20.5 | 24.3 | 3419 | 352 | 4M7K |
| Ab09 | $P2_12_12_1$ (1) | 2.1 (2.2–2.1) | 93.6 (89.8) | 4.9 (3.5) | 7.9 (27.8) | 16.8 (5.7) | 17.1 | 22.0 | 3278 | 398 | 4KMT |
| Ab10 | $P2_1$ (2) | 1.6 (1.7–1.6) | 96.0 (95.4) | 3.4 (2.1) | 4.6 (34.7) | 28.7 (1.9) | 19.0 | 23.0 | 6828 | 864 | 4M61 |
| Ab11 | $P2_12_12$ (1) | 1.8 (1.9–1.8) | 99.7 (99.6) | 4.5 (4.0) | 7.9 (51.1) | 16.4 (2.9) | 19.4 | 23.2 | 3320 | 165 | 4M43 |

Values in parentheses are for the highest resolution shell.

of main-chain conformations, termed "canonical structures."[6] These conformations are determined by the length of the loops and by the presence of key residues at specific positions either within the loops or in the FW regions. The specific pattern of residues that determines each canonical structure forms a "signature" whereby a canonical structure can be recognized in the sequence of an immunoglobulin of unknown structure and can therefore be predicted from sequence alone.[11]

The sixth loop, CDR-H3, is the most variable in sequence and conformation, and has not been fully categorized. Some predictions can be made for the "torso" region of CDR-H3 that includes 10 residues proximal to the FW.[12–16] The remaining "head" portion of the loop can be modeled ab initio, that is, without templates, by one of the standard methods of protein loop prediction.[17–19]

To assess the state-of-the-art in antibody structure modeling, a blinded study was conducted. Eleven unpublished Fab crystal structures covering a wide range of antigen-binding site conformations were used as a benchmark to compare Fv models generated by seven structure prediction methodologies. The crystal structures for the assessment, or the reference structures, were provided by Janssen R&D and by The Scripps Research Institute. The participants included Accelrys, Inc. (http://accelrys.com/), Chemical Computing Group (CCG) (http://www.chemcomp.com/), Jeffrey Gray's group at John Hopkins University (http://graylab.jhu.edu/), Hiroki Shirai's group at Osaka University and Astellas Pharma (http://sysimm.ifrec.osaka-u.ac.jp/kotaiab/), Macromoltek, LLC (https://www.macromoltek.com/), and Schrödinger, LLC (http://www.schrodinger.com/). In addition, the PIGS web-based server (http://cassandra.med.uniroma1.it/pigs/) was used to generate models using default parameters.[20] A description of the modeling methods is provided by the participants in the accompanying articles in this issue of *Proteins*.

In the first round, each of the participants submitted three non-ranked Fv models for each of the 11 reference antibodies. Only one model for each antibody was generated using PIGS. In the second round, the participants were provided with the reference structures without CDR-H3, so that the modeling of CDR-H3 can be performed in the context of the correct environment. Each participant submitted five non-ranked models for each antibody. PIGS was not involved in the second round. In this report, we describe the reference structures and present our analysis of the results of the modeling exercise. All models generated in this study are available on http://www.3dabmod.com.

## MATERIALS AND METHODS

### Selection of crystal structures

Eleven crystal structures of the Fab fragments of different antibodies were selected for the modeling exercise. The criteria for selection were the following: (1) there are no disordered regions in the CDRs and (2) the structure must be of sufficient resolution to unambiguously define the conformation of CDRs. The selected structures represented a variety of species (human, mouse, rabbit). One antibody (Ab05) had a lambda light chain, all other were kappa. The summary of the structures is given in Table I. All structures and X-ray data have been deposited in the PDB.

In most cases, the asymmetric unit of the crystal cell contained only one Fab molecule. In two cases, Ab01 and Ab10, there were two independent molecules and their structures were virtually identical, which is usually true for Fab structures. In other words, we assumed that the CDR conformations observed in the crystal structures are typically not affected by crystal contacts. In contrast, binding of antigen often causes an induced fit of the CDRs. Therefore, structures of Fabs present in complexes were excluded from consideration and only Fab structures in the unbound form were selected.

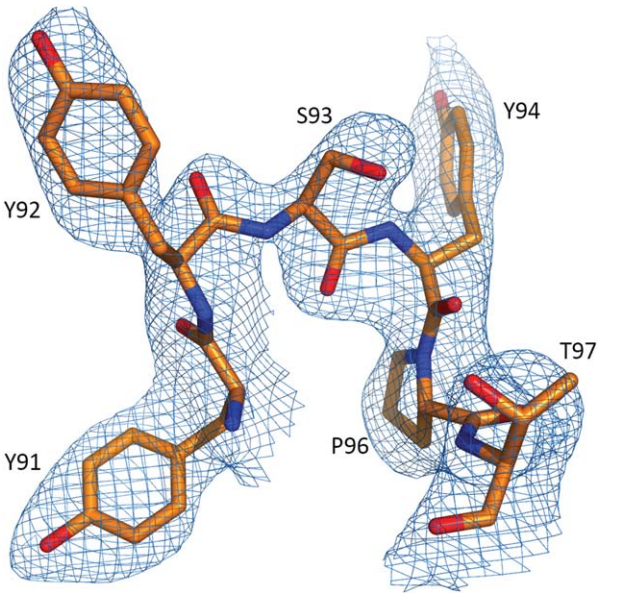


**Figure 1**

Electron density (2Fo-Fc contoured at 1.2 RMSD) for CDR-L3 of Ab02 determined at 2.7 Å resolution. This figure and figures with structural superpositions were prepared with PyMOL (DeLano Scientific, LLC).

Most structures were determined at high resolution. However, we included two medium-resolution structures (Ab02 and Ab06) since they met our criteria. Figure 1 shows a representative fragment of electron density for Ab02 CDR-L3 with clearly defined orientations of the peptide carbonyl groups.

## Canonical structures

The canonical structures for the reference antibodies were assigned following the nomenclature of North *et al.*[21] Their analysis of over 300 unique antibody structures identified the most common CDR conformations that can be described in terms of the main-chain angles φ and ψ. The notation specifies the canonical structure for a particular CDR of a given length, one being the most common, two being the second most common, and so on. For instance, H1-13-1 is the major (87%) canonical structure for CDR-H1 of the 13-residue length. Their definition of CDRs is somewhat different from any of the widely accepted definitions. With respect to the AbM definition,[22] three CDRs are longer on the N-terminal side, CDR-H1 and CDR-H3 by two residues and CDR-L2 by one residue (Fig. 2).

The canonical structure analysis for the CDRs of the 11 reference structures is presented in Table II. CDR-H3 was not classified due to its conformational variability. For the other five CDRs, the majority of the observed CDR conformations correspond to the most common canonical structures. However, 7 out of 11 reference structures contain at least one CDR that is not the most

```
VL
          10        20        30              40        50        60        70        80        90           100
      .    |    .    |    .    |abcdef    .    |    .    |    .    |    .    |    .    |    .abcd    |    . a
01  AQVLTQTPSSVSAAVGGTVTINCQASQSVYNK----NYLAWYQQKPGQPPKRLIYSASTLASGVSSRFKGSGSGTQFTLTISDVQCDDVATYYCLGSYDCNRAECHAFGGGTKVVV-EV
02  DIVMSQSPSSLAVSVGEKVTMSCKSSQSLLYSSNQKNYLAWYQQKPGQSPKLLIYWASTRESGVPDRFTGSGSGTDFTLTISSVKAEDLAVYYCQQYYSY-----PTFGAGTKLEL-K
03  DIQMTQSPSSVSASVGDRVTITCRASQGIS------NWLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQYSDD-----PTFGQGTKVEI-K
04  DIVMTQSHKFMSTSVGDRVSITCKASQDVS------TAVAWYQQKPGQSPKLLIYWASTRHTGVPDRFTGSGSGTDYTLTISSVQAEDLALYYCQQHYSTP----YTFGGGTKLEI-K
05  QSVLTQPPS-VSGAPGQRVTISCTGSSSNIGSG---YDVHWYQQLPGTAPKLLIYGNSKRPSGVPDRFSGSKSGTSASLAITGLQSEDEADYYCASWTDGLSL--VVFGGGTKLTVL-
06  AIQLTQSPSSLSASVGDRVTITCRASQGIS------SALAWYQQKPGKAPKLLIYDASSLESGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQFNSYP----LTFGGGTKVEI-K
07  DIVLTQSPASLAISLGQRATISCRASKSVSTSGS--SYMFWYQQKPGQPPKLLIYLASNLESGVPARFSGSGSGTDFTLNIHPVEEEDAAAYYCQHSREIP----YTFGGGTKLEI-K
08  DIVMTQSPSSLTVTAGEKVTMSCKSSQSLLSSGNQKNYLTWYQQIPGQPPKLLIYWASTRESGVPDRFTGSGSGTDFTLTINSVQAEDLAVYYCQNDYTYP----LTFGAGTKLEL-K
09  DIQMTQSPSSLSASVGDRVTITCRASQSIS------SYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYSTP----LTFGQGTKVEI-K
10  NIMMTQSPSSLAVSAGEKVTMSCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTRESGVPDRFTGSGSGTDFTLTISSVQAEDLAVYYCHQHLSS-----WTFGGGTKLEI-K
11  ELVMTQSPAILSVSPGERVSFSCRASQIIG------TSIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGTDFTLTINSVESDDIADYYCQQSNSWP----VTFGAGTKLEL-K
AbM                        ******************              *******                                  *************
NLD                        ******************              ********                                 *************
RMS --BBBBBBBBBBB----BBBBBBBB-LLLLLLLLLLLLBBBBBB----BBBBBBBLLLL-------BBBBBBB--BBBBBBB---------BBBBBBLLLLLLLLLLBBBBBBB------

VH
          10        20        30        40        50        60        70        80           90        100             110
      .    |    .    |    .    | a    .    |    .    | a  .    |    .    |    .    |  abc  .    |    .    |abcdef    .    |
01  -QSVEESGGRLVTPGTPLTLTCTVSGFSLSS-YPMNWVRQAPGKGLEWIGGIG-TSGNIWYASWAKGRFIISRAS-STTVDLKVTSPTTEDTATYFCARGLYNDY------TVWGPGTLVTVSS
02  EVQLQQSGPELVKPGASVKISCKASGYTFTD-YYMNWVKQSHGKSLEWIGGINPNNGGTSYNQKFKGKATLTVDKSSSTAYMELRSLTSEDSAVYYCARNDGYRGYAM---DYWGQGTSVTVSS
03  QVQLVQSGAEVKKPGSSVKVSCKASGGTFSS-YAISWVRQAPGQGLEWMGSIIPWFGTTNYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARDSEYYF------DHWGQGTLVTVSS
04  QVQLQQSGAELVRPGVSVRISCKGSGYTFTD-YAMHWVKQSHAKSLEWIGVISTYSGDARFNQKFTGKATMTVDKSSSTAYMELARLTSEDSAIYYCAREVRRSM------DYWGQGTSVTVSS
05  EVQLVQSGAEVKKPGESLKISCKGSGYSFSN-YWIGWVRQMPGKGLEWMGIIDPSNSYTRYSPSFQGQVTISADKSISTAYLQWSSLKASDTAMYYCARWYYKPF------DVWGQGTLVTVSS
06  EVQLVESGGGLVQPGGSLRLSCAASGFTLRS-YSMNWVRQAPGKGLEWVSYISRSSHTIFYADSVKGRFTISRDNAKNSLYLQMDSLRDEDTAMYYCARVYSSGWHVSDYFDYWGQGILVTVSS
07  EVQLQESGPGLVKPSQSLSLTCTVTGFSITSDYAWNWIRQFPGSKLEWMGFIS-YSGDTSFNPSLKSRISVTRDTSKNQFFLQLNSVTTEDTATYYCASYDGYSF------DYWGQGTTLTVSS
08  QVHLQQSGAELMKPGASVKISCKASGYTFIT-YWIEWVKQRPGHGLEWIGDILPGSGSTNYNENFKGKATFTADSSSNTAYMQLSSLTSEDSAVYYCARSGYYGNSGF---AYWGQGTLVTVSA
09  EVQLVQSGAEVKKPGESLKISCKGSGYSFTS-YWIGWVRQMPGKGLEWMGIIYPGDSDTRYSPSFQGQVTISADKSISTAYLQWSSLKASDTAMYYCARYDGIYGEL----DFWGQGTLVTVSS
10  EIQLQQSGAELVKPGASVKISCKASGYSFTG-YNMNWVKQSHGKSLEWIGKINPYYGSTSYNQKFKGQATLTVDKSSSTAYMQLNSLTSEDSAVYYCARGRLRRGGYF---DYWGQGTTLTVSS
11  EVQLQQFGAELVKPGASVKISCKASGYTFTD-YNMDWVKQSHGKSLQWIGDISPYYGSTGYSQKFKGKATLTVDRSSSTAYMELRSLTSEDTAVYYCARRNYDGSWF----AYWGQGTLVTVSS
AbM                        ***********             **********                                  **************
NLD                      **************            **********                                ****************
RMS --BBBBB----------BBBBBBB-LLLLLLLLLBBBBBBB---BBBBBBBBLLLLLBBBB-------BBBBB-LLLLBBBBBBB------BBBBBBBBLLLLLLLLLLLL-BBBBBBBBB---
```

**Figure 2**

Amino acid sequence alignment of the reference antibodies (variable domains). Residue numbering is Chothia's. CDR definitions according to AbM[22] and NLD[21] are indicated by stars. The line at the bottom (RMS) shows the regions used for calculation of RMSDs, "B" for the β-core, "L" for CDR loops.

**Table II**
Canonical Structures and Germlines in Reference Antibodies

| Ab | CDR-L1 | CDR-L2 | CDR-L3 | CDR-H1 | CDR-H2 | CDR-H3 length | Ab source | GL(VL) | PDB(VL) | GL(VH) | PDB(VH) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Ab01** | L1-13-NC | L2-8-1 | L3-13-NC | H1-13-1 | H2-9-1 | 10 | rabbit | IGKV1S41*01 | 2X7L(68) | IGHV1S69*01 | 2X7L(71) |
| **Ab02** | L1-17-1 | L2-8-1 | L3-8-2 | H1-13-1 | H2-10-1 | 13 | mouse | IGKV8-30*01 | 3MBX(95) | IGHV1-26*01 | 1F11(71) |
| **Ab03** | L1-11-1 | L2-8-1 | L3-8-2 | H1-13-NC | H2-10-1 | 10 | phage | IGKV1-12*01 | 2BX5(92) | IGHV1-69*01 | 3QOT(90) |
| **Ab04** | L1-11-2 | L2-8-1 | L3-9-cis7-1 | H1-13-1 | H2-10-1 | 10 | mouse | IGKV6-25*01 | 1ZTX(95) | IGHV1-67*02 | 1YNT(87) |
| **Ab05** | L1-14-2 | L2-8-1 | L3-11-NC | H1-13-1 | H2-10-1 | 10 | phage | IGLV1-40*01 | 3H42(89) | IGHV5-51*01 | 3NA9(87) |
| **Ab06** | L1-11-1 | L2-8-1 | L3-9-cis7-1 | H1-13-1 | H2-10-2 | 16 | TGM | IGKV1-13*02 | 1VGE(94) | IGHV3-48*02 | 3H42(80) |
| **Ab07** | L1-15-1 | L2-8-1 | L3-9-cis7-1 | H1-14-1 | H2-9-1 | 10 | mouse | IGKV3-12*01 | 3DGG(91) | IGHV3-2*02 | 1S5I(87) |
| **Ab08** | L1-17-1 | L2-8-1 | L3-9-cis7-1 | H1-13-1 | H2-10-3 | 13 | mouse | IGKV8-19*01 | 1MVU(95) | IGHV1-9*01 | 2ZJS(85) |
| **Ab09** | L1-11-1 | L2-8-1 | L3-9-cis7-1 | H1-13-1 | H2-10-1 | 12 | human | IGKV1-39*01 | 2BX5(99) | IGHV5-51*01 | 2H32(92) |
| **Ab10** | L1-17-1 | L2-8-1 | L3-8-1 | H1-13-1 | H2-10-1 | 13 | mouse | IGKV8-27*01 | 2DTG(87) | IGHV1-39*01 | 1KB5(86) |
| **Ab11** | L1-11-1 | L2-8-1 | L3-9-cis7-1 | H1-13-1 | H2-10-1 | 12 | mouse | IGKV5-48*01 | 1YY8(88) | IGHV1-18*01 | 3CVH(82) |
| **CDR** | 24-34 | 49-56 | 89-97 | 23-35 | 50-58 | 93-102 | | | | | |

Light gray shading indicates minor canonical structures, dark gray shading indicates non-canonical structures.
CDR definition is according to North *et al.*[21]
Ab source: phage, human antibody from phage display; TGM, human antibody from transgenic mouse.
GL(VL) and GL(VH) are the closest germlines for VL and VH.
PDB(VL) and PDB(VH) are the closest sequences available in PDB with the sequence identity in parentheses.

typical for the given length. CDR-L1 in Ab04 adopts a minor canonical structure L1-11-2 that is present in 40% of all crystal structures. CDR-L3 in Ab02 and Ab03 belongs to the minor canonical cluster L3-8-2 observed in 18% structures and characterized by a Pro at position 96. However, the peptide bond 92–93 is flipped in both Ab02 and Ab03 with respect to L3-8-2. Nine antibodies have a 10-residue CDR-H2. Most of them belong to the major canonical cluster H2-10-1, but there are two exceptions. Ab06 has canonical structure H2-10-2 (19% frequency of occurrence) and Ab08 has H2-10-3 observed in only 5% structures.

Three antibodies (Ab01, Ab03, and Ab05) present a particularly difficult modeling challenge since they contain CDRs in non-canonical conformations. Ab01, the rabbit antibody, has non-canonical CDR-L1 and CDR-L3, the latter of which contains an internal disulfide bond. Ab03 has non-canonical CDR-H1 in addition to the unusual CDR-L3 with a flipped peptide in a minor canonical structure L3-8-2. Ab05 has an unusually long CDR-L3 with a non-canonical structure and a minor canonical structure for CDR-L1 (L1-14-2), which occurs in 22% structures.

Therefore, the set of 11 reference structures provides a spectrum of canonical and non-canonical CDR conformations for modeling. The length of CDR-H3 in most of these structures varies from 10 to 13 residues, which is within the normal range, except for Ab06, which has a relatively long CDR-H3 of 16 residues. Only one CDR-H3, that of Ab09, has an exact sequence match in the current PDB.

## Conventions and definitions

The models are labeled by the three-letter name of the participant (**acc** for Accelrys, **ccg** for CCG, **jhu** for Jeffrey Gray's group, **joa** for Hiroki Shirai's group, **mmt** for Macromoltek, **sch** for Schrödinger, and **pig** for PIGS) followed by the reference antibody (Ab01–Ab11) followed by the model number (m1–m3). For example, **accAb06m2** is the second Accelrys model for reference structure Ab06.

The Chothia antibody numbering scheme[6,23] is used throughout the article.

The β-sheet core used for domain superposition is defined as follows. VL: 3–13, 18–25, 33–38, 43–49, 61–67, 70–76, 85–90, 97–103. VH: 3–7, 18–24, 34–40, 44–51, 56–59, 67–72, 77–82a, 87–94, 102–110. The CDR boundaries used in the calculation of RMSDs are defined as the following. For comparison, AbM definitions[22] are given in parentheses. CDR-L1: 27–32 (24–34), CDR-L2: 50–53 (50–56), CDR-L3: 91–96 (89–97), CDR-H1: 26–33 (26–35), CDR-H2: 52–55 (50–58), CDR-H3: 95–100*x* (95–102) where 100*x* is the residue preceding position 101, CDR-H4: 73–76. These CDRs are somewhat shorter than normally used because the ends of CDRs are located in the structurally invariant β-strands. We excluded those residues from comparisons to emphasize the deviations in the structurally variable part of CDRs.

## MolProbity scores

The molecular models were assessed with MolProbity 4[24] as implemented in PHENIX version 1.8.2.[25] For MolProbity score calculations, the VL and VH models were truncated to exactly the same length because some of the models contain fewer residues at the N- or C-termini due to differences in modeling algorithms. Hydrogen atoms were ignored for the clash score and MolProbity score calculations.

## Carbonyl RMSD

Structural models are typically assessed by comparison to their respective crystal structures. Root-mean-square deviations (RMSD) between models and structures are

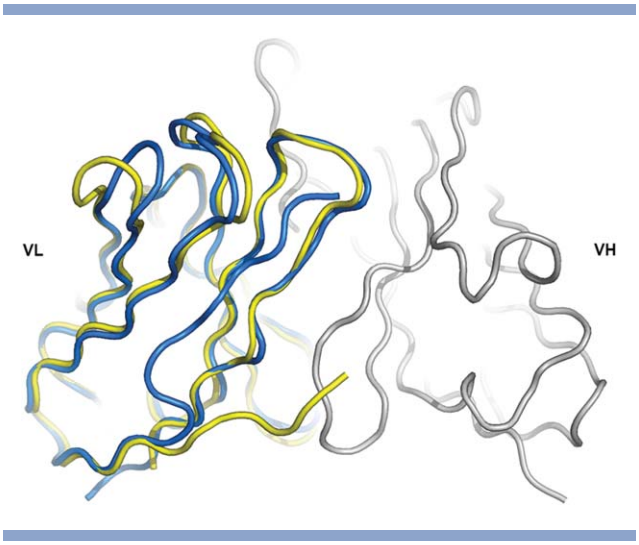


**Figure 3**

Superposition of **schAb06m2** (*yellow*) on the VL domain of Ab06 (*blue*). The VH domain of Ab06 is shown in *gray*.

calculated using either all atoms or just Cα atoms. Since antibody modeling is primarily focused on the fine details of the CDR loops rather than on the overall prediction of the protein fold, we calculated RMSDs based upon the peptide carbonyl atoms C and O to emphasize relatively small deviations in the CDR backbone conformation. For example, a peptide flip in the middle of a four-residue β-turn dislocates the carbonyl oxygen by 3.5 Å, which would result in a RMSD of 1.3 Å calculated across the four carbonyl groups. In contrast, the corresponding RMSD on Cα atoms is only 0.4 Å. Thus, carbonyl RMSDs provide a more sensitive measure of the model versus structure deviations.

To calculate RMSDs for the CDRs, the models produced by the participants were superimposed on the reference structures using Cα atoms of the β-sheet core of the corresponding variable domain, VL or VH, with the help of the program LSQKAB.[26] The RMSDs were calculated for carbonyl atoms C and O of the CDR residues as defined above.

### VL/VH tilt angle

Antibody VH and VL pairing is generally considered random[27] except a small number of germlines, for which pairing preferences do exist.[28] The VL and VH domains associate through the interface residues in both the framework and the CDRs. In addition to measuring the RMSDs between the models and the structures for the individual domains and CDRs, we also calculated the deviation in the tilt angle, which reflects the relative orientation between the VH and VL domains. The difference with respect to the reference structure is calculated by sequential RMS superposition of the VL and VH domains using β-sheet core Cα positions as defined above. The κ angle in spherical polar angular system (ω, φ, κ) of the latter transformation is the difference in the tilt angle.

The VL/VH orientation parameters according to Dunbar *et al.*[29] are also calculated using a script downloaded from the website (http://opig.stats.ox.ac.uk/webapps/abangle).

### CDR-H3

CDR-H3 typically presents the greatest challenge in antibody modeling due to its conformational diversity.[12–14] These difficulties are elevated further by the inaccuracies in modeling the FW region and other CDRs that affect the conformation of CDR-H3. In order to decouple these effects, we devised a two-round procedure for modeling CDR-H3. In the first round, the entire variable domain of the antibody is modeled based solely on the amino acid sequence. In the second round, the participants are provided with the reference structures without CDR-H3, so that the modeling of CDR-H3 can be performed in the context of the correct environment. Each participant submitted five non-ranked models for each antibody. PIGS was not involved in the second round.

## RESULTS

### Overall structural assessment with MolProbity

The quality of the molecular models was assessed with MolProbity,[24] which evaluates protein geometry and clashes in the structures (Supporting Information Table S1). The PIGS models tend to have high scores because they were not subjected to energy minimization. These models are not considered in the following descriptions of MolProbity scores. With a few exceptions, all models have good overall backbone geometry. Average Ramachandran favored are in the range of 92%–96%, average outliers are <0.8% excluding the PIGS models. The MolProbity clash scores are also very good, in line with those of the reference X-ray structures. Several of the models have fairly high clash scores and poor geometry. For example, **schAb07m3** has a clash score of 30.8 with only 75.8% of the backbone in the favored region. As described later, some of these errors were derived from the choice of modeling templates whereas the minimization protocol was not able to remove the clashes. In general, the MolProbity scores have improved since the first assessment[30] as the energy minimization step was included in modeling algorithms. On the other hand, these good scores do not reflect the correctness of the models. For Ab01 for example, all models deviate from the structure in significant ways (discussed below), yet the MolProbity scores are quite acceptable. Thus,

**Table III**
RMSD for the β-sheet Core and CDRs and the VL/VH Tilt Angle Deviation Between the Models and the Reference Structures

| **Ab01** | VL | VH | L1 | L2 | L3 | H1 | H2 | H3 | H4 | Tilt |
|---|---|---|---|---|---|---|---|---|---|---|
| accAb01m1 | 0.4 | 0.6 | 3.6 | 0.5 | 5.5 | 0.7 | 1.5 | 6.0 | 3.3 | 6.5 |
| accAb01m2 | 0.4 | 0.6 | 3.6 | 0.6 | 5.5 | 0.7 | 1.5 | 6.2 | 3.3 | 6.5 |
| accAb01m3 | 0.4 | 0.4 | 4.4 | 0.4 | 5.7 | 1.2 | 1.9 | 6.3 | 0.7 | 6.0 |
| ccgAb01m1 | 0.5 | 0.6 | 4.2 | 0.6 | 3.3 | 1.5 | 1.6 | 6.0 | 2.4 | 6.5 |
| ccgAb01m2 | 0.5 | 0.7 | 1.6 | 0.6 | 3.8 | 1.5 | 1.8 | 6.7 | 2.1 | 6.9 |
| ccgAb01m3 | 0.5 | 0.4 | 4.4 | 0.5 | 4.1 | 1.6 | 1.7 | 6.7 | 2.3 | 7.0 |
| jhuAb01m1 | 0.5 | 0.4 | 4.0 | 0.5 | 5.1 | 0.8 | 2.0 | 6.8 | 3.2 | 9.2 |
| jhuAb01m2 | 0.5 | 0.4 | 4.0 | 0.5 | 5.2 | 0.8 | 2.0 | 7.7 | 3.2 | 10.7 |
| jhuAb01m3 | 0.5 | 0.4 | 4.0 | 0.5 | 5.4 | 0.8 | 2.2 | 7.0 | 3.2 | 9.1 |
| joaAb01m1 | 0.5 | 0.4 | 3.8 | 0.4 | 3.8 | 1.0 | 1.8 | 6.8 | 3.2 | 4.4 |
| joaAb01m2 | 0.6 | 0.4 | 3.7 | 0.5 | 3.8 | 0.9 | 1.9 | 6.8 | 3.2 | 4.3 |
| joaAb01m3 | 0.6 | 0.4 | 3.8 | 0.5 | 3.8 | 0.9 | 1.9 | 6.7 | 3.2 | 4.5 |
| mmtAb01m1 | 0.8 | 0.6 | 3.5 | 1.2 | 9.8 | 1.7 | 0.9 | 5.6 | 3.2 | 10.1 |
| mmtAb01m2 | 0.9 | 0.6 | 3.5 | 1.4 | 9.2 | 1.9 | 1.2 | 5.5 | 3.2 | 10.0 |
| mmtAb01m3 | 0.7 | 0.8 | 3.5 | 1.7 | 8.5 | 1.6 | 1.2 | 5.4 | 3.4 | 7.3 |
| schAb01m1 | 0.8 | 0.8 | 3.8 | 1.3 | 4.4 | 1.1 | 2.1 | 5.8 | 2.9 | 2.7 |
| schAb01m2 | 0.8 | 0.8 | 3.9 | 1.2 | 4.1 | 1.1 | 2.0 | 5.9 | 2.8 | 2.8 |
| schAb01m3 | 0.8 | 0.8 | 4.0 | 1.3 | 6.0 | 1.1 | 2.1 | 6.0 | 2.8 | 2.9 |
| pigAb01m1 | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a |
| **Ab02** | VL | VH | L1 | L2 | L3 | H1 | H2 | H3 | H4 | Tilt |
| accAb02m1 | 0.4 | 0.6 | 0.6 | 0.3 | 1.4 | 1.1 | 1.1 | 4.8 | 1.2 | 6.0 |
| accAb02m2 | 0.4 | 0.6 | 0.6 | 0.3 | 1.1 | 1.1 | 1.1 | 4.8 | 1.2 | 6.1 |
| accAb02m3 | 0.4 | 0.6 | 0.8 | 0.3 | 0.6 | 1.1 | 1.1 | 6.0 | 1.2 | 6.1 |
| ccgAb02m1 | 0.6 | 0.6 | 1.1 | 0.6 | 2.2 | 0.9 | 1.8 | 5.2 | 1.0 | 4.4 |
| ccgAb02m2 | 0.4 | 0.5 | 1.0 | 0.3 | 2.0 | 0.6 | 1.6 | 4.0 | 0.8 | 4.2 |
| ccgAb02m3 | 0.4 | 0.5 | 0.6 | 0.3 | 1.4 | 0.7 | 1.5 | 4.9 | 1.0 | 4.1 |
| jhuAb02m1 | 0.4 | 0.6 | 0.9 | 0.6 | 1.8 | 1.0 | 0.9 | 3.9 | 0.8 | 5.9 |
| jhuAb02m2 | 0.4 | 0.6 | 0.8 | 0.6 | 1.7 | 1.0 | 0.9 | 2.5 | 0.8 | 4.9 |
| jhuAb02m3 | 0.4 | 0.6 | 0.8 | 0.5 | 1.7 | 0.9 | 0.8 | 2.6 | 0.8 | 6.1 |
| joaAb02m1 | 0.4 | 0.5 | 0.9 | 0.4 | 1.5 | 0.9 | 1.0 | 3.5 | 0.9 | 9.2 |
| joaAb02m2 | 0.4 | 0.5 | 1.0 | 0.4 | 1.4 | 0.9 | 1.0 | 4.0 | 0.9 | 9.2 |
| joaAb02m3 | 0.4 | 0.4 | 0.5 | 0.3 | 1.6 | 0.8 | 0.9 | 4.3 | 0.8 | 2.8 |
| mmtAb02m1 | 0.6 | 0.7 | 1.4 | 0.9 | 1.9 | 1.6 | 0.8 | 3.4 | 1.4 | 10.8 |
| mmtAb02m2 | 0.6 | 0.7 | 1.8 | 0.8 | 2.2 | 1.8 | 0.6 | 3.8 | 1.4 | 8.1 |
| mmtAb02m3 | 0.6 | 0.8 | 1.3 | 0.9 | 2.0 | 1.5 | 0.8 | 2.9 | 1.2 | 6.3 |
| schAb02m1 | 0.6 | 0.6 | 0.7 | 0.4 | 2.0 | 0.8 | 0.8 | 7.2 | 0.8 | 3.3 |
| schAb02m2 | 0.4 | 0.5 | 1.3 | 0.4 | 2.1 | 0.7 | 1.4 | 6.0 | 0.8 | 4.3 |
| schAb02m3 | 0.6 | 0.5 | 0.7 | 0.4 | 2.0 | 0.8 | 0.8 | 2.4 | 0.8 | 3.3 |
| pigAb02m1 | 0.4 | 0.5 | 0.9 | 0.4 | 0.9 | 0.8 | 0.8 | 5.0 | 0.7 | 4.6 |

**Table III**
Continued

| Ab03 | VL | VH | L1 | L2 | L3 | H1 | H2 | H3 | H4 | Tilt |
|---|---|---|---|---|---|---|---|---|---|---|
| accAb03m1 | 0.4 | 0.3 | 0.3 | 0.4 | 0.6 | 2.9 | 0.4 | 2.3 | 0.9 | 5.7 |
| accAb03m2 | 0.3 | 0.4 | 0.3 | 0.3 | 0.7 | 2.7 | 1.1 | 3.5 | 0.6 | 6.8 |
| accAb03m3 | 0.3 | 0.4 | 0.4 | 0.3 | 0.5 | 6.1 | 3.8 | 10.3 | 0.8 | 3.7 |
| ccgAb03m1 | 0.7 | 0.5 | 1.2 | 0.9 | 1.9 | 2.1 | 0.8 | 5.1 | 0.8 | 14.1 |
| ccgAb03m2 | 0.6 | 0.5 | 0.8 | 1.0 | 1.7 | 2.3 | 1.1 | 3.1 | 0.9 | 14.3 |
| ccgAb03m3 | 0.4 | 0.3 | 1.3 | 0.5 | 0.8 | 4.1 | 0.5 | 1.4 | 0.6 | 2.3 |
| jhuAb03m1 | 0.4 | 0.3 | 0.4 | 0.4 | 1.9 | 3.1 | 2.6 | 3.1 | 0.5 | 9.0 |
| jhuAb03m2 | 0.4 | 0.3 | 0.4 | 0.4 | 1.8 | 3.2 | 2.4 | 3.1 | 0.6 | 8.8 |
| jhuAb03m3 | 0.4 | 0.3 | 0.4 | 0.4 | 1.9 | 3.1 | 2.4 | 2.8 | 0.6 | 5.1 |
| joaAb03m1 | 0.3 | 0.4 | 0.4 | 0.4 | 0.7 | 2.0 | 0.5 | 2.2 | 0.4 | 3.6 |
| joaAb03m2 | 0.4 | 0.4 | 0.5 | 0.4 | 0.6 | 2.1 | 0.5 | 2.5 | 0.4 | 3.7 |
| joaAb03m3 | 0.3 | 0.4 | 0.4 | 0.4 | 1.2 | 2.1 | 0.3 | 1.9 | 0.5 | 7.6 |
| mmtAb03m1 | 0.5 | 0.5 | 1.3 | 0.9 | 2.4 | 2.0 | 1.5 | 2.7 | 1.5 | 10.1 |
| mmtAb03m2 | 0.5 | 0.5 | 1.4 | 0.6 | 2.9 | 2.5 | 1.1 | 3.5 | 1.4 | 13.1 |
| mmtAb03m3 | 0.5 | 0.5 | 1.5 | 0.6 | 2.6 | 2.2 | 1.6 | 2.4 | 1.7 | 8.4 |
| schAb03m1 | 0.3 | 0.4 | 1.4 | 0.3 | 1.5 | 3.2 | 1.6 | 2.0 | 0.9 | 3.9 |
| schAb03m2 | 0.3 | 0.4 | 1.3 | 0.3 | 2.6 | 3.1 | 1.6 | 1.9 | 0.9 | 3.8 |
| schAb03m3 | 0.3 | 0.4 | 1.3 | 0.3 | 1.5 | 3.2 | 1.6 | 2.0 | 1.0 | 3.9 |
| pigAb03m1 | 0.3 | 0.3 | 0.3 | 0.2 | 0.9 | 1.4 | 0.6 | 2.6 | 0.6 | 11.7 |

| Ab04 | VL | VH | L1 | L2 | L3 | H1 | H2 | H3 | H4 | Tilt |
|---|---|---|---|---|---|---|---|---|---|---|
| accAb04m1 | 0.7 | 0.5 | 1.2 | 0.4 | 0.4 | 1.0 | 1.0 | 1.9 | 0.5 | 6.8 |
| accAb04m2 | 0.7 | 0.4 | 1.2 | 0.4 | 0.4 | 1.0 | 1.0 | 2.1 | 0.5 | 6.8 |
| accAb04m3 | 0.5 | 0.4 | 1.3 | 0.3 | 0.7 | 0.7 | 0.9 | 1.8 | 0.5 | 7.3 |
| ccgAb04m1 | 0.8 | 0.5 | 1.4 | 0.9 | 0.7 | 1.3 | 1.0 | 2.0 | 0.5 | 6.8 |
| ccgAb04m2 | 0.8 | 0.5 | 1.2 | 0.8 | 0.3 | 1.4 | 1.1 | 2.7 | 0.6 | 6.7 |
| ccgAb04m3 | 0.5 | 0.6 | 1.3 | 0.5 | 1.1 | 1.5 | 2.7 | 5.3 | 1.0 | 3.7 |
| jhuAb04m1 | 0.5 | 0.4 | 0.8 | 0.7 | 1.2 | 0.8 | 0.3 | 1.7 | 0.8 | 4.0 |
| jhuAb04m2 | 0.5 | 0.4 | 0.7 | 0.6 | 1.2 | 0.8 | 0.3 | 2.0 | 0.8 | 7.7 |
| jhuAb04m3 | 0.5 | 0.4 | 0.7 | 0.8 | 1.2 | 0.8 | 0.3 | 1.6 | 0.8 | 5.4 |
| joaAb04m1 | 0.5 | 0.4 | 0.5 | 0.4 | 0.8 | 0.9 | 1.5 | 1.8 | 0.5 | 4.8 |
| joaAb04m2 | 0.6 | 0.4 | 0.6 | 0.4 | 0.7 | 0.9 | 1.4 | 1.5 | 0.5 | 4.8 |
| joaAb04m3 | 0.6 | 0.5 | 0.5 | 0.4 | 0.7 | 1.1 | 1.3 | 1.7 | 0.6 | 8.1 |
| mmtAb04m1 | 0.6 | 0.6 | 1.1 | 1.0 | 0.9 | 1.0 | 1.0 | 2.0 | 0.7 | 8.0 |
| mmtAb04m2 | 0.5 | 0.6 | 1.1 | 1.1 | 0.9 | 0.8 | 0.9 | 1.7 | 0.8 | 11.9 |
| mmtAb04m3 | 0.5 | 0.6 | 1.1 | 1.0 | 0.8 | 0.9 | 0.9 | 2.5 | 0.7 | 12.4 |
| schAb04m1 | 0.7 | 0.4 | 1.2 | 1.3 | 0.4 | 1.0 | 1.3 | 3.6 | 0.6 | 5.8 |
| schAb04m2 | 0.6 | 0.5 | 1.3 | 0.9 | 0.7 | 1.1 | 0.9 | 3.6 | 0.7 | 2.9 |
| schAb04m3 | 0.7 | 0.4 | 1.2 | 1.2 | 0.4 | 1.0 | 1.3 | 1.9 | 0.6 | 5.8 |
| pigAb04m1 | 0.5 | 0.4 | 0.7 | 0.3 | 0.5 | 0.8 | 0.4 | 1.4 | 0.7 | 3.6 |

**Table III**
Continued

| Ab05 | VL | VH | L1 | L2 | L3 | H1 | H2 | H3 | H4 | Tilt |
|---|---|---|---|---|---|---|---|---|---|---|
| accAb05m1 | 0.4 | 0.8 | 1.0 | 0.4 | 3.8 | 0.4 | 0.4 | 3.7 | 0.9 | 8.4 |
| accAb05m2 | 0.4 | 0.3 | 0.8 | 0.5 | 2.5 | 0.2 | 0.4 | 3.7 | 0.5 | 12.0 |
| accAb05m3 | 0.4 | 0.8 | 0.9 | 0.3 | 3.9 | 0.4 | 0.5 | 3.7 | 0.8 | 8.4 |
| ccgAb05m1 | 0.5 | 0.4 | 1.1 | 0.7 | 2.3 | 0.7 | 0.6 | 4.0 | 0.5 | 6.5 |
| ccgAb05m2 | 0.4 | 0.4 | 0.9 | 0.6 | 2.0 | 0.9 | 0.8 | 2.8 | 0.6 | 6.6 |
| ccgAb05m3 | 0.5 | 0.3 | 0.8 | 0.4 | 2.5 | 0.8 | 0.5 | 2.9 | 0.5 | 3.7 |
| jhuAb05m1 | 0.4 | 0.3 | 1.5 | 0.5 | 2.6 | 0.9 | 0.7 | 3.6 | 0.4 | 8.0 |
| jhuAb05m2 | 0.4 | 0.3 | 1.5 | 0.6 | 2.6 | 1.2 | 1.0 | 3.6 | 0.4 | 8.1 |
| jhuAb05m3 | 0.4 | 0.3 | 1.6 | 0.5 | 2.6 | 0.9 | 0.7 | 3.5 | 0.4 | 8.4 |
| joaAb05m1 | 0.4 | 0.4 | 0.8 | 0.4 | 3.0 | 1.5 | 0.6 | 3.3 | 0.6 | 6.3 |
| joaAb05m2 | 0.4 | 0.4 | 0.8 | 0.4 | 3.0 | 1.4 | 0.5 | 1.7 | 0.7 | 6.5 |
| joaAb05m3 | 0.6 | 0.7 | 5.2 | 0.5 | 2.5 | 0.7 | 1.2 | 3.0 | 0.9 | 14.1 |
| mmtAb05m1 | 0.5 | 0.8 | 1.1 | 0.6 | 2.6 | 0.7 | 0.6 | 4.0 | 0.7 | 6.8 |
| mmtAb05m2 | 0.5 | 0.8 | 1.6 | 0.7 | 2.7 | 1.2 | 0.8 | 2.6 | 0.7 | 8.8 |
| mmtAb05m3 | 0.5 | 0.8 | 1.4 | 0.8 | 2.5 | 0.8 | 0.6 | 2.4 | 0.6 | 9.2 |
| schAb05m1 | 0.5 | 0.5 | 2.3 | 0.6 | 2.6 | 0.8 | 0.8 | 2.7 | 1.2 | 4.9 |
| schAb05m2 | 0.3 | 0.4 | 0.9 | 0.8 | 2.5 | 1.2 | 0.6 | 6.2 | 0.3 | 4.7 |
| schAb05m3 | 0.5 | 0.5 | 2.3 | 0.6 | 2.6 | 0.8 | 0.8 | 2.4 | 1.2 | 4.9 |
| pigAb05m1 | 0.4 | 0.9 | 1.1 | 0.5 | 2.3 | 1.1 | 1.0 | 3.6 | 0.7 | 6.0 |

| Ab06 | VL | VH | L1 | L2 | L3 | H1 | H2 | H3 | H4 | Tilt |
|---|---|---|---|---|---|---|---|---|---|---|
| accAb06m1 | 0.4 | 0.3 | 0.4 | 0.3 | 0.7 | 0.7 | 0.4 | 4.8 | 0.5 | 4.9 |
| accAb06m2 | 0.4 | 0.3 | 0.4 | 0.3 | 0.7 | 0.6 | 0.3 | 3.9 | 0.5 | 4.8 |
| accAb06m3 | 0.3 | 0.4 | 0.5 | 0.3 | 0.8 | 0.4 | 0.3 | 3.0 | 0.5 | 4.7 |
| ccgAb06m1 | 0.4 | 0.5 | 0.6 | 0.6 | 1.1 | 0.8 | 1.0 | 4.5 | 0.7 | 4.0 |
| ccgAb06m2 | 0.4 | 0.5 | 0.4 | 0.4 | 1.0 | 0.8 | 0.8 | 5.2 | 0.8 | 3.8 |
| ccgAb06m3 | 0.4 | 0.4 | 0.6 | 0.4 | 0.6 | 1.4 | 0.6 | 2.5 | 0.6 | 5.5 |
| jhuAb06m1 | 0.4 | 0.5 | 0.5 | 0.6 | 0.8 | 1.1 | 0.4 | 4.2 | 0.8 | 4.8 |
| jhuAb06m2 | 0.4 | 0.5 | 0.5 | 0.6 | 0.8 | 1.1 | 0.5 | 5.3 | 0.8 | 5.8 |
| jhuAb06m3 | 0.4 | 0.5 | 0.5 | 0.6 | 1.0 | 1.1 | 0.4 | 4.2 | 0.8 | 5.7 |
| joaAb06m1 | 0.3 | 0.4 | 0.5 | 0.2 | 0.8 | 0.7 | 0.5 | 2.5 | 0.5 | 8.2 |
| joaAb06m2 | 0.4 | 0.4 | 0.4 | 0.3 | 0.8 | 0.7 | 0.6 | 5.0 | 0.6 | 8.2 |
| joaAb06m3 | 0.3 | 0.4 | 0.3 | 0.3 | 0.8 | 0.6 | 0.4 | 2.2 | 0.4 | 4.5 |
| mmtAb06m1 | 0.6 | 0.5 | 0.5 | 0.5 | 0.7 | 0.8 | 0.6 | 3.3 | 0.8 | 4.2 |
| mmtAb06m2 | 0.7 | 0.5 | 0.6 | 0.6 | 0.8 | 0.9 | 0.6 | 6.7 | 0.6 | 5.4 |
| mmtAb06m3 | 0.7 | 0.6 | 0.8 | 0.6 | 0.8 | 0.8 | 0.7 | 4.5 | 0.8 | 3.9 |
| schAb06m1 | 0.3 | 0.5 | 0.5 | 0.5 | 1.2 | 0.9 | 0.6 | 3.4 | 0.6 | 5.6 |
| schAb06m2 | 1.0 | 0.5 | 1.0 | 0.4 | 1.2 | 1.2 | 2.3 | 3.1 | 0.5 | 7.8 |
| schAb06m3 | 0.3 | 0.5 | 0.5 | 0.5 | 1.1 | 0.9 | 0.6 | 5.1 | 0.6 | 5.6 |
| pigAb06m1 | 0.5 | 0.4 | 0.5 | 0.4 | 0.4 | 1.2 | 0.5 | 4.0 | 0.3 | 5.7 |

**Table III**
Continued

| Ab07 | VL | VH | L1 | L2 | L3 | H1 | H2 | H3 | H4 | Tilt |
|---|---|---|---|---|---|---|---|---|---|---|
| accAb07m1 | 0.3 | 0.4 | 0.8 | 0.2 | 0.7 | 0.3 | 0.3 | 3.1 | 0.4 | 3.3 |
| accAb07m2 | 0.3 | 0.4 | 0.8 | 0.2 | 0.7 | 0.3 | 0.3 | 3.2 | 0.4 | 3.3 |
| accAb07m3 | 0.3 | 0.4 | 0.8 | 0.2 | 0.7 | 0.3 | 0.3 | 3.0 | 0.4 | 3.3 |
| ccgAb07m1 | 0.8 | 0.6 | 2.1 | 0.5 | 0.6 | 0.5 | 0.6 | 3.8 | 0.3 | 1.7 |
| ccgAb07m2 | 0.7 | 0.5 | 2.1 | 0.4 | 0.6 | 0.4 | 0.7 | 3.5 | 0.5 | 1.4 |
| ccgAb07m3 | 0.5 | 0.4 | 0.5 | 0.4 | 1.0 | 0.6 | 0.4 | 2.6 | 0.3 | 2.3 |
| jhuAb07m1 | 0.3 | 0.3 | 1.0 | 0.4 | 0.7 | 0.4 | 0.4 | 1.5 | 0.5 | 4.3 |
| jhuAb07m2 | 0.3 | 0.3 | 0.9 | 0.4 | 0.6 | 0.5 | 0.6 | 2.4 | 0.5 | 4.0 |
| jhuAb07m3 | 0.3 | 0.3 | 1.0 | 0.4 | 0.7 | 0.4 | 0.6 | 1.9 | 0.5 | 7.1 |
| joaAb07m1 | 0.3 | 0.3 | 1.8 | 0.2 | 1.6 | 0.4 | 0.7 | 1.1 | 0.3 | 1.3 |
| joaAb07m2 | 0.3 | 0.3 | 1.8 | 0.3 | 1.6 | 0.4 | 0.6 | 2.9 | 0.4 | 1.3 |
| joaAb07m3 | 0.4 | 0.3 | 1.9 | 0.6 | 0.6 | 0.4 | 0.2 | 1.0 | 0.3 | 4.2 |
| mmtAb07m1 | 0.5 | 0.4 | 0.5 | 0.5 | 0.7 | 1.5 | 0.4 | 2.6 | 0.3 | 6.0 |
| mmtAb07m2 | 0.5 | 0.5 | 0.8 | 0.7 | 0.8 | 1.5 | 0.3 | 3.0 | 0.4 | 3.8 |
| mmtAb07m3 | 0.5 | 0.4 | 0.6 | 0.8 | 0.7 | 1.6 | 0.6 | 3.6 | 0.4 | 3.9 |
| schAb07m1 | 0.4 | 0.3 | 1.9 | 0.8 | 1.0 | 1.7 | 0.8 | 2.3 | 0.5 | 3.3 |
| schAb07m2 | 0.9 | 0.5 | 1.3 | 0.4 | 1.5 | 0.7 | 0.8 | 3.8 | 0.8 | 3.4 |
| schAb07m3 | 0.4 | 0.3 | 2.0 | 0.8 | 1.0 | 1.7 | 0.8 | 1.7 | 0.5 | 3.3 |
| pigAb07m1 | 0.9 | 0.5 | 3.7 | 0.9 | 0.7 | 0.9 | 0.7 | 3.6 | 0.9 | 3.8 |

| Ab08 | VL | VH | L1 | L2 | L3 | H1 | H2 | H3 | H4 | Tilt |
|---|---|---|---|---|---|---|---|---|---|---|
| accAb08m1 | 0.3 | 0.4 | 0.7 | 0.2 | 0.9 | 0.5 | 1.1 | 2.3 | 0.3 | 1.7 |
| accAb08m2 | 0.3 | 0.4 | 0.7 | 0.2 | 0.9 | 0.5 | 1.1 | 4.7 | 0.3 | 1.7 |
| accAb08m3 | 0.3 | 0.4 | 0.7 | 0.2 | 0.9 | 0.5 | 1.1 | 4.3 | 0.3 | 1.7 |
| ccgAb08m1 | 0.6 | 0.6 | 0.9 | 0.5 | 0.7 | 0.5 | 0.6 | 3.8 | 0.4 | 5.4 |
| ccgAb08m2 | 0.6 | 0.5 | 0.9 | 0.9 | 0.5 | 0.3 | 0.8 | 3.9 | 0.5 | 5.5 |
| ccgAb08m3 | 0.5 | 0.4 | 0.6 | 0.4 | 0.8 | 0.4 | 1.2 | 3.5 | 0.3 | 2.5 |
| jhuAb08m1 | 0.5 | 0.3 | 0.8 | 0.5 | 0.6 | 0.6 | 0.7 | 3.2 | 0.2 | 8.6 |
| jhuAb08m2 | 0.5 | 0.3 | 0.9 | 0.5 | 0.6 | 0.6 | 0.8 | 3.1 | 0.2 | 2.3 |
| jhuAb08m3 | 0.5 | 0.3 | 0.8 | 0.5 | 0.6 | 0.5 | 0.7 | 3.1 | 0.2 | 7.8 |
| joaAb08m1 | 0.6 | 0.5 | 1.4 | 0.2 | 0.8 | 1.2 | 0.5 | 2.4 | 0.4 | 4.8 |
| joaAb08m2 | 0.6 | 0.5 | 1.4 | 0.2 | 0.9 | 1.3 | 0.7 | 3.1 | 0.5 | 4.8 |
| joaAb08m3 | 0.6 | 0.4 | 0.9 | 0.1 | 0.8 | 0.6 | 0.6 | 2.4 | 0.3 | 4.4 |
| mmtAb08m1 | 0.7 | 0.5 | 0.8 | 0.9 | 0.8 | 0.7 | 0.9 | 5.6 | 0.6 | 5.3 |
| mmtAb08m2 | 0.6 | 0.4 | 0.6 | 0.6 | 0.6 | 0.8 | 1.1 | 5.0 | 0.6 | 7.3 |
| mmtAb08m3 | 0.7 | 0.4 | 0.9 | 0.5 | 0.6 | 0.7 | 0.7 | 3.1 | 0.4 | 4.8 |
| schAb08m1 | 0.5 | 0.4 | 0.8 | 0.3 | 0.6 | 0.6 | 0.8 | 3.3 | 0.3 | 2.6 |
| schAb08m2 | 0.5 | 0.6 | 0.9 | 0.4 | 0.7 | 0.6 | 0.9 | 3.0 | 0.4 | 2.7 |
| schAb08m3 | 0.5 | 0.4 | 0.9 | 0.4 | 0.6 | 0.6 | 0.8 | 4.2 | 0.3 | 2.6 |
| pigAb08m1 | 0.5 | 0.7 | 1.6 | 0.8 | 0.5 | 0.5 | 1.0 | 3.5 | 1.0 | 4.0 |

**Table III**
Continued

| Ab09 | VL | VH | L1 | L2 | L3 | H1 | H2 | H3 | H4 | Tilt |
|---|---|---|---|---|---|---|---|---|---|---|
| accAb09m1 | 0.3 | 0.2 | 0.3 | 0.3 | 0.5 | 0.4 | 0.7 | 3.0 | 0.7 | 1.9 |
| accAb09m2 | 0.3 | 0.2 | 0.3 | 0.3 | 0.5 | 0.4 | 0.7 | 4.2 | 0.7 | 1.9 |
| accAb09m3 | 0.3 | 0.2 | 0.3 | 0.3 | 0.5 | 0.4 | 0.7 | 4.6 | 0.7 | 1.9 |
| ccgAb09m1 | 0.5 | 0.3 | 1.3 | 0.6 | 0.9 | 0.7 | 1.0 | 1.8 | 0.7 | 2.0 |
| ccgAb09m2 | 0.4 | 0.3 | 1.3 | 0.5 | 0.8 | 0.7 | 1.0 | 2.0 | 0.7 | 1.8 |
| ccgAb09m3 | 0.4 | 1.0 | 0.8 | 0.6 | 1.0 | 0.7 | 1.0 | 1.4 | 0.7 | 5.7 |
| jhuAb09m1 | 0.3 | 0.3 | 0.4 | 0.5 | 0.5 | 0.4 | 0.5 | 2.1 | 0.6 | 5.7 |
| jhuAb09m2 | 0.3 | 0.3 | 0.4 | 0.4 | 0.5 | 0.4 | 0.6 | 1.1 | 0.6 | 4.3 |
| jhuAb09m3 | 0.3 | 0.3 | 0.4 | 0.4 | 0.4 | 0.5 | 1.0 | 1.5 | 0.6 | 5.5 |
| joaAb09m1 | 0.5 | 0.5 | 0.5 | 0.3 | 0.7 | 0.4 | 1.4 | 2.2 | 0.9 | 2.9 |
| joaAb09m2 | 0.5 | 0.5 | 0.5 | 0.3 | 0.6 | 0.4 | 1.3 | 1.6 | 0.9 | 2.9 |
| joaAb09m3 | 0.3 | 0.9 | 0.6 | 0.4 | 0.4 | 1.1 | 1.6 | 2.7 | 1.3 | 5.1 |
| mmtAb09m1 | 0.7 | 0.8 | 0.5 | 0.5 | 2.5 | 0.9 | 1.3 | 4.7 | 0.5 | 5.5 |
| mmtAb09m2 | 0.5 | 0.9 | 0.4 | 0.5 | 2.4 | 0.6 | 1.3 | 3.9 | 0.6 | 6.5 |
| mmtAb09m3 | 0.5 | 0.8 | 0.5 | 0.4 | 2.1 | 0.7 | 0.8 | 4.9 | 0.5 | 5.8 |
| schAb09m1 | 0.4 | 0.4 | 0.6 | 0.8 | 0.8 | 0.6 | 0.7 | 2.7 | 0.4 | 3.3 |
| schAb09m2 | 0.3 | 1.0 | 0.7 | 0.3 | 2.1 | 1.4 | 1.8 | 3.7 | 1.1 | 4.8 |
| schAb09m3 | 0.4 | 0.4 | 0.6 | 0.8 | 0.8 | 0.6 | 0.6 | 4.3 | 0.4 | 3.3 |
| pigAb09m1 | 0.7 | 0.9 | 0.8 | 1.3 | 1.5 | 0.9 | 0.7 | 2.5 | 0.9 | 7.2 |

| Ab10 | VL | VH | L1 | L2 | L3 | H1 | H2 | H3 | H4 | Tilt |
|---|---|---|---|---|---|---|---|---|---|---|
| accAb10m1 | 0.3 | 0.4 | 0.6 | 0.3 | 1.6 | 0.4 | 0.4 | 1.9 | 1.3 | 6.5 |
| accAb10m2 | 0.3 | 0.4 | 0.6 | 0.3 | 1.6 | 0.4 | 0.4 | 2.2 | 1.3 | 6.5 |
| accAb10m3 | 0.3 | 0.4 | 0.6 | 0.3 | 1.6 | 0.4 | 0.4 | 3.1 | 1.3 | 6.5 |
| ccgAb10m1 | 0.4 | 0.5 | 1.2 | 0.4 | 0.8 | 0.4 | 1.0 | 4.1 | 1.0 | 5.0 |
| ccgAb10m2 | 0.3 | 0.5 | 1.1 | 0.4 | 0.7 | 0.4 | 0.8 | 3.5 | 1.0 | 5.2 |
| ccgAb10m3 | 0.4 | 0.3 | 0.7 | 0.3 | 0.6 | 0.7 | 0.9 | 2.3 | 0.7 | 5.8 |
| jhuAb10m1 | 0.5 | 0.5 | 0.7 | 0.3 | 2.6 | 1.2 | 1.0 | 2.1 | 0.9 | 5.7 |
| jhuAb10m2 | 0.5 | 0.5 | 0.7 | 0.3 | 2.6 | 1.3 | 1.1 | 1.4 | 0.9 | 2.8 |
| jhuAb10m3 | 0.5 | 0.5 | 0.8 | 0.3 | 2.6 | 1.6 | 1.4 | 1.7 | 0.9 | 2.1 |
| joaAb10m1 | 0.3 | 0.4 | 0.9 | 0.4 | 2.1 | 0.8 | 0.6 | 2.8 | 1.2 | 2.3 |
| joaAb10m2 | 0.3 | 0.4 | 0.9 | 0.5 | 2.1 | 0.8 | 0.6 | 4.2 | 1.3 | 2.3 |
| joaAb10m3 | 0.4 | 0.5 | 0.8 | 0.5 | 1.9 | 0.8 | 0.6 | 5.3 | 1.3 | 2.3 |
| mmtAb10m1 | 0.5 | 0.6 | 1.2 | 0.9 | 2.0 | 1.0 | 0.6 | 3.3 | 1.1 | 6.0 |
| mmtAb10m2 | 0.5 | 0.6 | 1.4 | 0.9 | 2.2 | 1.2 | 0.3 | 3.7 | 1.1 | 6.1 |
| mmtAb10m3 | 0.5 | 0.5 | 1.0 | 0.6 | 2.2 | 0.6 | 0.6 | 3.1 | 1.0 | 8.3 |
| schAb10m1 | 0.3 | 0.5 | 2.1 | 0.3 | 0.6 | 1.2 | 0.7 | 2.8 | 1.0 | 5.0 |
| schAb10m2 | 0.5 | 0.4 | 1.2 | 0.5 | 1.3 | 1.4 | 0.7 | 2.5 | 0.9 | 5.7 |
| schAb10m3 | 0.3 | 0.5 | 2.1 | 0.3 | 0.6 | 1.2 | 0.7 | 2.7 | 1.0 | 5.0 |
| pigAb10m1 | 0.2 | 0.5 | 1.3 | 0.3 | 1.6 | 0.5 | 1.4 | 4.2 | 1.1 | 4.7 |

**Table III**
Continued

| Ab11 | VL | VH | L1 | L2 | L3 | H1 | H2 | H3 | H4 | Tilt |
|---|---|---|---|---|---|---|---|---|---|---|
| accAb11m1 | 0.4 | 0.7 | 0.5 | 0.3 | 0.9 | 2.6 | 2.3 | 3.6 | 2.4 | 11.9 |
| accAb11m2 | 0.4 | 0.7 | 0.5 | 0.3 | 0.9 | 2.6 | 2.3 | 3.4 | 2.4 | 11.9 |
| accAb11m3 | 0.4 | 0.7 | 0.5 | 0.5 | 0.9 | 2.6 | 2.3 | 2.7 | 2.4 | 11.9 |
| ccgAb11m1 | 0.5 | 0.6 | 0.8 | 0.5 | 0.5 | 0.9 | 2.0 | 4.3 | 0.9 | 6.7 |
| ccgAb11m2 | 0.4 | 0.5 | 0.4 | 0.5 | 0.5 | 0.6 | 1.8 | 3.0 | 0.9 | 7.0 |
| ccgAb11m3 | 0.5 | 0.5 | 0.4 | 0.3 | 0.5 | 0.7 | 1.8 | 3.3 | 0.9 | 6.3 |
| jhuAb11m1 | 0.4 | 0.4 | 0.4 | 0.4 | 0.7 | 1.2 | 2.4 | 3.9 | 0.5 | 5.1 |
| jhuAb11m2 | 0.4 | 0.4 | 0.4 | 0.4 | 0.7 | 0.9 | 2.6 | 3.6 | 0.5 | 5.5 |
| jhuAb11m3 | 0.4 | 0.4 | 0.3 | 0.3 | 0.7 | 0.9 | 2.5 | 0.8 | 0.5 | 5.7 |
| joaAb11m1 | 0.5 | 0.7 | 0.4 | 0.3 | 0.8 | 0.7 | 1.0 | 2.7 | 1.0 | 5.4 |
| joaAb11m2 | 0.6 | 0.7 | 0.3 | 0.3 | 0.8 | 0.7 | 1.0 | 1.2 | 1.0 | 5.2 |
| joaAb11m3 | 0.5 | 0.6 | 0.4 | 0.2 | 0.8 | 0.7 | 2.7 | 2.8 | 1.0 | 5.7 |
| mmtAb11m1 | 0.5 | 0.6 | 0.6 | 0.4 | 0.8 | 1.1 | 0.8 | 3.5 | 0.3 | 6.7 |
| mmtAb11m2 | 0.5 | 0.7 | 0.5 | 0.5 | 0.8 | 1.0 | 0.5 | 2.4 | 0.3 | 7.2 |
| mmtAb11m3 | 0.5 | 0.7 | 0.5 | 0.5 | 1.1 | 1.2 | 0.5 | 5.0 | 0.3 | 10.9 |
| schAb11m1 | 0.6 | 0.7 | 0.5 | 0.8 | 0.6 | 0.9 | 2.1 | 3.9 | 1.1 | 6.3 |
| schAb11m2 | 0.5 | 0.6 | 0.4 | 0.4 | 1.4 | 0.7 | 2.7 | 2.7 | 0.8 | 8.3 |
| schAb11m3 | 0.6 | 0.7 | 0.4 | 0.8 | 0.7 | 1.0 | 2.1 | 3.9 | 1.1 | 6.3 |
| pigAb11m1 | 0.3 | 0.5 | 0.3 | 0.2 | 0.6 | 0.6 | 0.4 | 2.7 | 0.4 | 7.3 |

Columns "VL" and "VH" are the carbonyl RMSDs (in Å) for the β-sheet core as defined in Methods section; "L1" to "H4" are the carbonyl RMSDs (in Å) for CDRs; "Tilt" (in degrees) is the difference in the VL/VH angle between the model and the reference structure; Three-color scheme: top 10 percentile red, bottom 10 percentile green, 50 percentile yellow.

MolProbity seems to be a good tool to find errors in molecular models but is not good to measure model accuracy.

## Framework region

The models produced by the participants were superimposed on the reference structures using Cα atoms of the β-sheet core (defined above) of the corresponding variable domain, VL or VH. Residues 1–8 were excluded from superposition of VL because several models had incorrectly placed the N-terminal strand and this could skew the overall alignment. As an example of such a misplacement of the N-terminal strand, Figure 3 shows a superposition of Ab06 and **schAb06m2**. The template for this model was taken from 3DVG,[31] which had this portion of the Fab mistraced.

Overall, the core of the variable domains was modeled well, with the RMSD (on the carbonyl groups) mostly in the range 0.4–0.7 Å (Table III). Although the task of modeling the framework region is not particularly difficult, a few deviations of up to 1.0 Å point to some potential pitfalls. For instance, several models of Ab09 VH show high RMSDs in contrast to other models that are very accurate. Ab09 is an antibody composed of human germlines IGHV5-51 and IGKV1-39. A unique feature of the germline family VH5 (which includes germlines 5–51 and 5-a) is Trp at position 82, whereas other germlines have either Leu or Met. To accommodate the bulky indole group buried in the VH domain, β-strands E and B are displaced by approximately 2 Å with respect to their positions in other germlines. Selection of antibodies with non-VH5 germlines as a template for modeling Ab09 resulted in higher overall deviations. This example illustrates the importance of using templates with matching germlines.

## Cis-trans isomers

Analysis of the *cis*-isomers reveals another source of errors in the FW region. An incorrect isomer may lead to an error of up to 5 Å for the carbonyl oxygen. Thus a single carbonyl group may account for a approximately 0.2 Å higher RMSD for the FW region. Most of the incorrect isomers in the FW region originate from wrong templates, whereas similar errors in the CDRs are probably due to the loop-building algorithm. Table IV presents the data on the *cis*-isomers in the reference structures and in the models.

The conserved Pro8 in kappa light chains adopts a *cis*-conformation. There are of course lambda chains with *trans*-Pro8 and kappa chains with a non-proline residue at position 8 in the *trans*-conformation. The range of options at position 8 caused some confusion in several

**Table IV**
*Cis*-isomers in the Reference Structures and in the Models

| | Ab01 | Ab02 | Ab03 | Ab04 | Ab05 | Ab06 | Ab07 | Ab08 | Ab09 | Ab10 | Ab11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **xray** | P8L | P8L | P8L | P95L | none | P8L, P95L | P8L, P77L, P95L | P8L, P95L | P8L, P95L | P8L | P8L, P95L |
| **acc-m1** | | | G104H | H8L | | | | | | G100aH | |
| **acc-m2** | | | | H8L | | | | | | | |
| **acc-m3** | | | G104H | | | | | | | | |
| **ccg-m1** | N30aL, N98H | | | | | | (P8L, P77L) | S77L | | | |
| **ccg-m2** | | | | | | | (P8L, P77L) | S77L | | | |
| **ccg-m3** | | | G26H | | | | | | | | (P8L) |
| **jhu-m1** | | | | H8L | | | | | | | |
| **jhu-m2** | | | | H8L | | | | | | | |
| **jhu-m3** | | | | H8L | | | | | | | |
| **joa-m1** | | | | | | | | | | | |
| **joa-m2** | | | | | | | | | | | |
| **joa-m3** | | | | | | G42H | | | | | |
| **mmt-m1** | T7L | | S7L | | | S7L | | | S7L | | |
| **mmt-m2** | T7L | | S7L | | | S7L | | | S7L | | |
| **mmt-m3** | T7L | | S7L | | | S7L | | | S7L | | |
| **sch-m1** | T77H | (P8L) | G104H | H8L | | | | | | | |
| **sch-m2** | T77H | | G104H | | | (P8L) | (P77L) | | G104H | | |
| **sch-m3** | T77H | (P8L) | G104H | H8L | | | | | | | |
| **pig-m1** | | R99H | | | | | (P77L) | | | | |

Residues are labeled with a one-letter amino acid code, position in Chothia's numbering and a letter, L or H, for the chain. Only deviations from the reference structures are shown. Residues in parentheses indicate trans-isomers where cis-isomers are due.

cases. Pro8 was incorrectly modeled in the *trans*-conformation in **ccgAb07m1**, **ccgAb07m2**, **ccgAb11m3**, **schAb02m1**, **schAb02m3,** and **schAb06m2**, all of which were based on the templates with non-Pro residues.

An opposite mistake is the introduction of the *cis*-peptide where it should be *trans*. Ab04 has *trans*-His at position 8, however, seven models have His8 in the *cis*-conformation. For instance, **schAb04m1** was based on 3MXV,[32] which has mouse germline KV6-32 with *cis*-Pro8, and this *cis*-peptide was incorporated into the model. A better choice for modeling Ab04 would be germline KV6-17 from 4H20[33] with *trans*-His8 that was utilized in **schAb04m2**.

A special case is the **mmt** models for Ab01, Ab03, Ab06, and Ab09 where the first eight residues of VL were placed out of register so that the residue preceding Pro8 happened to be in the *cis*-conformation. This was one of the reasons to exclude residues 1–8 of VL from the overall superposition.

Another conserved proline in kappa light chains, Pro95 in CDR-L3, adopts a *cis*-conformation. There was no difficulty in modeling *cis*-Pro95, all models had it right, which probably reflects the lack of variation at this position. More problem was with residue 77 of VL that can be either *cis* of *trans*. The majority of the models have the correct isomer at position 77. Four models (**ccgAb07m1**, **ccgAb07m2**, **schAb07m2,** and **pigAb07m1**) incorrectly have Pro77 in the *trans*-conformation in Ab07. Alternatively, *trans*-Ser77 in Ab08 was modeled in the *cis*-configuration in **ccgAb08m1** and **ccgAb08m2**. Once again, these examples point to the choice of a template with the right isomer at certain positions.

None of the reference structures has *cis*-isomers in VH. Nevertheless, 14 models exhibit various residues in the *cis*-conformation. Most of these residues occur in the tight turns of CDRs and may have resulted from energy minimization of the models. Curiously, six models have suffered from the use of a single template 2XTJ,[34] which has Gly104 (a conserved residue in the WGQG motif following CDR-H3) in the *cis*-conformation. It looks like this *cis*-Gly has originated from the SWISS-MODEL server[35] that was used for constructing the search model for 2XTJ structure determination.

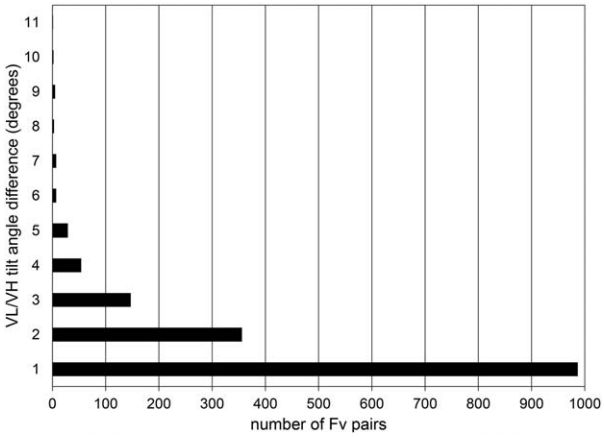


**Figure 4**
Distribution of VL/VH tilt angle differences in the PDB entries containing multiple copies of the Fv domains.

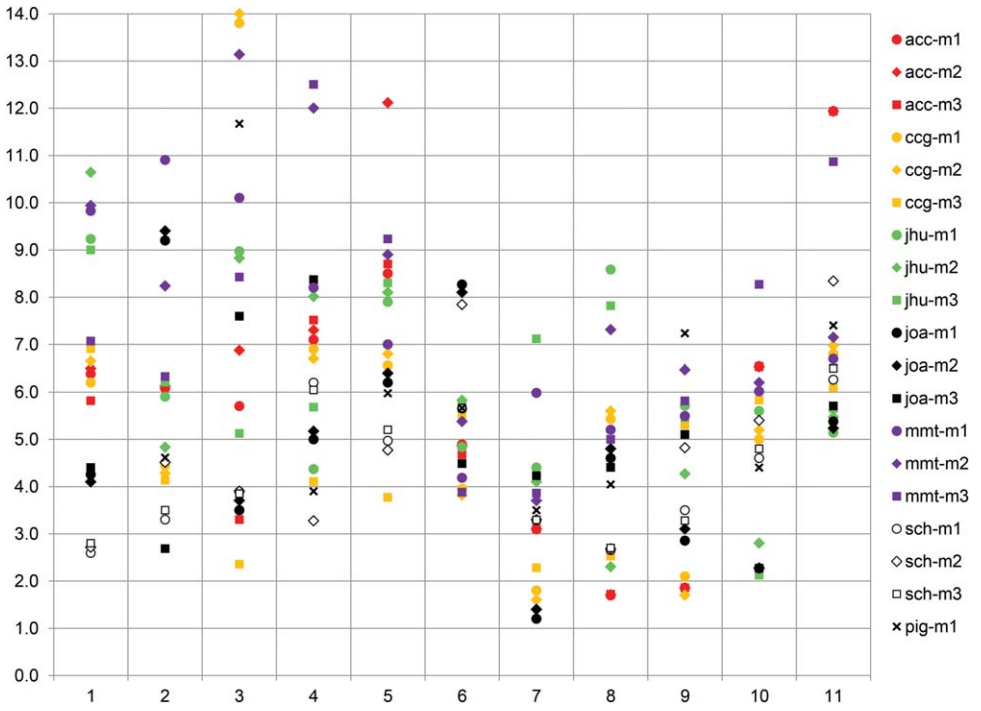


**Figure 5**

Deviation in VL/VH tilt angle from reference structures. Values are taken from Table III and displayed so that each marker represents one model, circle for model 1, rhomb for model 2, square for model 3. Models from the same modeler are in one color.

## VL/VH tilt angle

The variable domains VL and VH have the β-sandwich fold that allows only minimal conformational freedom. However, their relative orientation can vary over a wide range due to sequence differences at the domain interface. Several methods have been proposed to characterize the VH/VL orientation.[8,29,36] The method of Dunbar *et al.*[29] produces six parameters that describe the VH/VL orientation. These parameters were calculated for the reference structures and the models and the results are presented in Supporting Information Table S2. All models have these parameters within the expected range derived from the analysis of antibody structures in the PDB.[29] Because the physical meaning

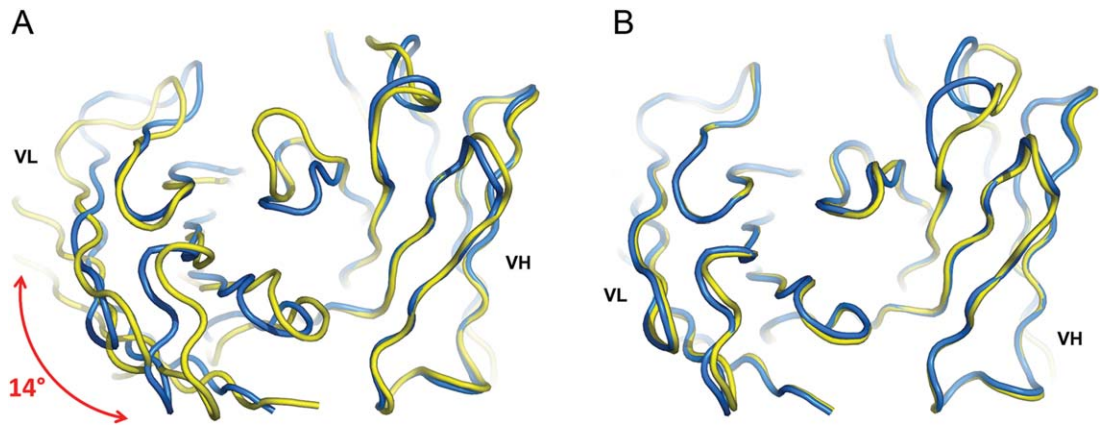


**Figure 6**

Superposition of the **ccg** models 1 and 3 (*yellow*) on the VH domain of Ab03 (*blue*) as viewed from the antigen-binding side. (**A**) **ccgAb03m1**. The VL domain is off by a rigid-body rotation of 14°. (**B**) **ccgAb03m3**. The VL domain is within 2° from Ab03.

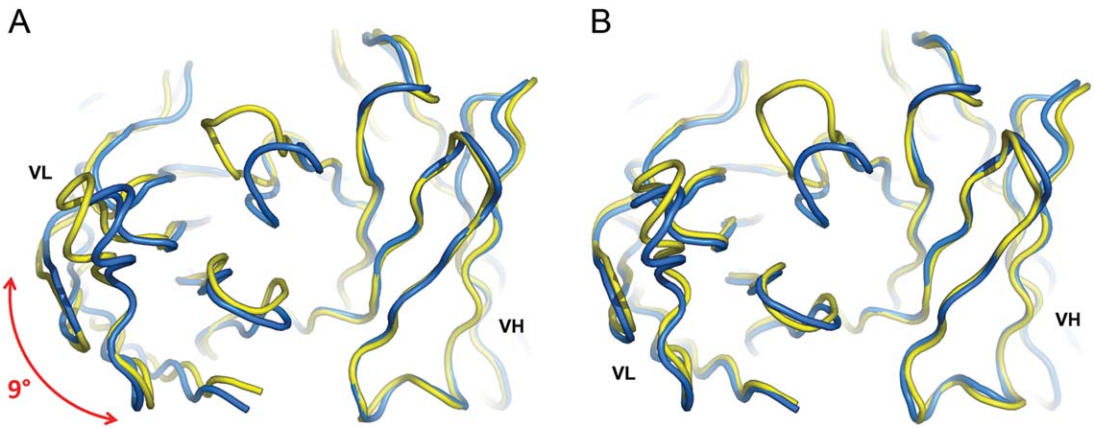


**Figure 7**

Superposition of **joaAb02m1** and **joaAb02m3** (*yellow*) on the VH domain of Ab02 (*blue*) as viewed from the CDR side. In **joaAb02m1** the VL domain is off by a rigid-body rotation of 9° while in **joaAb02m3** it is within 3° from Ab02.

of the six parameters is not intuitive, we use a simple relative parameter to characterize the difference in VL/VH orientation rather than the absolute values.

The measure of the VL/VH orientation used in this study is the difference in the VL/VH tilt angle between the model and the reference structure calculated as described in the Methods section. To estimate what range of tilt angle deviations to expect for the models, we carried out a survey of the entire set of antibody structures in the PDB (as of February 2014) using a customized MOE script (CCG). All PDB structures with multiple copies of the same Fab or Fv in the asymmetric unit were included in the calculations. Only the molecules present in the same crystal were compared. The VH and VL domains were superimposed using all atoms; otherwise the procedure was as described above. The results are presented in Figure 4. Although there is some "breathing" of the Fv domain, the VL/VH tilt angles rarely deviate by more than 3° and typically stay in the range of 1°–2° despite different crystal environment. Deviations >5° constitute about 1.5% of all cases and deviations exceeding 7° occur in only four structures.

The comparisons of the tilt angles in the models with the corresponding reference structures are shown in Table III. These values are brought together in Figure 5 where each model is represented by a marker. While the individual domains VL and VH were modeled quite well, the tilt angles in many cases deviate significantly from the expected values. Very few models have tilt angles within the expected range of 3°. Structure Ab07 is the best-modeled example, particularly for models **ccg** and **joa**.

It is of interest to analyze the largest deviations in the tilt angle. Superposition of the **ccg** models 1 and 3 on Ab03 suggests a relationship between the tilt angle and the conformation of CDR-H3 (Fig. 6). In **ccgAb03m1**, CDR-H3 was modeled in the "out" configuration (the apex of the loop points away from the center), whereas in Ab03 it is "in" (the apex of the loop points toward the center), which requires a wider opening of the VL/VH cleft. Oppositely, **ccgAb03m3** correctly locates CDR-H3 in the "in" configuration and accordingly has the minimal tilt deviation from Ab03 (this is one of the best CDR-H3 predictions). From this example one may conclude that the conformation of CDR-H3 affects the VL/VH packing angle.

On the other hand, the relationship between CDR-H3 and the tilt angle may be inverted. Comparison of the **joa** models for Ab02 (Fig. 7) indicates that the tilt angle

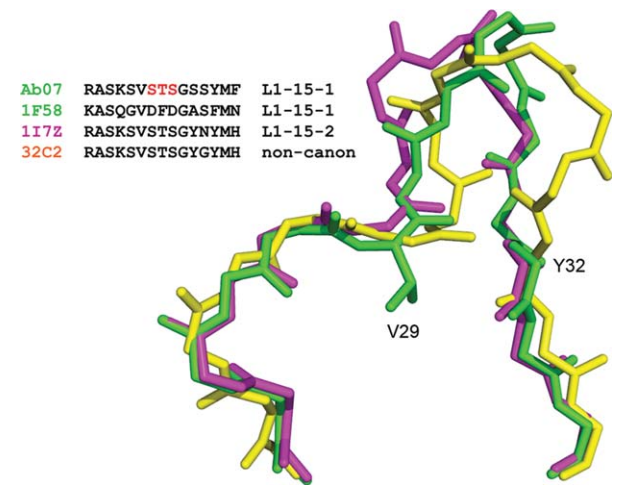


**Figure 8**

Superposition of CDR-L1 in Ab07 (*green*), 1I7Z (*magenta*), and 32C2 (*yellow*). Side chains except Val29 are not shown for clarity. Insert on the left shows the sequence alignment of the segment Cys23-Trp35 and the canonical structures.

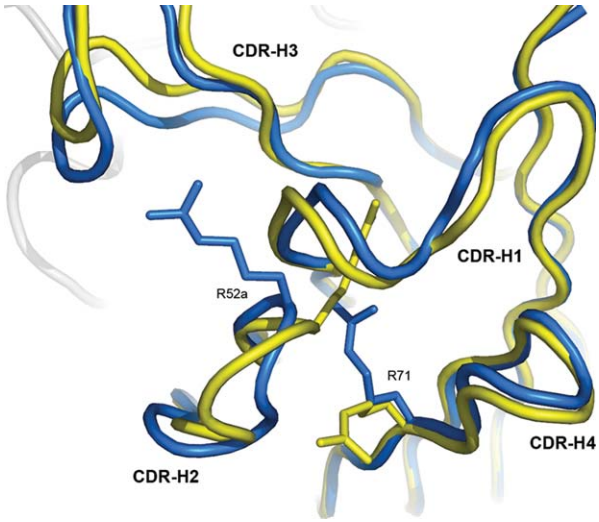


**Figure 9**

Superposition of VH domains of Ab06 (*blue*) and **schAb06m2** (*yellow*). The conformation of Arg52a and Arg71 is, respectively, "out"/"in" in Ab06, "in"/"out" in **schAb06m2**.

may be inherited from the template regardless of the CDR-H3 conformation. Both models, **joaAb02m1** and **joaAb02m3**, have CDR-H3 correctly in the "out" configuration, which was built into the fixed template, 3O2D[37] for **joaAb02m1** and 3Q3G[38] for **joaAb02m3**. The templates have substantially different tilt angles though. This example points to the importance of the adjustment of the VL/VH orientation, perhaps simultaneously with CDR-H3 modeling.

## Canonical CDRs

Since it was discovered that five out of six CDRs adopt only a limited number of canonical conformations[6] and the rules relating the sequence and the structure of these CDRs were established,[11] their modeling was considered straightforward. However, every rule has an exception. Results of CDR modeling presented in Table III indicate difficulties not only with the CDRs that adopt non-canonical conformations, but surprisingly with a number of CDRs that have major canonical structures. Examples of the latter are CDR-L1 in Ab07 and Ab08, CDR-L1 and CDR-L3 in Ab10, CDR-H1 and CDR-H2 in Ab11.

Based on the amino acid sequence, possible templates for CDR-L1 in Ab07 are either 1F58[39] or 1I7Z.[40] Figure 8 illustrates the difference between the templates. CDR-L1 in 1F58 belongs to the major canonical structure L1-15-1 that covers 85% of all cases in the PDB including the structure in question, Ab07. 1I7Z is a representative of the other 15% with a slightly different conformation of the β-hairpin. The third variant utilized in **pigAb07m1** is a non-canonical conformation observed in

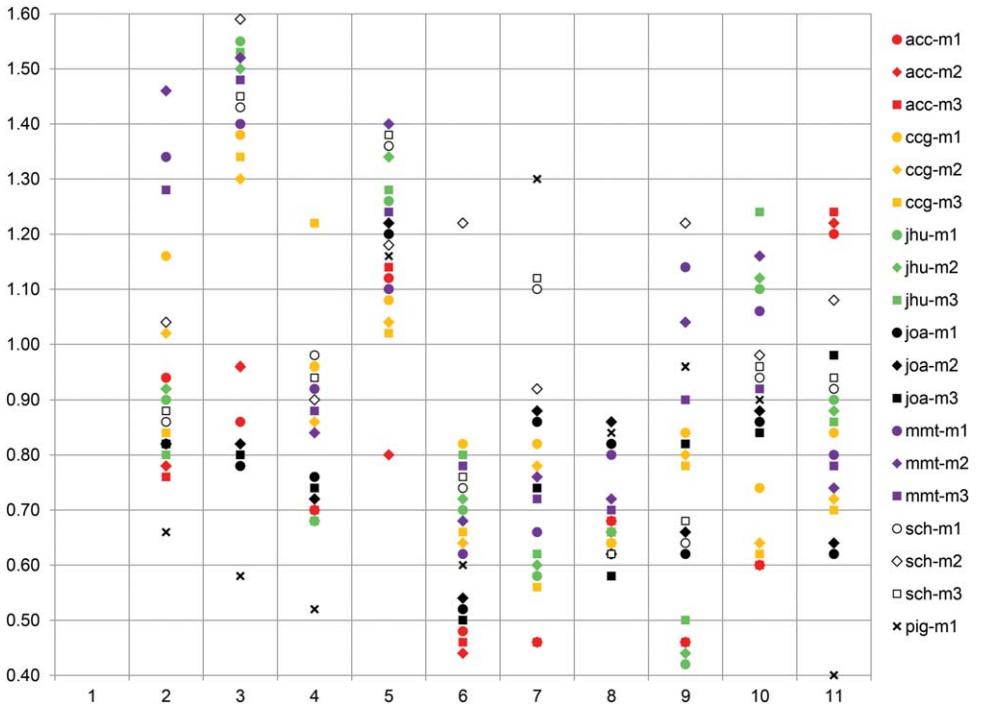


**Figure 10**

RMSD from reference structures for five canonical CDRs. Each point is an average of five RMSDs (from Table III) for a given model. The markers are as described for Fig. 5.

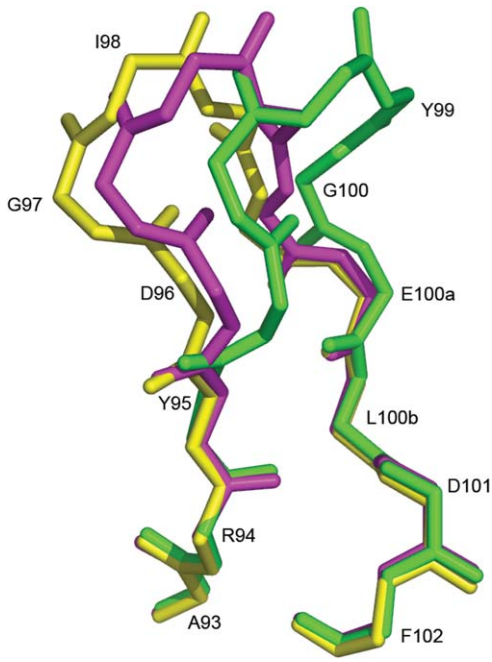


**Figure 11**

Superposition of CDR-H3 in Ab09 (*green*), 3QOT (*yellow*), and 4DN3 (*magenta*). Side chains are not shown.

32C2.[41] This variant should have been excluded from consideration because the structure 32C2 is of poor quality, which is evident from its MolProbity scores and a Ramachandran plot. The choice between 1F58 and 1I7Z is not easy. On the one hand, L1-15-1 occurs six times more frequently, on the other hand, the STS motif in the middle of the sequence suggests 1I7Z to be a better choice. In fact, it was not, and those models that were based on 1I7Z (**ccg**, **joa**, **sch**) show large deviations from the reference structure with RMSD values of approximately 2 Å.

An unexpected difficulty was presented by CDR-L3 of Ab10, which has canonical structure L3-8-1. Although none of the participants has chosen a minor canonical structure L3-8-2, only the **sch** models were close to the reference structure. A peculiar feature of canonical structure L3-8-1 is the unusual conformation of residue 94 with $\varphi = 60°$, $\psi = -150°$. In most models the peptide is flipped to maintain favorable Ramachandran plot geometry, which results in larger RMSDs. The **ccg** models are correct in terms of the canonical structure but the RMSD is relatively high due to the incorrect rotamer of His91, which forces a shift of the CDR loop by approximately 1 Å. This example points to the problem of positioning one CDR in the context of neighboring CDRs, in this case CDR-H3.

Another example on this theme reveals the interplay of CDR-H2 and CDR-H4 (strictly speaking, the latter is not a CDR since it is never involved in antigen recognition). The effect of CDR-H4 on CDR-H2 was noted earlier.[42] There is a pocket between these CDRs, and there are two contenders for the space, residues 52a of CDR-H2 and 71 of CDR-H4. When residue 71 is Arg, which is true for half of human germlines, there is no room for residue 52a, which then takes the "out" orientation. As a result, the canonical structure of CDR-H2 very much depends on residue 71. Ab06 has Arg in both positions and naturally Arg52a is "out" (Fig. 9). Model **schAb06m2** is based on 3DVG,[31] which has Pro52a and Ala71, both in the pocket. An attempt to squeeze both arginines into the pocket in **schAb06m2** was unsuccessful. The other model from the same participant, **schAb06m1**, was based on 3HR5,[43] which had Asp52a and Arg71, and was correct. This example points to yet another factor to be considered when selecting a template.

One aspect of modeling CDR-H2 is highlighted by the Ab11 case. The RMSD is very high for most of the models (Table III) despite the fact that all of them have the correct canonical structure H2-10-1. However, the top half of the β-hairpin loop (residues 51–57) is displaced by 2–3 Å in a number of models (**acc**, **ccg**, **jhu**, **sch**). At the same time some other models (**pig**, **mmt**, **joa**) are quite accurate. Examination of the case did not reveal any crystallographic artifacts either in Ab11 or in the templates. Although some templates were derived from the complexes with antigens, no correlation between the presence of the antigen and the conformation of CDR-H2 could be found. Undoubtedly, CDR-H2 is different from other CDRs in that it can move as a rigid body within a few angstroms, opening or closing the gap to CDR-H4. For example, in the templates used for **joaAb11m1** (2OZ4)[44] and **joaAb11m3** (1A3R)[45] the gap widens from 7.7 to 11.4 Å, as measured between residues 53 and 73.

## Non-canonical CDRs

Despite the pitfalls in modeling canonical structures, a real challenge is to model the CDR with a non-canonical conformation. There were four such CDRs in the reference structures (Table II). Rabbit antibody Ab01 was a particularly difficult challenge. CDR-L3 in Ab01 is unusually long and has an internal disulfide bond 94–95d. No model came close to it (Table III). CDR-L1 also deviates from canonical structures, but not as dramatically as was presumed in the models. Both anchoring residues, Val29 and Tyr32, occupy their usual positions within the Fv. The unique feature of this antibody is that the β-hairpin insertion in CDR-L1 occurs between residues 29 and 30, and not between 30 and 31. Only one model, **ccgAb01m2**, identified this as the case.

Ab05 is the only reference antibody with the lambda light chain. It has an 11-residue CDR-L3 without

**Table V**
Carbonyl RMSD of CDR-H3 in the Round-2 Models from the Reference Structures

| | Ab02 | Ab03 | Ab04 | Ab05 | Ab06 | Ab07 | Ab08 | Ab09 | Ab10 | Ab11 |
|---|---|---|---|---|---|---|---|---|---|---|
| acc-m0 | 4.8 | 2.3 | 1.8 | 3.7 | 3.0 | 3.0 | 2.3 | 3.0 | 1.9 | 2.7 |
| acc-m1 | 3.0 | 1.0 | 1.5 | 2.8 | 3.6 | 2.0 | 2.3 | 2.3 | 4.6 | 3.6 |
| acc-m2 | 3.4 | 0.7 | 1.0 | 0.8 | 3.1 | 2.2 | 3.3 | 2.9 | 4.7 | 5.3 |
| acc-m3 | 4.4 | 1.9 | 1.3 | 3.1 | 3.8 | 2.5 | 3.6 | 3.1 | 1.3 | 4.9 |
| acc-m4 | 2.7 | 2.9 | 1.9 | 3.2 | 4.8 | 1.7 | 4.0 | 3.5 | 3.5 | 4.2 |
| acc-m5 | 4.6 | 3.8 | 1.0 | 0.8 | 2.8 | 2.6 | 3.6 | 3.0 | 3.0 | 3.0 |
| ccg-m0 | 4.0 | 1.4 | 2.0 | 2.8 | 2.5 | 2.6 | 3.5 | 1.4 | 2.3 | 3.0 |
| ccg-m1 | 3.2 | 2.2 | 1.0 | 1.4 | 3.9 | 1.8 | 6.9 | 3.1 | 3.2 | 1.8 |
| ccg-m2 | 2.9 | 2.8 | 1.6 | 2.7 | 3.3 | 1.1 | 3.8 | 3.2 | 4.2 | 3.3 |
| ccg-m3 | 3.4 | 1.0 | 0.9 | 1.2 | 4.9 | 4.2 | 4.0 | 4.0 | 3.1 | 3.4 |
| ccg-m4 | 5.1 | 2.3 | 1.2 | 3.2 | 6.9 | 2.6 | 5.3 | 4.1 | 3.5 | 2.1 |
| ccg-m5 | 3.5 | 3.0 | 1.3 | 1.9 | 4.7 | 1.4 | 3.9 | 1.8 | 3.5 | 5.3 |
| jhu-m0 | 2.5 | 2.8 | 1.6 | 3.5 | 4.2 | 1.5 | 3.1 | 1.1 | 1.4 | 0.8 |
| jhu-m1 | 3.3 | 1.7 | 1.2 | 1.9 | 3.9 | 0.7 | 3.4 | 1.4 | 1.3 | 4.1 |
| jhu-m2 | 1.7 | 1.5 | 1.2 | 4.0 | 3.8 | 2.5 | 2.2 | 1.8 | 3.6 | 4.0 |
| jhu-m3 | 2.6 | 2.8 | 1.2 | 2.7 | 5.1 | 1.8 | 2.4 | 3.1 | 3.1 | 4.7 |
| jhu-m4 | 3.4 | 2.5 | 1.4 | 2.9 | 4.7 | 5.3 | 4.3 | 2.0 | 2.7 | 5.1 |
| jhu-m5 | 3.0 | 2.4 | 1.5 | 4.2 | 8.0 | 5.4 | 4.7 | 3.9 | 7.1 | 7.7 |
| joa-m0 | 3.5 | 1.9 | 1.5 | 1.7 | 2.2 | 1.0 | 2.4 | 1.6 | 2.8 | 1.2 |
| joa-m1 | 3.3 | 1.3 | 1.0 | 0.9 | 3.6 | 1.5 | 2.0 | 1.3 | 1.1 | 0.6 |
| joa-m2 | 0.7 | 1.1 | 1.1 | 0.4 | 4.5 | 1.2 | 3.6 | 2.7 | 4.3 | 4.3 |
| joa-m3 | 3.4 | 1.2 | 1.2 | 3.3 | 4.0 | 1.9 | 3.0 | 2.8 | 2.9 | 2.2 |
| joa-m4 | 2.2 | 0.6 | 1.0 | 4.5 | 4.5 | 2.1 | 3.3 | 3.6 | 4.7 | 2.8 |
| joa-m5 | 2.2 | 1.1 | 1.4 | 2.3 | 3.9 | 2.2 | 4.5 | 2.6 | 2.4 | 3.4 |
| mmt-m0 | 2.9 | 2.4 | 1.7 | 2.4 | 3.3 | 2.6 | 3.1 | 3.9 | 3.1 | 2.4 |
| mmt-m1 | 3.2 | 2.1 | 1.0 | 3.1 | 3.4 | 6.5 | 3.5 | 2.8 | 2.6 | 3.9 |
| mmt-m2 | 3.1 | 2.3 | 1.1 | 2.9 | 3.8 | 3.7 | 5.2 | 2.8 | 2.4 | 3.4 |
| mmt-m3 | 2.6 | 2.2 | 1.2 | 3.2 | 6.3 | 3.5 | 3.6 | 3.1 | 3.4 | 4.2 |
| mmt-m4 | 4.0 | 1.5 | 1.1 | 1.7 | 4.3 | 4.6 | 3.7 | 2.2 | 4.7 | 3.3 |
| mmt-m5 | 3.2 | 2.1 | 1.2 | 2.6 | 3.0 | 3.7 | 3.5 | 2.5 | 2.1 | 3.0 |
| sch-m0 | 2.4 | 1.9 | 1.9 | 2.4 | 3.1 | 1.7 | 3.0 | 2.7 | 2.5 | 2.7 |
| sch-m1 | 3.2 | 0.5 | 1.1 | 3.2 | 3.1 | 0.4 | 1.8 | 0.6 | 1.0 | 0.5 |
| sch-m2 | 2.1 | 1.9 | 1.4 | 3.7 | 4.8 | 2.4 | 1.8 | 4.1 | 2.5 | 1.7 |
| sch-m3 | 3.2 | 1.3 | 1.1 | 2.4 | 4.2 | 2.0 | 3.8 | 3.5 | 4.3 | 2.6 |
| sch-m4 | 2.6 | 1.0 | 1.3 | 0.4 | 8.8 | 1.6 | 2.5 | 4.0 | 4.6 | 1.1 |
| sch-m5 | 3.3 | 1.1 | 0.8 | 4.4 | 6.6 | 0.5 | 2.0 | 3.4 | 6.2 | 0.6 |

Models "m0" are the best from round 1; Three-color scheme: top 10 percentile red, bottom 10 percentile green, 50 percentile yellow.

prolines. All models are based on the few lambda templates available in the PDB that belong to canonical structure L3-11-1, which is a nearly perfect β-hairpin. However, CDR-L3 in Ab05 forms a wide loop stabilized by external rather by internal interactions. Admittedly, this was a tough call.

Ab03 presented a challenge in CDR-H1. Almost 90% of all Fabs in the PDB exhibit canonical structure H1-13-

1 in CDR-H1.[21] Residues 27 (typically Phe or Tyr) and 29 (Phe or Ile or Val) anchor the CDR in the hydrophobic core of VH and define the conformation of the CDR. Ab03 has neither this nor any of the other described canonical structures. However, there is one structure, 2CMR,[46] that is very close and in fact together with Ab03 may now be "canonized." In both structures, CDR-H1 forms two turns of a regular α-helix whereas in H1-13-1 it is mostly irregular with a single β-turn in the middle. A small difference between Ab03 and 2CMR is that in Ab03 peptide 26–27 is flipped. **pigAb03m1** appeared to be the only model that was based on 2CMR and therefore has got a nice score (Table III). All other models used canonical structure H1-13-1, although Gly at position 27 should have been a warning.

The CDR modeling results are summarized in Figure 10. RMSDs for all CDRs except CDR-H3 were averaged for each model and plotted the same way as in Figure 5. There are no markers for Ab01 because all of them fall outside the plot due to poor prediction of CDR-L1 and CDR-L3. Among other structures, most difficult happened to be Ab03 and Ab05 both of which contain non-canonical CDR structures. Curiously, these are synthetic antibodies that were derived from phage display libraries. The easiest cases were apparently Ab06 and Ab08 even though both have CDR-H2 in a minor canonical conformation. A much wider spread of the models for other targets indicates certain difficulties in choosing the right variant. Examples of these are Ab02 and Ab09 characterized by a large deviation between the models.

### CDR-H3

Not unexpectedly, the RMSDs for CDR-H3 shown in Table III are mostly in red (top 10 percentile) compared with the other five CDRs, which for the most part have well-defined canonical structures. Only a few good models were obtained (**ccgAb03m3**, **pigAb04m1**, **joaAb07m1**, **jhuAb07m1**, **jhuAb09m2**, **ccgAb09m3**, **jhuAb10m2**, **jhuAb11m3**, **joaAb11m2**), which is not surprising given the difficulty of the modeling task. Each participant submitted three models for each target, and in several instances one of the models was closer to the correct conformation than the others. A good example is the **jhu** models for Ab11. **jhuAb11m3** is much closer to the crystal structure (RMSD = 0.8 Å) than the other two, **jhuAb11m1** and **jhuAb11m2**, which deviate significantly (RMSD = 3.9 Å). The question here is how to select the best model in the absence of crystal structure.

An interesting case was presented by Ab09. It is a synthetic antibody constructed from human germlines IGHV5-51 and IGKV1-39 that has CDR-H3 from an unrelated antibody 4DN3.[47] Another synthetic antibody, 3QOT,[30] with different germlines but the same CDR-H3 is also available from the PDB. A superposition of CDR-H3 for these structures is shown in Figure 11. In all three structures, CDR-H3 forms a β-hairpin with a β-turn at the top; however, the orientation of the loop with respect to VH is different to adjust for the environment. The span at Cα of Ile98 for the entire set of predictions is over 5 Å. So, even in this rare case when the exact CDR-H3 sequence is present in the PDB, modeling was not easy. More interestingly, these CDR-H3 structures from the available PDB entries were not chosen as modeling templates in any one of the submitted models, presumably due to the fact that ab initio methods rather than template based approaches were used for CDR-H3 modeling. This example also illustrates a special difficulty of CDR-H3, for its structure or orientation with respect to the rest of Fv is influenced by its environment a lot as compared with other CDRs.

To evaluate the effect of the environment on the CDR-H3 modeling, a second round was organized where participants were provided with the reference structures, in which CDR-H3 was omitted. That would eliminate problems related to the VL/VH tilt angle and misplaced canonical CDRs. Ab01 was removed from the list because of its unusual difficulty in the first round. The results of the second round are given in Table V in comparison to the best model from round 1 (PIGS does not have a CDR-H3 refinement mode and thus did not participate in this round). There are obviously some improvements for a number of targets, but even with this additional information, targets that proved to be difficult to model in the first round remained difficult. These are the targets with longer CDR-H3, particularly Ab06 with a 16-residue CDR-H3. On the other hand, shorter CDRs (10 residues) of Ab03, Ab04, Ab05, and Ab07 have seen some really good models, although their ranking remains an open question. For instance, Ab05 models **accAb05m2**, **joaAb05m2,** and **schAb05m4** are quite accurate with the RMSD below 1.0 Å, but other models from the same participants, for example, **accAb05m4**, **joaAb05m4,** and **schAb05m5**, are inaccurate (RMSD of 3.2–4.5 Å).

## DISCUSSION AND CONCLUSIONS

Antibody modeling basically involves three steps: (1) the choice of a template from the PDB for the VL, VH, and Fv when VL and VH template structures are from different antibodies, (2) grafting of five canonical CDRs from the Fab structures of choice onto each template, and (3) the template-based or ab initio modeling of CDR-H3 optionally followed by energy minimization. The current assessment reveals some bottlenecks and pitfalls in this process.

The templates provide grafting points for CDRs but more importantly they also define the packing of the VL and VH domains which can be characterized by a

relative tilt angle such as the one defined in this work. The importance of the VL/VH tilt angle for antibody modeling has been recognized recently.[17,36,48] Since this is extremely important for the relative positioning of CDRs, the choice of the Fv template must include specific filters besides the overall sequence similarity. Packing prediction tools[8,29] may benefit the template selection.

Comparison of the models, their templates and the corresponding X-ray structures also reveals that many of the errors in the models are derived from the errors in the structures in the PDB. For a small sampling, we inspected the electron density maps (from ED server) of the templates used in the **acc** models (27 Fab structures) as well as six high resolution structures used in the elbow angle survey.[49] Of the 33 structures, 10 structures have various structure refinement issues including peptide bond flips, missing residues, solvent molecules, incorrect rotamer conformations, and poor backbone geometry (data not shown). Two examples of incorrectly built peptide bonds are shown in Supporting Information Figure S1 for 2XA8 and 1NGZ.[50] These errors were propagated in the current models and possibly in the crystal structures that used them as search models for molecular replacement. Thus, careful curation of the PDB structures to generate a set of templates devoid of errors and suitable for modeling would improve antibody modeling.

Correct identification and use of the CDR canonical structures is the key element in antibody modeling. One of the best selection algorithms appears to be in the PIGS modeling server, which generated the lowest-RMSD backbone structures for at least five of the targets. The other PIGS models are also within the common range of agreement, with the only outlier (Ab07) resulting from the unfortunate use of a template that contains serious errors, which could have been avoided if a curated database of templates had been available. Classification of canonical structures is an evolving process, and great progress has been made recently with the growing number of Fab structures in the PDB. This work undoubtedly needs to continue with more focus on the rules that relate the CDR sequence and its conformation as well as key positions in the neighbor CDR and framework regions. There may also be a correlation between canonical structures of neighbor CDRs in terms of the preference for one in the presence of the other, much like the preferred VL/VH pairing.

Modeling non-canonical structures including CDR-H3 remains the biggest challenge. For CDR-H3, variations in sequence and conformation are so vast that it seems impractical to look for a template. Building CDR-H3 from scratch is perhaps the most promising path, especially in view of a number of successful cases in this assessment. Modern programs seem to be able to generate correct conformations, unfortunately together with many incorrect ones. Ranking of the models is currently imperfect, so there is clearly a need to devise an extremely fine weighting scheme that would be able to select a correct model from the ensemble of structures. It is worth noting that algorithms that take advantage of large computational resources for extensive conformational sampling of CDR-H3 (**sch**, **jhu**, **joa**) seem to produce better agreement between models and reference structures (Table V).

A variety of methods were employed by participants to construct models of the target antibodies. Automated and manual approaches were combined to different degrees, and the use of energy minimization varied. Overall, these methods have been able to predict Fv structures that have reasonably good templates in the database. Such models, though imperfect, are still useful for antibody engineering in the absence of X-ray structures. Compared with the first assessment, some advances have been achieved based upon the CDR RMSDs, although many challenges remain. Lessons learned from this exercise will help develop new algorithms for structure prediction that will be useful not only for antibodies but also for other classes of proteins.

## ACKNOWLEDGMENTS

We thank Johannes Maier (CCG) for providing a script to analyze the tilt angles in the PDB structures.

## REFERENCES

1. Sircar A. Methods for the homology modeling of antibody variable regions. Methods Mol Biol 2012;857:301–311.
2. Ramos OH. Computer-assisted modeling of antibody variable domains. Methods Mol Biol 2012;907:39–55.
3. Kuroda D, Shirai H, Jacobson MP, Nakamura H. Computer-aided antibody design. Protein Eng Des Sel 2012;25:507–521.
4. Al-Lazikani B, Jung J, Xiang Z, Honig B. Protein structure prediction. Curr Opin Chem Biol 2001;5:51–56.
5. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The protein data bank. Nucleic Acids Res 2000;28:235–242.
6. Chothia C, Lesk A. Canonical structures for the hypervariable regions of immunoglobulins. J Mol Biol 1987;196:901–917.
7. Morea V, Tramontano A, Rustici M, Chothia C, Lesk AM. Antibody structure, prediction and redesign. Biophys Chem 1997;68:9–16.
8. Abhinandan KR, Martin AC. Analysis and prediction of VH/VL packing in antibodies. Protein Eng Des Sel 2010;23:689–697.
9. Davies DR, Metzger H. Structural basis of antibody function. Annu Rev Immunol 1983;1:87–117.
10. Kabat EA. The structural basis of antibody complementarity. Adv Protein Chem 1978;32:1–75.
11. Morea V, Lesk AM, Tramontano A. Antibody modeling: implications for engineering and design. Methods 2000;20:267–279.
12. Reczko M, Martin AC, Bohr H, Suhai S. Prediction of hypervariable CDR-H3 loop structures in antibodies. Protein Eng 1995;8:389–395.
13. Morea V, Tramontano A, Rustici M, Chothia C, Lesk AM. Conformations of the third hypervariable region in the VH domain of immunoglobulins. J Mol Biol 1998;275:269–294.
14. Shirai H, Kidera A, Nakamura H. Structural classification of CDR-H3 in antibodies. FEBS Lett 1996;399:1–8.
15. Oliva B, Bates PA, Querol E, Avilés FX, Sternberg MJ. Automated classification of antibody complementarity determining region 3 of

the heavy chain (H3) loops into canonical forms and its application to protein structure prediction. J Mol Biol 1998;279:1193–1210.
16. Kuroda D, Shirai H, Kobori M, Nakamura H. Structural classification of CDR-H3 revisited: a lesson in antibody modeling. Proteins 2008;73:608–620.
17. Sircar A, Kim ET, Gray JJ. RosettaAntibody: antibody variable region homology modeling server. Nucleic Acids Res 2009;37: W474–W479.
18. Choi Y, Deane CM. Predicting antibody complementarity determining region structures without classification. Mol Biosyst 2011;7: 3327–3334.
19. Zhu K, Day T. Ab initio structure prediction of the antibody hypervariable H3 loop. Proteins 2013;81:1081–1089.
20. Marcatili P, Rosi A, Tramontano A. PIGS: automatic prediction of antibody structures. Bioinformatics 2008;24:1953–1954.
21. North B, Lehmann A, Dunbrack RL. A new clustering of antibody CDR loop conformations. J Mol Biol 2011;406:228–256.
22. Abhinandan KR, Martin AC. 2008. Analysis and improvements to Kabat and structurally correct numbering of antibody variable domains. Mol Immunol 2008;45:3832–3839.
23. Chothia C, Lesk AM, Tramontano A, Levitt M, Smith-Gill SJ, Air G, Sheriff S, Padlan EA, Davies D, Tulip WR, Colman PM, Spinelli S, Alzari PM, Poljak RJ. Conformations of immunoglobulin hypervariable regions. Nature 1989;342:877–883.
24. Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr 2010;D66:12–21.
25. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr 2010;D66:213–221.
26. Kabsch W. A solution for the best rotation to relate two sets of vectors. Acta Crystallogr 1976;A32:922–923.
27. de Wildt RM, Hoet RM, van Venrooij WJ, Tomlinson IM, Winter G. Analysis of heavy and light chain pairings indicates that receptor editing shapes the human antibody repertoire. J Mol Biol 1999;285: 895–901.
28. Jayaram N, Bhowmick P, Martin AC. Germline VH/VL pairing in antibodies. Protein Eng Des Sel 2012;25:523–529.
29. Dunbar J, Fuchs A, Shi J, Deane CM. ABangle: characterising the VH-VL orientation in antibodies. Protein Eng Des Sel 2013;26:611–620.
30. Almagro JC, Beavers MP, Hernandez-Guzman F, Maier J, Shaulsky J, Butenhof K, Labute P, Thorsteinson N, Kelly K, Teplyakov A, Luo J, Sweet R, Gilliland GL. Antibody modeling assessment. Proteins 2011;79:3050–3066.
31. Newton K, Matsumoto ML, Wertz IE, Kirkpatrick DS, Lill JR, Tan J, Dugger D, Gordon N, Sidhu SS, Fellouse FA, Komuves L, French DM, Ferrando RE, Lam C, Compaan D, Yu C, Bosanac I, Hymowitz SG, Kelley RF, Dixit VM. Ubiquitin chain editing revealed by polyubiquitin linkage-specific antibodies. Cell 2008;134: 668–678.
32. Maun HR, Wen X, Lingel A, de Sauvage FJ, Lazarus RA, Scales SJ, Hymowitz SG. Hedgehog pathway antagonist 5E1 binds hedgehog at the pseudo-active site. J Biol Chem 2010;285:26570–26580.
33. Zhao Z, Worthylake D, LeCour L Jr, Maresh GA, Pincus SH. Crystal structure and computational modeling of the fab fragment from a protective anti-ricin monoclonal antibody. PLoS One 2012;7:e52613.
34. Ni YG, Di Marco S, Condra JH, Peterson LB, Wang W, Wang F, Pandit S, Hammond HA, Rosa R, Cummings RT, Wood DD, Liu X, Bottomley MJ, Shen X, Cubbon RM, Wang SP, Johns DG, Volpari C, Hamuro L, Chin J, Huang L, Zhao JZ, Vitelli S, Haytko P, Wisniewski D, Mitnaul LJ, Sparrow CP, Hubbard B, Carfí A, Sitlani A. A PCSK9-binding antibody that structurally mimics the EGF(A) domain of LDL-receptor reduces LDL cholesterol in vivo. J Lipid Res 2011;52:78–86.
35. Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, Schwede T. Protein structure homology modeling using SWISS-MODEL workspace. Nat Protoc 2009;4:1–13.
36. Narayanan A, Sellers BD, Jacobson MP. Energy-based analysis and prediction of the orientation between light- and heavy-chain antibody variable domains. J Mol Biol 2009;388:941–953.
37. Freeman MM, Seaman MS, Rits-Volloch S, Hong X, Kao CY, Ho DD, Chen B. Crystal structure of HIV-1 primary receptor CD4 in complex with a potent antiviral antibody. Structure 2010;18:1632–1641.
38. Mahalingam B, Ajroud K, Alonso JL, Anand S, Adair BD, Horenstein AL, Malavasi F, Xiong JP, Arnaout MA. Stable coordination of the inhibitory Ca2+ ion at the metal ion-dependent adhesion site in integrin CD11b/CD18 by an antibody-derived ligand aspartate: implications for integrin regulation and structure-based drug design. J Immunol 2011;187:6393–6401.
39. Stanfield R, Cabezas E, Satterthwait A, Stura E, Profy A, Wilson I. Dual conformations for the HIV-1 gp120 V3 loop in complexes with different neutralizing fabs. Structure 1999;7:131–142.
40. Larsen NA, Zhou B, Heine A, Wirsching P, Janda KD, Wilson IA. Crystal structure of a cocaine-binding antibody. J Mol Biol 2001; 311:9–15.
41. Sawicki MW, Ng PC, Burkhart BM, Pletnev VZ, Higashiyama T, Osawa Y, Ghosh D. Structure of an activity suppressing Fab fragment to cytochrome P450 aromatase: insights into the antibody-antigen interactions. Mol Immunol 1999;36:423–432.
42. Tramontano A, Chothia C, Lesk AM. Framework residue 71 is a major determinant of the position and conformation of the second hypervariable region in the VH domains of immunoglobulins. J Mol Biol 1990;215:175–182.
43. Brightbill HD, Jeet S, Lin Z, Yan D, Zhou M, Tan M, Nguyen A, Yeh S, Delarosa D, Leong SR, Wong T, Chen Y, Ultsch M, Luis E, Ramani SR, Jackman J, Gonzalez L, Dennis MS, Chuntharapai A, DeForge L, Meng YG, Xu M, Eigenbrot C, Lee WP, Refino CJ, Balazs M, Wu LC. Antibodies specific for a segment of human membrane IgE deplete IgE-producing B cells in humanized mice. J Clin Invest 2010;120:2218–2229.
44. Chen X, Kim TD, Carman CV, Mi LZ, Song G, Springer TA. Structural plasticity in Ig superfamily domain 4 of ICAM-1 mediates cell surface dimerization. Proc Natl Acad Sci USA 2007;104:15358–15363.
45. Tormo J, Blaas D, Parry NR, Rowlands D, Stuart D, Fita I. Crystal structure of a human rhinovirus neutralizing antibody complexed with a peptide derived from viral capsid protein VP2. EMBO J 1994;13:2247–2256.
46. Luftig MA, Mattu M, Di Giovine P, Geleziunas R, Hrin R, Barbato G, Bianchi E, Miller MD, Pessi A, Carfí A. Structural basis for HIV-1 neutralization by a gp41 fusion intermediate-directed antibody. Nat Struct Mol Biol 2006;13:740–747.
47. Obmolova G, Teplyakov A, Malia TJ, Grygiel TL, Sweet R, Snyder LA, Gilliland GL. Structural basis for high selectivity of anti-CCL2 neutralizing antibody CNTO 888. Mol Immunol 2012;51:227–233.
48. Sivasubramanian A, Sircar A, Chaudhury S, Gray JJ. Toward high-resolution homology modeling of antibody Fv regions and application to antibody-antigen docking. Proteins 2009;74:497–514.
49. Stanfield RL, Zemla A, Wilson IA, Rupp B. Antibody elbow angles are influenced by their light chain class. J Mol Biol 2006;357:1566–1574.
50. Yin J, Andryski SE, Beuscher AE 4th, Stevens RC, Schultz PG. Structural evidence for substrate strain in antibody catalysis. Proc Natl Acad Sci USA 2003;100:856–861.