

Assessment of fully automated antibody homology modeling protocols in molecular operating environment

Johannes K. X. Maier* and Paul Labute

Chemical Computing Group, Inc., Montreal, Quebec, H3A 2R7, Canada

ABSTRACT

The success of antibody-based drugs has led to an increased demand for predictive computational tools to assist antibody engineering efforts surrounding the six hypervariable loop regions making up the antigen binding site. Accurate computational modeling of isolated protein loop regions can be quite difficult; consequently, modeling an antigen binding site that includes six loops is particularly challenging. In this work, we present a method for automatic modeling of the F_V region of an immunoglobulin based upon the use of a precompiled antibody x-ray structure database, which serves as a source of framework and hypervariable region structural templates that are grafted together. We applied this method (on common desktop hardware) to the Second Antibody Modeling Assessment (AMA-II) target structures as well as an experimental specialized CDR- H_3 loop modeling method. The results of the computational structure predictions will be presented and discussed.

Proteins 2014; 82:1599–1610.

© 2014 The Authors. Proteins published by Wiley Periodicals, Inc.

Key words: homology modeling; comparative modeling; antibody structure; antibody engineering; CDR- H_3 loop modeling; loop enrichment; loop refinement.

INTRODUCTION

The hallmark feature of antibodies is their ability to bind an almost unlimited collection of target structures with remarkably high affinity. This high target specificity combined with modular composition into distinct functional and structural domains makes immunoglobulins particularly attractive for use as drugs. Antibodies have been successfully applied in therapeutic contexts for the last 20 years and there are currently hundreds of antibody-based drugs in the late stage of clinical trials.¹ The focus of engineering efforts revolves around the F_V region of the immunoglobulin, the smallest fragment of the antibody that retains antigen binding ability. One particular challenge in modeling an antigen-binding site is that the regions of interest—the six hypervariable loop regions—are designed by nature to be diverse, whereas the region outside the antigen binding loops, the framework region (FR), is structurally characterized by a conserved beta barrel fold.

Accurate computational modeling of isolated protein loop regions can be quite difficult; consequently, modeling an antigen binding site that includes six loops is particularly challenging. Fortunately, five of the six hypervariable “CDR” loops, three of the light chain

(CDR- L_1 , CDR- L_2 , and CDR- L_3) and two of the heavy chain (CDR- H_1 and CDR- H_2) assume a rather small repertoire of main chain conformations and are therefore called “canonical” loops^{2–4} leaving the central challenge of modeling the CDR- H_3 loop. A further complication is the fact that the antigen-binding site has to be modeled as a dimer composed of the variable domains of the light V_L and heavy V_H chains. Given that the CDR loops bind to the antigen at the interface of the V_L and V_H chains, their relative orientation also affects binding and can create an additional source of diversity.

A computational assessment of various antibody modeling programs was held in 2009,⁵ the First Antibody Modeling Assessment (AMA-I). Our various submissions in AMA-I were characterized by the amount of user intervention necessary to model the target as well as a control protocol that used no antibody-specific

*Correspondence to: Johannes K. X. Maier, Schrodinger Inc, 120W 45th Street, NY 10036-4041. E-mail: johannes.maier@schrodinger.com

Received 5 December 2013; Revised 12 March 2014; Accepted 29 March 2014

Published online 23 April 2014 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.24576

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

knowledge. Our primary fully automated protocol used in AMA-I, referred to as “autoFv,” has undergone several modifications since its first implementation and is the standard for automated antibody homology modeling in the Molecular Operating Environment⁶ (MOE) software.

In this work, we report on our contributions to the Second Antibody Modeling Assessment (AMA-II) that consisted of two stages: Stage 1 involved full structure modeling (as in the AMA-I) while Stage 2 focused on H₃ modeling. The autoFv protocol (also referred to as “cgg3” in this work) was used along with other prototype protocols to be described below. Our efforts for AMA-II were directed at rapid automated modeling of antibodies; that is, running on common desktop hardware without user intervention. We describe our methods in the Materials and Methods section. In the next section, the results of AMA-II computational experiments are presented and discussed. Conclusions are drawn in the final section.

MATERIALS AND METHODS

The MOE antibody modeler is based upon the use of a precompiled antibody structure database, the “Fab Database” (see below). This database is used as a source of framework and hypervariable region structural templates, which are grafted together (see below). Typically, the Fab Database will contain Protein Data Bank^{7,8} (PDB) structures augmented with additional proprietary structures, if available. In this work, only PDB structures available as of December 2012 were used, a total of 1,969 structures. Given the V_L and V_H amino acid sequences as input, the automated antibody modeling protocol (autoFv) proceeds as follows:

1. Search the Fab Database sequences for collections of candidate templates for each of V_L, V_H, and the FR $F_V = (V_L, V_H)$ as well as individual CDR loops.
2. Build 10 crude models from combinations of the template candidates. Each model is built by grafting the loop templates onto the template V_L and V_H framework chains. In case of hetero template composition, the V_L and V_H grafted templates are transposed via backbone superposition to F_V dimer coordinates of the highest ranking (see below) F_V dimer of the V_L + V_H search results. Side chains are modeled with a rotamer library and coordinates are refined using forcefield energy minimization to relieve strained geometries and clashes.
3. Build a consensus model from the pool of crude models. The model with the lowest F_V binding energy is used as the consensus structural core.⁹ The binding energy of the F_V complex is calculated as $E(V_L:V_H) - E(V_L) - E(V_H)$. Cluster the CDR conformations in the pool and select the best representative (see below) for each CDR loop. Graft the selected loops onto the

consensus structural core using the protocol for building the crude models.

This autoFv procedure is termed “cgg3” for AMA-II; the “cgg1” and “cgg2” prototype protocols are derivatives of cgg3 that will be described after the details of the cgg3/autoFv protocol are described. All three protocols are summarized in the flow chart of Figure 1.

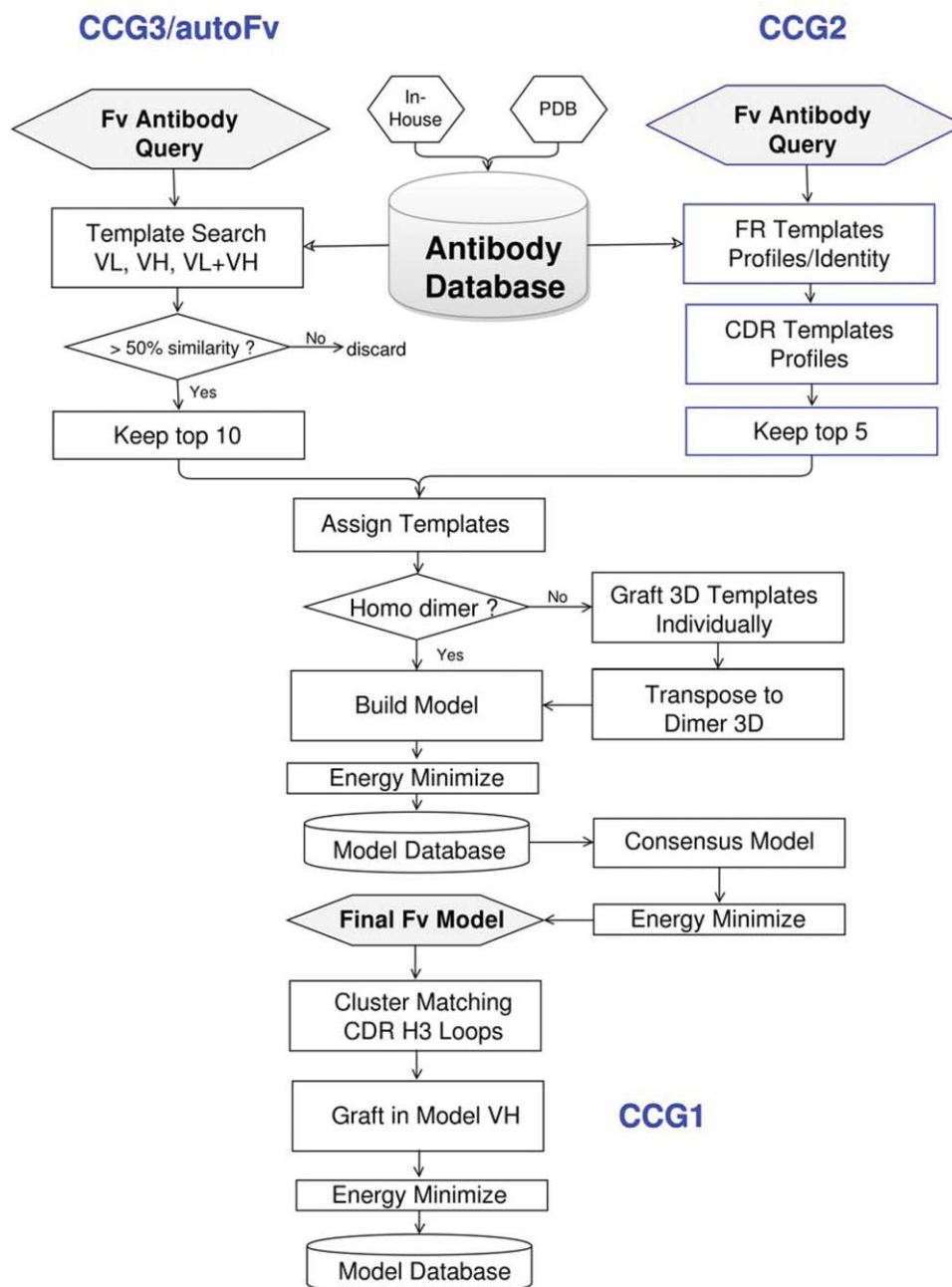
Fab database compilation

The Fab Database is a collection of antibody structures determined by x-ray crystallography. It was assembled using an automated procedure that identifies F_V domains from the PDB (or collection of proprietary structures). The antibody detection and classification system is based on a previously published¹⁰ collection of V_L and V_H reference sequences enriched with antibody sequences submitted to the PDB with no restrictions concerning type or species. The reference sequences were aligned using MOE, followed by single-linkage clustering to extract antibody clusters of maximum diversity in the framework sequence. The V_L reference set is composed of 101 antibody structures with a maximum framework sequence identity of 85% and the V_H set of 85 structures at a maximum framework identity of 75%.

The entire PDB was then screened for immunoglobulin structures. The sequence of each protein chain is aligned against the V_L and V_H reference sets and the sequence identity is calculated as a percentage of identical residues within the aligned F_V domain as well as within the respective FRs FR₁, FR₂, and FR₃. Protein chains are classified as antibody chains if the overall sequence identity with one chain of the reference collections exceeds 40%. Classifying an F_V chain as either V_L or V_H requires a framework sequence identity of at least 60% with one class and <50% with the opposite class. The cutoffs in the antibody detection as described above ensured no contamination of other immunoglobulin-like structures such as T-cell receptor sequences. If detected as a V_L chain, further subclassification into either κ or λ subclasses is achieved by aligning to reference germline sequences originating from the Immunogenetics database.^{11,12} For structural clustering, the Structural Composition of Protein database SCOP^{13,14} was used as seed clusters followed augmented with single linkage clustering.

Template search

Given the V_L and V_H amino acid sequences of the antibody to be modeled, the Fab database is searched for suitable templates. Separate results are collected for V_L, V_H, and V_L + V_H as well as the hypervariable loop regions. A sequence alignment with the BLOSUM62¹⁵ substitution matrix is performed. Two different scores are used for framework and hypervariable loop regions.

**Figure 1**

Workflow diagram of the automated antibody homology modeling protocol. The F_V consensus model is built based on a collection of either 10 (ccg3) or 5 (ccg2) individual models. The steps that are specific to the ccg2 protocol are outlined in dark blue color; ccg1 is an extension step to ccg2 or ccg3 in generating additional loop conformation in CDR- H_3 .

The scores for the FRs are calculated as the identity with prealigned framework clusters in the Fab Database, factoring gaps into the calculation. The hypervariable loop scores are calculated as similarity of the query loop compared to the corresponding hypervariable loop sequences of the same length in the Fab Database.¹⁶ Since the template search is confined to the content of the Fab Database, the time required to find suitable templates

takes on average <0.5 s per query. Only the top 10 scoring framework candidates from each of respective result set are retained. If the sequence identity of a $V_L + V_H$ result is within 10% of the target sequence then a homo template (V_L and V_H belong to the same structure) is used, otherwise hetero templates (V_L and V_H from different structures) are considered. For the CDRs, only the top 10 ranking candidates at a minimum sequence

similarity of 50% are retained. In case the similarity cutoff in a CDR region is not met, and no templates are found as a result, the loop will be modeled using the general loop building procedure of the MOE Homology Modeler instead of the particular antibody techniques described here.

Template and model construction

Antibody template structures are built by grafting the framework and hypervariable backbone geometries together. The selected CDR template chains are superimposed in the backbone atoms (N, C α , C, and O) onto the target framework template within three framework residues on either side of the CDR region followed by tethered energy minimization to relax strained geometries in the transition. In case of a hetero template composition, the templates are transposed via backbone superposition to F_V dimer coordinates of the highest-ranking F_V dimer of the $V_L + V_H$ results of the template search. The MOE general Homology Modeler was used to perform sidechain repacking of CDR loop sidechain rotamers as well as framework residues with van der Waals clashes (>1.5 kcal/mol) with the grafted loop residues. For this modeling exercise, all CDR residues were repacked using the conformations from the MOE rotamer library exclusively.

For the consensus model, single-linkage clustering is applied to determine the conformational spread within the respective CDR loops of the models in the pool built above. The near neighbor cutoff for the CDR conformation clusters is set to 0.4 Å for the canonical loops L_1 , L_2 , L_3 , and H_1 , 0.6 Å for the H_2 and 1.5 Å for H_3 . The loop candidate with the highest structure score within the most frequent CDR conformation was selected as input for the consensus core structure. The structure score, S , is a geometric mean with $S = (t^3 g^2 ob)^{1/7}$ where t , g , o , and b are values between 0 and 1 as follows; t assesses backbone topology, bond lengths, angles etc.; g is the Ramachandran phi/psi probability; o assesses crystallographic occupancies; b assesses temperature factors. Each of t , g , o , and b were calibrated by setting 0 to correspond to the lower range of outliers and 1 the upper range of outliers as measured from PDB statistics. Values of S are interpreted as 0 meaning completely unsuitable for modeling purposes and 1 meaning ideal. Only structures with $S > 0.65$ were used in this work.

For the canonical CDR loops, the loop conformations of the crude models are used as input for conformational analysis in each CDR loop category, recruiting the most common conformation as input for the consensus model. The conformational analysis helps filter CDR loop conformations that fit particularly well within a distinct modeling environment, shifting the selection process away from general sequence criteria to structurally more specific ones. Consequently, CDR loops with lower

sequence identity (but higher structure scores) may be selected for assembly in the consensus model. In addition, the crude models serve as a pool to address complex structural issues such as the $V_L:V_H$ dimer orientation and the variability in CDR loop conformations that are challenging to predict based on sequence alone.

Protocol ccg2

The ccg2 protocol is a variation of the autoFv/ccg3 protocol described above. First, only five crude models are built (instead of the usual 10), which reduces the overall modeling time significantly. Second, the canonical CDR loops templates are selected based on position-based profile scores of the respective CDR loop clusters in a special canonical loop database (see below) that match the length of the CDR query; only the top five most sequence-similar candidates are retained. For H_3 , only the top five template candidates based on sequence similarity calculations are retained.

The canonical CDR loop templates of the ccg2 protocol are selected from a dedicated loop database. All canonical CDR loops extracted from the Fab Database were grouped by CDR type and loop length. Templates with structural issues in the protein backbone and/or low phi/psi scores were discarded. The remainder were clustered based on backbone RMSDs at a resolution of 0.3 Å employing a single-linkage clustering algorithm followed by a sequence profile calculation that determined the frequencies of unique residues in each cluster column. The resulting values were normalized and used as weights for compiling a position score for each residue in the cluster and used when matching the template queries to the sequence profiles of each cluster. A similar approach using Hidden Markov models has been previously reported.¹⁷ In this way, residues that are more preserved (and thus more relevant for determining the actual conformation) within a cluster have a stronger impact on the selection.

The selection of framework candidates is based on position-based profile scores with residues shown to be involved in F_V dimer interactions. The orientation between V_L and V_H is a particular challenge in modeling an antigen-binding site.^{18,19} Narayanan *et al.* (2009) used the C α RMSD in the FR of V_L when superposed on the V_H framework as a metric to quantify the relative changes in the $V_L:V_H$ orientation and also demonstrated that the native $V_L:V_H$ orientations correspond with the energy minima measured as F_V complex interaction energies. We make use of this observation along with the residues identified by Abhinandan and Martin²⁰ in the framework (L38, L40, L41, L44, L46, L87, H33, H42, H45, H60, H62, H91, H105, and Chothia numbering) for the profile scores. With this distinct and reduced set of residues in place that are (intended to be)

determinants of the F_V dimer orientation, the framework template candidates were identified based on the number of identical residues within this distinct set between the templates. The overall framework sequence identity was used as a secondary metric to determine the final FR template selection.

Protocol ccg1

The ccg1 is a derivative of the ccg2 protocol in which the CDR- H_3 loop diversification step is applied using the Fab Database described above. The database was first searched for H_3 loops identical in length with the query CDR- H_3 . Single-linkage clustering was applied based on backbone RMSD at a resolution of 0.5 Å to determine distinct sets of CDR- H_3 loop conformations. The candidate with the best structure score within each conformation cluster was used as input for template assembly into the V_H chain of the consensus model or native structure. A model based on each cluster was built in the manner described above. For each resulting model, $V_L:V_H$ binding energies and conformational energies were calculated. Models with more than 10 clashes >0.5 kcal/mol or positive conformation energies were discarded. The three models with the lowest binding energies were output.

Manual modeling of target MA2-1 in AMA-II Stage 1

The lack of suitable structure templates for CDR L_1 , L_3 , and H_3 combined with low sequence identity in the FR across species precluded a fully automated modeling approach for target MA2-1. Instead, this target was modeled using the interactive MOE Antibody Modeler application using the default settings. This application suggested PDB:2VUO—one of the few rabbit immunoglobulins in the PDB—as the only suitable structure template. The top 3 ranking models based on geometry assessments in the protein backbone and overall energy scores of 100 coarse models were used for CDR conformation diversification in CDR- L_1 , CDR- L_3 , and CDR- H_3 using LowModeMD.²¹ The backbone atoms N, $C\alpha$, C, and O of the CDR loops L_2 , H_1 , and H_2 as well as the framework atoms were kept fixed, allowing only the sidechain atoms to move and thus accommodating novel conformations. The procedure was stopped after 300 distinct conformations were generated. The candidates with the lowest energies were selected and subjected to a final structure refinement procedure as described above.

CDR- H_3 modeling in AMA-II Stage 2

Stage 2 focused on modeling CDR- H_3 when given x-ray crystal structures (with H_3 removed) provided by the organizers of the modeling assessment. Our efforts involved the use of a dedicated CDR- H_3 loop database.

This database was searched to find H_3 loops (of matching length) that superposed to an RMSD of 0.25 Å to the surrounding three framework backbone atoms. Sidechains were repacked as described above and the coordinates were refined using a tethered protocol. During refinement, all heavy atoms of the input x-ray structure were held fixed allowing only the H_3 loop atoms to move within the antigen-binding complex. The minimization procedure was repeated five times, gradually releasing the tethers on the backbone atoms of H_3 . Finally, models with unresolvable clashing energies >2.5 kcal/mol were discarded and the best scoring (see below) conformation was selected. On average, the processing time of one conformation consumed ~1 CPU minute.

For the compilation of the CDR- H_3 loop database, all H_3 loops were extracted from the Fab Database and clustered according to length and backbone RMSD. Single-linkage clustering applied based on a backbone RMSD cutoff of 0.4 Å to determine distinct sets of CDR- H_3 loop conformations. Structures with positive forcefield energies, poor (<85%) Ramachandran map probability, or a clash count of >5 were discarded. The cluster member with the highest score (see below) was used as the representative. The antibody dimer structures of these CDR- H_3 representatives were then extracted from the Fab Database for further diversification of the loop conformation space by subjecting them to conformational search using LowModeMD. All residue atoms beyond a perimeter of 5.5 Å of CDR- H_3 loop atoms were kept fixed, as well as the backbone atoms of non- H_3 residues within the perimeter. This left the H_3 loop atoms, including three flanking framework residues and all sidechain atoms within the perimeter, free. LowModeMD was terminated after 100 attempts to generate a novel conformation up to a maximum or when 50 unique structures were generated. The structural scores of the newly generated H_3 conformations were calculated discarding conformations with inappropriate structure/clashing scores. The remainders were clustered at a resolution of 0.4 Å and the conformation representatives were merged with the corresponding conformation pool extracted from the Fab Database.

As a control alternative to energy-based scores used for ranking output models in Stage 1, a scoring function was developed factoring in H_3 -specific properties like accessible surface area, psi/phi values and interaction energies of CDR- H_3 to the F_V dimer complex. The prepared representative F_V dimer structures extracted above from the Fab Database were used as raw data. For each loop length category (ranging from 4 to 14 residues), the respective components were calculated, normalized, and the frequency data was fitted to a Gaussian distribution. The final CDR- H_3 loop score was calculated as the geometric mean of the three individual loop scores.

Hardware and software

All calculations (and validation analyses) in this work were conducted using the MOE version 2012.10 on two computers, a Mac Pro, Dual Core Intel Xeon at 2.66 GHz, OS X 10.5 (Leopard) and a Dell Inspiron, Intel 2 Dual Core on OS Linux Mint 13. Energy calculations and coordinate refinement were performed using the AMBER 12²² forcefield parameters augmented with the GB/VI implicit solvent model²³ with a nonbonded cutoff of 8 Å. PDB structures were prepared for simulation work using the MOE Structure Preparation application followed by Protonate3D.²⁴

RESULTS AND DISCUSSION

AMA-II Stage 1

The Stage 1 models (cgg1, cgg2, and cgg3) were built using the methods described in the previous section with the exception of target MA2-1, a rabbit antibody, which was modeled manually due to a lack of templates in the PDB. Each final cgg3 model required 20–50 CPU minutes to compute and 10–25 CPU minutes per final model for cgg2 and cgg3 protocols with most of the time spent in coordinate refinement. All of the protocols presented here produced models of good structural quality evidenced by an overall MolProbity score average of 1.5, lower than the 1.8 average we have calculated for 1,969 x-ray structures from the PDB (our Fab Database). A detailed listing of the various MolProbity measurement components and comparisons to the other participants are provided in the Supporting Information Tables of Topalev *et al.* in this issue.

Table I presents a comparison of the cgg1–3 structures with the (previously unknown) x-ray coordinates of the targets of the second antibody modeling assessment. The RMSD values measured over the entire framework of all 10 models (FR) reveals a more robust and overall slightly better performance of the established cgg3/autoFv protocol (0.80 ± 0.19 Å RMSD), compared with cgg2 (0.87 ± 0.24 Å RMSD). However, the discrepancy is largely caused by a poor cgg2 template choice for modeling target MA2-3 with a RMSD of 1.35 Å compared with 0.57 Å of the corresponding cgg3 model. The low agreement in dimer orientation is further evidenced by large tilt angles (14.3°) and large discrepancies in the HL torsion and HC₂ bend angles (7.8° and 7.9° difference to the x-ray structure). Our analysis revealed that the failure in selecting an appropriate framework template for this target was not solely rooted in a failure of the FR scoring function that was applied in selecting the five template candidates to build the crude models, but rather in the unusual, “open” conformation of CDR-H₃ in one of the models. Unfortunately, this unusual conformation also measured the lowest F_V complex binding energy among

the five crude models that is the driving parameter for the framework selection of the consensus model. Further analysis of the MolProbity scores (Topalev *et al.* in this issue) and the all-atom energies of the cgg2 model are reasonable for all of the measured components suggest that lower interaction energies were not obtained at the expense of unrealistic structures. It is possible that a larger collection of crude models as used by cgg3/autoFv (ten instead of five) offers better protection against selecting a particularly poor framework template. However, given our findings that CDR-H₃ residues are strongly influential in the interaction between V_L and V_H chains, the issues might be more complex and point potentially to a chicken and egg problem by which the orientation of the V_L and V_H chains may only be predicted accurately if the correct conformation of CDR-H₃ is known. Even with inclusion of the poor modeling result of target MA2-3, the measurements in the FR are promising and comparable with the other modeling approaches presented in this contest (see Almagro overview in this issue). As the cgg2 modeling protocol was still under development at the time of AMA-II, we are optimistic that further fine-tuning of the scoring parameters will produce more stable results that are in keeping with those demonstrated with cgg3/autoFv over the last few years. The even more experimental cgg1 protocol fared worst of the three. Since cgg1 used the output of cgg2 it would appear that the additional steps were detrimental and further investigation is required to diagnose the cause. An interesting possibility is cgg1's additional coordinate refinement with CDR-H₃ backbone atoms fixed (letting the other CDR loops adapt to the backbone of CDR-H₃), which may introduce negative effects on the other CDR loops (e.g., CDR-L₁ in MA2-2, MA2-3 and MA2-10 and to a lesser degree in CDR-L₃ in MA2-3).

Since the distinct arrangement of CDR loops is a critical factor in the formation of the antigen/antibody complex, CDR measurements must be assessed on two levels—the first with respect to their orientation within the dimer and the second with respect to the accuracy of the individual loop conformations. Assessment is compromised, however, when the entire antigen-binding surface (all Loops) is considered due to the influence of the CDR-H₃ loop component. Accordingly, the canonical loops are assessed first (Table I, Canonical) and, in this measurement component, these show competitive results with an average cgg3/autoFv of 1.00 ± 0.29 Å RMSD for all 10 models. The RMSDs are almost identical between the protocols ($\sim 1.0 \pm 0.25$ Å) with some interesting variations in specific cases. For example, larger RMSD outliers of single models are responsible for the variations in the light chain averages (CDR-L₁ in MA2-7, 1.25 vs. 0.37 Å and 0.58 vs. 0.27 Å in MA2-8 for CDR-L₂), whereas overall improvement in modeling CDR-H₁ and CDR-H₂ is observed with cgg2 over cgg3. As expected, the less

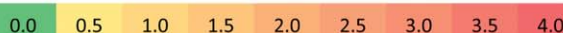
Table I

RMSD Values in Å Between Three CCG Protocols (See Text) and the Stage 1 Target X-Ray Crystal Coordinates; “All” is the RMSD on the Entire F_V Region; “FR” Is the RMSD on the Framework (Excluding the CDR Loops); “Canonical” Is the RMSD on CDR L_1-L_3 and H_1-H_2 ; “ L_1-L_3 ” and “ H_1-H_3 ” Are the RMSDs of Each Individual CDR Loop Region

	All / All			FR / FR			Canonical / Canonical		
	cgg1	cgg2	cgg3	cgg1	cgg2	cgg3	cgg1	cgg2	cgg3
MA2-1	1.72	1.61	1.81	0.88	0.87	0.80	2.52	2.05	2.71
MA2-2	1.31	1.05	1.16	0.80	0.74	0.64	0.95	0.85	0.84
MA2-3	1.61	1.53	0.93	1.35	1.35	0.57	1.37	1.35	1.58
MA2-4	1.16	1.2	1.31	1.13	1.13	0.86	1.15	1.10	1.14
MA2-5	1.09	0.99	1.13	0.87	0.88	1.01	1.33	1.16	1.21
MA2-6	1.12	1.44	0.97	0.68	0.69	0.75	0.73	0.70	1.14
MA2-7	1.04	1.03	0.79	0.85	0.86	0.62	0.94	0.92	0.72
MA2-8	1.11	1.13	1.03	0.84	0.83	0.71	0.94	1.01	0.63
MA2-9	0.62	0.59	1.14	0.48	0.46	1.14	0.67	0.62	0.86
MA2-10	1.13	1.05	0.87	0.78	0.76	0.68	0.79	0.79	0.77
MA2-11	1.23	1.1	1.12	0.97	0.96	0.93	1.21	1.07	1.14
average	1.14	1.11	1.05	0.88	0.87	0.79	1.01	0.96	1.00
stdev	0.25	0.26	0.16	0.24	0.24	0.19	0.25	0.22	0.29

	L1 / L1			L2 / L2			L3 / L3		
	cgg1	cgg2	cgg3	cgg1	cgg2	cgg3	cgg1	cgg2	cgg3
MA2-1	2.09	0.93	3.13	0.27	0.24	0.25	2.39	2.94	2.85
MA2-2	0.61	0.39	0.35	0.39	0.29	0.26	0.86	0.76	0.72
MA2-3	0.55	0.24	0.73	0.36	0.29	0.19	0.87	0.59	0.21
MA2-4	0.76	0.65	0.68	0.45	0.49	0.33	0.44	0.29	0.48
MA2-5	0.51	0.48	0.45	0.26	0.24	0.21	1.48	1.47	1.45
MA2-6	0.46	0.36	0.38	0.34	0.25	0.28	0.44	0.51	0.35
MA2-7	1.21	1.25	0.37	0.30	0.25	0.29	0.49	0.42	0.42
MA2-8	0.45	0.49	0.35	0.39	0.58	0.27	0.36	0.26	0.31
MA2-9	0.71	0.62	0.38	0.35	0.18	0.25	0.43	0.35	0.42
MA2-10	0.69	0.37	0.43	0.33	0.36	0.22	0.48	0.42	0.31
MA2-11	0.49	0.27	0.21	0.20	0.17	0.22	0.39	0.34	0.29
average	0.64	0.51	0.43	0.34	0.31	0.25	0.62	0.54	0.50
stdev	0.23	0.29	0.16	0.07	0.13	0.04	0.35	0.36	0.36

	H1 / H1			H2 / H2			H3 / H3		
	cgg1	cgg2	cgg3	cgg1	cgg2	cgg3	cgg1	cgg2	cgg3
MA2-1	1.00	0.92	0.80	0.20	0.25	0.34	2.88	2.33	2.94
MA2-2	0.51	0.35	0.36	0.50	0.74	1.08	1.89	2.70	1.75
MA2-3	1.79	1.91	2.64	0.49	0.49	0.29	2.02	1.26	0.93
MA2-4	0.53	0.40	0.69	0.52	0.43	0.57	1.01	1.33	3.11
MA2-5	0.26	0.24	0.28	0.25	0.23	0.15	1.94	1.22	1.62
MA2-6	0.41	0.30	0.57	0.47	0.34	0.44	2.58	2.17	1.45
MA2-7	0.25	0.22	0.20	0.21	0.18	0.11	2.23	2.05	1.24
MA2-8	0.28	0.16	0.18	0.34	0.37	0.90	1.87	2.25	2.51
MA2-9	0.32	0.18	0.27	0.18	0.32	0.39	1.34	1.17	0.99
MA2-10	0.37	0.28	0.31	0.40	0.26	0.38	1.96	2.07	1.42
MA2-11	0.55	0.24	0.37	0.50	0.32	0.43	1.69	1.82	1.82
average	0.53	0.43	0.59	0.39	0.37	0.47	1.85	1.80	1.68
stdev	0.46	0.53	0.74	0.13	0.16	0.31	0.44	0.53	0.68



All RMSD measurements (in Å) are based on backbone atoms N, C α , C, and O.

canonical and particularly challenging CDR loops (CDR- L_1 and CDR- L_3 of target MA2-1, CDR- H_1 of MA2-3 and CDR- L_3 of MA2-5) show lower precision but within the range of the results in the modeling competition (see Table I and Teplyakov and Almagro this issue). Promising results are seen in CDR loops with low agreements to the North CDR clustering scheme such as the CDR- L_3 loops of MA2-2 and MA2-3, CDR- L_1 of MA2-4 and MA2-5 and CDR- H_2 of MA2-6 and MA2-8 where the

RMSD values are within the modeling precision of less challenging cases. An exception is CDR- H_2 of MA2-8 that has only been well modeled with the cgg2 protocol (0.35 vs. 0.90 Å RMSD). As mentioned in Teplyakov and Almagro (this issue), improvements could be made in addressing the *cis/trans* configuration of certain proline residues in the CDRs in more detail.

A quality assessment of the antigen-binding surface is not complete without the analysis of the key CDR- H_3 .

Table II

RMSD Values in Å (Local and Global) of H₃ Between Three CCG Protocols (See Text) and the Stage 1 Target X-Ray Crystal Coordinates; “Lowest” Is the Smallest RMSD Value of Intermediate Models Generated in the Protocols; “cgg1–3” Are the RMSD Values of the Three Protocols; “Lowest E” Is the RMSD of the Intermediate Structure with the Lowest Forcefield Energy for cgg1

	lowest	cgg1	cgg2	cgg3	lowest	cgg1	cgg2	cgg3	lowest E
MA2-2	1.15	1.89	2.70	1.75	2.51	4.78	3.17	4.56	3.80
MA2-3	0.53	2.02	1.26	0.93	1.24	3.99	2.73	1.11	3.42
MA2-4	0.68	1.01	1.33	3.11	1.09	2.06	1.87	4.84	1.66
MA2-5	0.70	1.94	1.22	1.62	1.09	3.30	2.48	2.40	3.29
MA2-6	1.58	2.58	2.17	1.45	2.45	3.95	4.87	2.34	3.04
MA2-7	0.44	2.23	2.05	1.24	1.16	3.11	2.91	2.09	3.07
MA2-8	1.23	1.87	2.25	2.51	1.61	3.19	3.54	3.37	2.97
MA2-9	0.35	1.34	1.17	0.99	0.66	1.52	1.49	1.19	2.65
MA2-10	0.90	1.96	2.07	1.42	1.91	3.65	3.18	2.06	2.56
MA2-11	0.95	1.69	1.82	1.82	1.72	3.52	2.46	2.70	2.45
average	0.85	1.85	1.80	1.68	1.54	3.31	2.87	2.67	2.89
stdev	0.39	0.44	0.53	0.68	0.61	0.94	0.94	1.26	0.60
	local rmsd				global rmsd				

Global RMSDs denote loop measurements after superposition of the framework first.

The unusually broad range in loop lengths, the conformational flexibility as well as its influence on other areas of the antigen-binding complex warrants a separate evaluation. Since it is not yet possible to employ any type of loop conformation dictionary in modeling this loop accurately, it is probably best to offer a set of plausible conformations as a representation of CDR-H₃ rather than a single conformation providing only the sequence as input. To this end, the main objective of the cgg1 protocol is to provide a diverse, structurally sound, and energetically plausible collection of CDR-H₃ conformations. On average, 55 alternate and diverse conformations were generated in approximately one CPU hour of computation on a common desktop computer. While this approach demonstrates that the primary objective was satisfied, the H₃ modeling results in Table II reveal that it was not possible to improve the modeling predictions of CDR-H₃ compared with the output conformations of cgg2 and cgg3 (which performed best). The cgg1 conformation diversification step does not fail in producing reasonably accurate conformations (average global RMSD of 1.54 ± 0.61 Å) but rather in selecting the best, or close to the best, solution within the generated conformation pool. The H₃ scoring function, trained on loop-specific properties like psi/phi torsion values, F_V interaction energies and surface area was implemented as a relative metric for ranking CDR-H₃ loop conformations but failed in predicting CDR-H₃ conformations accurately (3.31 ± 0.94 Å). By default, the cgg1 protocol uses energies as a ranking criterion for the models in addition to the loop score. Selecting the lowest energy model for predicting H₃ would have improved the average RMSD value to 2.89 ± 0.57 Å (and thus within the

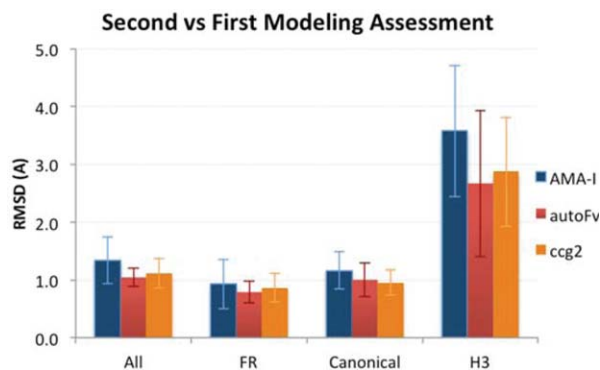
assessment) but would have also failed to improve the prediction accuracy of CDR-H₃ of the input models cgg2 or cgg3/autoFv. Figure 2(A) lists the AMBER12 energies of the target x-ray structures and the global RMSDs in CDR H3 of the top 3 ranking cgg1 models (Models 1, 2, and 3) followed by the lowest RMSD model. The suffix indicates the position of the low RMSD conformer within the energy-ranked conformation set. The comparison of AMBER12 all-atom energies shows that the x-ray structures would have been discarded in all cases if present in the pool of cgg1 conformations. The correlation analysis [global RMSD (y-Axis) vs. AMBER12 all-atom energies (x-Axis)] as depicted in the plots of Figure 2(A) does not support the notion that prediction accuracy is linked to low energies of the models further substantiating our problems in discriminating the best or close to the best possible solutions. Taken together, these results suggest that modeling of CDR-H₃ still represents a serious challenge both for us as well as for the participants that use more computationally expensive methods for modeling H₃.

In summary, the fully automated modeling protocols cgg3/autoFv and cgg2 show good prediction accuracy in the structural core and the canonical CDR loop regions. In AMA-II, the more seasoned and established cgg3/autoFv modeling method displayed a more robust and slightly improved performance compared with cgg2 (overall RMSD average of 1.05 vs. 1.11 Å for cgg2 and 1.14 Å for cgg1 Table I) in contrast to an internal benchmark (data not shown). The variations we observe among these modeling procedures are more evident in the framework/dimer orientation category (see Table I FR section as well as Topalev *et al.* and the Supporting Information Section of

(A) Stage 1				(B) Stage 2			
Sample	AMBER12	RMSD	RMSD vs AMBER12	Sample	AMBER12	RMSD	RMSD vs AMBER12
MA2_2	1330.99	0.00	R=0.06 n=59	MA2_2	1330.99	0.00	R=0.03 n=990
Model_1	-96.24	3.80		Model_1	984.47	2.95	
Model_2	-70.89	4.60		Model_2	985.35	2.64	
Model_3	-65.17	4.43		Model_3	985.75	4.04	
LowRMSD_27	14.40	2.51		LowRMSD_401	1138.71	1.53	
MA2_3	173.82	0.00	R=0.18 n=30	MA2_3	173.82	0.00	R=0.36 n=633
Model_1	-139.13	3.42		Model_1	137.38	1.67	
Model_2	-136.93	4.39		Model_2	152.12	2.31	
Model_3	-135.13	1.90		Model_3	157.67	2.27	
LowRMSD_27	-54.82	1.24		LowRMSD_55	194.49	0.84	
MA2_4	705.83	0.00	R=0.58 n=32	MA2_4	705.83	0.00	R=0.40 n=641
Model_1	-605.30	1.66		Model_1	601.38	0.87	
Model_2	-570.95	2.18		Model_2	609.91	1.29	
Model_3	-558.16	1.81		Model_3	618.08	1.87	
LowRMSD_6	-536.82	1.24		LowRMSD_28	643.93	0.68	
MA2_5	432.08	0.00	R=0.46 n=28	MA2_5	432.08	0.00	R=0.40 n=639
Model_1	-185.64	3.29		Model_1	413.09	1.15	
Model_2	-183.36	2.54		Model_2	413.71	2.34	
Model_3	-177.93	3.02		Model_3	416.40	1.39	
LowRMSD_15	-117.48	1.09		LowRMSD_5	421.82	0.93	
MA2_6	1887.98	0.00	R=0.51 n=36	MA2_6	1887.98	0.00	R=0.09 n=228
Model_1	-691.90	3.04		Model_1	1124.26	3.64	
Model_2	-682.46	4.84		Model_2	1127.10	3.10	
Model_3	-677.31	3.02		Model_3	1143.01	2.44	
LowRMSD_32	-548.54	2.45		LowRMSD_18	1181.74	1.66	
MA2_7	203.68	0.00	R=0.22 n=25	MA2_7	203.68	0.00	R=0.21 n=641
Model_1	-12.19	3.07		Model_1	98.73	1.40	
Model_2	0.88	2.76		Model_2	100.60	1.00	
Model_3	2.06	1.16		Model_3	118.21	1.08	
LowRMSD_3	2.06	1.16		LowRMSD_37	172.58	0.87	
MA2_8	275.46	0.00	R=0.09 n=72	MA2_8	275.46	0.00	R=0.07 n=1160
Model_1	-153.82	2.97		Model_1	286.53	6.36	
Model_2	-143.09	3.27		Model_2	288.04	3.63	
Model_3	-128.69	4.99		Model_3	289.07	2.56	
LowRMSD_9	-118.38	1.61		LowRMSD_1147	672.68	1.68	
MA2_9	117.43	0.00	R=0.18 n=80	MA2_9	117.43	0.00	R=0.30 n=1034
Model_1	-208.55	2.65		Model_1	122.78	2.51	
Model_2	-181.58	2.08		Model_2	147.57	2.68	
Model_3	-178.54	2.76		Model_3	150.85	3.24	
LowRMSD_26	-131.13	0.66		LowRMSD_72	189.34	0.66	
MA2_10	28.16	0.00	R=-0.03 n=59	MA2_10	28.16	0.00	R=-0.01 n=1160
Model_1	-64.76	2.56		Model_1	-66.95	2.72	
Model_2	-63.17	2.74		Model_2	-65.92	3.50	
Model_3	-54.20	5.65		Model_3	-52.63	3.67	
LowRMSD_56	113.60	1.91		LowRMSD_384	116.26	1.34	
MA2_11	146.47	0.00	R=0.61 n=82	MA2_11	146.47	0.00	R=0.33 n=1033
Model_1	-449.36	2.45		Model_1	77.38	1.58	
Model_2	-440.45	2.30		Model_2	84.92	2.64	
Model_3	-434.85	2.67		Model_3	86.60	2.39	
LowRMSD_5	-429.32	1.72		LowRMSD_12	106.53	1.11	

Figure 2

Comparison of AMBER12 all-atom energies (kcal/mol) of the target x-ray crystal structures with the Stage-1 (A) and Stage-2 models (B). Values referring to the X-ray structures are shown with gray background, followed by the lowest three energy models (Model-1, Model-2, and Model-3) and the model with the lowest global RMSD (LowRMSD_x). The suffix indicates the rank in the energy-sorted conformation pool. The RMSD values (Å) are based on measurements in the backbone (N, C α , C, and O) of CDR H3 and represent the global RMSD to the target H3. The Plots show the results of a correlation analysis (global RMSD against AMBER12 all-atom energies), “n” denotes the number of conformations left in the conformation pool after removal of models with bad energies or unresolvable clashes, “R” is the Pearson’s correlation coefficient of global RMSD (y-axis) and Amber12 all-atom energies (x-axis).

**Figure 3**

Comparison of the RMSDs of the automated protocols in the first and second antibody modeling assessment in the F_V region (All), the dimer framework (FR), Canonical Loops (Canonical), and CDR- H_3 (H3).

Almagro *et al.* in this issue) and are signs of a not yet fully fine-tuned and matured ccg2 modeling protocol.

Comparison to the first antibody modeling assessment (AMA-I)

The second assessment (AMA-II) provides an opportunity to measure the improvements made since the first assessment (AMA-I). The comparison is best done with the autoFv/ccg3 protocol since this protocol was also used in the first modeling assessment. As depicted in Figure 3, the current version of the autoFv protocol suggests enhancements in the important framework (FR), canonical loops (Canonical) measurements evidenced by the average RMSD values of 1.34 vs. 1.05 Å over the

entire F_V , 0.93 vs. 0.79 Å in the framework and 1.17 vs. 1.0 Å in the canonical loop categories. A similar degree of progress is seen with the ccg2 protocol. As discussed above, the differences observed in modeling the H_3 loop are not rooted in a systematic effort to improve modeling in this category but rather a result of refinements made in other areas of the binding site that result in better predictions of CDR- H_3 overall.

AMA-II Stage 2

Stage 2 focused on modeling of CDR- H_3 when given the coordinates of the remaining antibody dimer. Minor deviations in modeling the framework and canonical CDR loops will influence the architecture of the antigen binding surface and will in turn affect the modeling of CDR- H_3 . In this respect, Stage 2 can be viewed as a reduction of the modeling complexity by eliminating the uncertainty of modeling the remaining F_V region; in other words, Stage 2 is a CDR- H_3 modeling exercise under ideal conditions. Using the dedicated CDR- H_3 loop database and techniques described in the previous section (still under development at the time of AMA-II), we generated five models, CCG1–CCG5, selected on the basis of three different scoring regimes. The CCG1 and CCG2 models were the two lowest forcefield energy structures; CCG3 and CCG4 models represent the top two ranking models based on specialized H_3 loop scores; CCG5 models were selected manually (based on personal experience, atom clashes and protein geometry) from the pool of the top ten candidates of each group.

Table III presents the RMSD comparison (local and global) of each of the generated models with the experimental crystal structure. The number of CDR- H_3 conformations generated was ~800 conformations for CDR- H_3

Table III

RMSD Values in Å (Local and Global) Between Three Distinct Selection Methods (See Text) Employed in Stage 2 and the Target X-Ray Crystal Coordinates; “H3” Gives the Length of the CDR- H_3 Loop; “nConfs” Is the Total Number of Conformations Generated in the Respective Pool; “CCG1 and 2” Are the RMSDs of the Lowest AMBER12 All-Atom Energy Models, “CCG3 and 4” Are the RMSDs of the Models With the Highest Loop Scores and CCG5 Is Based on a Manual Selection, “Lowest” Is the Lowest RMSD Conformer in the Pool

Sample	H3	nConfs	AMBER12		Loop Score		Manual			AMBER12		Loop Score		Manual	
			CCG1	CCG2	CCG3	CCG4	CCG5	lowest	CCG1	CCG2	CCG3	CCG4	CCG5	lowest	
MA2-2	11	1452	2.31	1.82	2.17	1.63	2.23	1.24	2.95	2.51	3.03	4.47	2.87	1.53	
MA2-3	8	793	1.23	1.33	0.53	1.19	1.49	0.34	1.67	2.31	0.85	1.72	2.33	0.84	
MA2-4	8	803	0.62	0.91	0.63	0.72	0.79	0.54	0.87	1.29	0.92	1.06	1.21	0.68	
MA2-5	8	800	0.66	1.61	0.79	1.83	0.9	0.59	1.15	2.34	0.92	2.87	1.44	0.93	
MA2-6	14	303	2.66	1.75	2.42	3.63	1.8	1.61	3.64	3.09	4.26	6.37	4.15	1.66	
MA2-7	8	803	0.88	0.66	1.86	1.08	0.89	0.42	1.4	0.99	3.37	2.14	1.09	0.87	
MA2-8	11	1452	2.86	2.61	1.95	2.8	2.14	1.17	6.36	3.62	3.46	4.93	3.75	1.68	
MA2-9	10	1294	1.35	0.95	2.44	2.11	0.78	0.47	2.51	2.67	3.31	3.71	1.56	0.66	
MA2-10	11	1452	2.03	2.26	2.02	1.97	2.61	0.78	2.72	3.6	2.8	2.93	3.11	1.34	
MA2-11	10	1293	0.58	1.95	2.5	0.7	2.7	0.56	1.58	2.64	2.88	1.83	4.38	1.11	
average			1.52	1.59	1.73	1.77	1.63	0.77	2.49	2.51	2.58	3.2	2.59	1.13	
stdev			0.83	0.59	0.74	0.88	0.73	0.4	1.54	0.82	1.17	1.58	1.18	0.37	
						local RMSD				global RMSD					

All RMSD measurements are based the backbone atoms N, C α , C, and O.

lengths of 8, ~1,200 for lengths 10 and 11, and ~300 conformations for the 14 residue loop of target MA2-6. Conformation coverage was good with 7 of 10 targets having a conformation generated at <1.0 Å local RMSD to the experimental structure, two between 1.0 and 1.25 Å and target MA2-6 at 1.6 Å. It appears the conformation space of target MA2-6 was not adequately sampled suggesting that the iteration limits of LowModeMD should increase as a function of loop length. The energy-based selections, CCG1 and CCG2, were better than the scoring function selections, CCG3 and CCG4; this is true on average and in the individual cases with the exception of target MA2-3. In three cases, the manual selection, CCG5, outperformed the CCG1 energy based selections; however, in two out of the three the RMSD to the crystal structure was >2.0 Å and these cases (MA2-6 and MA2-8) also had no generated conformations under 1.0 Å RMSD. This suggests that in the absence of a close-to-experiment conformation, aesthetic qualities of the loop conformation, rather than energy, may lead to closer predictions. The average CCG1 RMSD to experiment on all 10 targets was 1.5 local and 2.5 Å global. Leaving out the targets for which no conformation <1 Å RMSD (local) was generated (MA2-2, MA2-6, and MA2-8) the averages were 1.1 Å RMSD local and 1.7 Å global. This supports the hypothesis that if there is a close conformation to experiment in the pool then the energy based selection can do well.

Some discussion is warranted surrounding the issues related to energy based model selection. First, atomic forcefields are only approximations and, as such, they have parameterization errors; for example, the bonded parameters, van der Waals and electrostatics/solvation treatment. The lowest energy structures can be quite different from each other. For example, Figure 4 shows the three lowest energy conformers of CDR-H₃ sampled near the x-ray structure of MA2-8; there is a large conformational spread but with a relatively small energy range of 2.6 kcal/mol (well within forcefield parameterization error). This presents a problem when ranking conformations and has been observed in small molecule docking calculations.²⁵ Second, forcefields are quite sensitive to small perturbations of the coordinates; experimental coordinates (which serve as the RMSD reference) are typically high in energy when compared to refined modeled structures, which, in some sense, is a built-in bias towards higher RMSD values. Given the premise of Stage 2 modeling with $>90\%$ of the heavy atom coordinates known, the energy differences among the models are caused exclusively by the conformational variability in CDR H3. Comparing the top energy models with the native x-ray structures as shown in Figure 2(B) reveals a systematic failure in discriminating the low RMSD conformer from the conformation pool. The results of a correlation analysis of AMBER12 all-atom energy against global RMSD do not support the notion of a linear

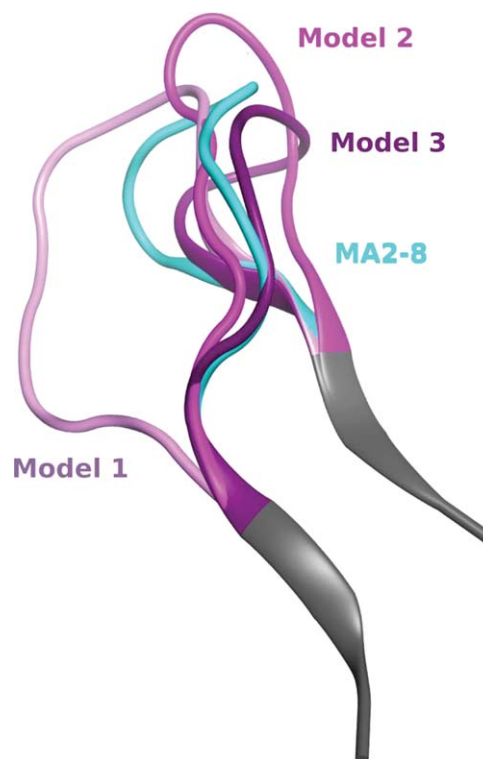


Figure 4

Conformational range in CDR-H₃ within a small energy window of 2.6 kcal/mol. The energies of Models 1, 2, and 3 are in order from lowest to highest and the deviation of the models from the x-ray coordinates were 6.37, 3.63, and 2.56 Å global RMSD.

relationship between the two variables [Fig. 2(B)]. This suggests that other—maybe more CDR H3-specific—terms are required to increase the loop prediction accuracy.

In retrospect, keeping all of the non-H₃ heavy atoms fixed in our protocol may have contributed to the higher RMSD values; for example, the join residues in the *F_V* (being fixed) may have contributed to increased strain energy that could have been relaxed if a tethering protocol was used. Such strain would confound the ranking and perhaps the Ramachandran component of the scoring function used in CCG3 and CCG4 (and implicitly in CCG5) may help in such cases if some means were developed to combine it with the energy value.

CONCLUSION

We have described a method for automatic modeling of the *F_V* region of an immunoglobulin. It is based upon the use of a precompiled antibody x-ray structure database. This database is used as a source of framework and hypervariable region structural templates, which are

grafted together. Multiple initial models from multiple templates are constructed. A final consensus model is built by selecting the templates used in one or more of the initial selected by considering forcefield binding energies between F_V components along with cluster analyses of the CDRs of the initial models. In addition, we have described a method for modeling the CDR- H_3 loop in the context of known coordinates for the remainder of an F_V . The method is based on using a computationally enriched database of H_3 conformations as a source for loop grafting and coordinate refinement.

The described methods were used in the Second Antibody Modeling Assessment (see Almagro, this issue) on commonly available desktop computer hardware; ~30 min of CPU time were required to build each model without user intervention (except for one manually modeled rabbit F_V). Stage 1 of AMA-II involved full F_V modeling and our models resulted in an average of 1.05 Å RMSD to the (previously withheld) x-ray crystal coordinates. Our previously established protocol (used in the First Antibody Modeling Assessment) proved the best when compared with two experimental new protocols still under development. Stage 2 of AMA-II focused on CDR- H_3 modeling with known remaining F_V coordinates. Our specialized CDR- H_3 modeling method (which was still under development at the time of the assessment) produced conformation predictions that were on average 1.5 Å local RMSD to the x-ray coordinates and 2.4 Å global RMSD. Leaving out a few problem structures resulted in a remaining average of 1.1 Å local RMSD local and 1.7 Å global.

The accuracy of CDR- H_3 predictions (whether under Stages 1 or 2 conditions) is lower than that of the other CDR loops and the remainder of the F_V region. Our methods are capable of generating diverse numbers of CDR- H_3 conformations and in the majority of cases considered in AMA-II there was a generated conformation <1.0 Å RMSD from the x-structure. The challenge appears to be the accurate ranking of the conformations. We present some evidence that the forcefield energies are a viable method for ranking keeping in mind that parameterization and coordinate perturbations can have a significant impact on such ranking.

REFERENCES

1. Maggon K. Monoclonal antibody “gold rush”. *Curr Med Chem* 2007;14:1978–1987.
2. Chothia C, Lesk AM. Canonical structures for the hypervariable regions of immunoglobulins. *J Mol Biol* 1987;196:901–917.
3. Tramontano A, Chothia C, Lesk AM. Structural determinants of the conformations of medium-sized loops in proteins. *Proteins* 1989;6:382–394.
4. Martin AC, Thornton JM. Structural families in loops of homologous proteins: automatic classification, modelling and application to antibodies. *J Mol Biol* 1996;263:800–815.

5. Almagro JC, Beavers MP, Hernandez-Guzman F, Shaulsky J, Butenhof K, Maier J, Kelly K, Labute P, Thorsteinson N, Teplyakov A, Luo J, Sweet R, Gilliland G. Antibody Modeling Assessment, *Proteins* 2011;79:3050–3066.
6. Molecular Operating Environment (MOE) software; Chemical Computing Group Inc. <http://www.chemcomp.com> (last accessed April 2014).
7. Bernstein FC, Koetzle TF, Williams GJB, Meyer EF, Jr, Brice MD, Rogers JR, Kennard O, Shimanouchi T, Tasumi M. The protein data bank: a computer-based archival file for macromolecular structures. *J Mol Biol* 1977;112:535–542.
8. Berman HM, Bhat TN, Bourne PE, Feng Z, Gilliland G, Weissig H, Westbrook J. The Protein Data Bank and the challenge of structural genomics. *Nat Struct Biol* 2000;7 Suppl:957–959.
9. 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.
10. Allcorn LC, Martin ACR. SACS – self-maintaining database of antibody crystal structure information. *Bioinformatics* 2002;18:175–181; <http://antibody.bath.ac.uk>.
11. Giudicelli V, Chaume D, Bodmer J, Müller W, Busin C, Marsh S, Bontrop R, Marc L, Malik A, Lefranc MP. IMGT, the international ImMunoGeneTics database. *Nucleic Acids Res* 1997;25:206–211.
12. Lefranc M-P, Giudicelli V, Ginestoux C, Jabado-Michaloud J, Folch G, Bellahcene F, Wu Y, Gemrot E, Brochet X, Lane J, Regnier L, Ehrenmann F, Lefranc G, Duroux P. IMGT, the international ImMunoGeneTics information system. *Nucleic Acids Res* 2009;37(Database issue):D1006–D1012.
13. Hubbard TJ, Murzin AG, Brenner SE, Chothia C. SCOP: a structural classification of proteins database. *Nucleic Acids Res* 1997;25:236–239.
14. Andreeva A, Howorth D, Brenner SE, Hubbard TJP, Chothia C, Murzin AG. SCOP database in 2004: refinements integrate structure and sequence family data. *Nucleic Acids Res* 2004;32(Database issue):D226–D229.
15. Henikoff S, Henikoff JG. Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci USA* 1992;89:10915–10919.
16. Sonnhammer ELL, Hollich V., Scoredist: a simple and robust protein sequence distance estimator, *BMC Bioinformatics* 2005;6:108.
17. North B, Lehmann A, Dunbrack RL, Jr. A New Clustering of Antibody CDR Loop Conformations. *J Mol Biol* 2011;406:228–256.
18. Chailyan A, Marcatili P, Tramontano A. The association of heavy and light chain variable domains in antibodies: implications for antigen specificity. *FEBS* 2011;278:2848–2866.
19. Teplyakov A, Obmolova G, Malia T, Gilliland G. Antigen recognition by antibody C836 through adjustment of VL/VH packing. *Acta Cryst* 2011;F67:1165–1167.
20. Abhinandan KR, Martin ACR. Analysis and improvements to Kabat and structurally correct numbering of antibody variable domains. *Mol Immunol* 2008;45:3832–3839.
21. Labute P. LowModeMD – implicit low-mode velocity filtering applied to conformational search of macrocycles and protein loops. *J Chem Inf Model* 2010;50:792–800.
22. Case DA, et al. AMBER 12. San Francisco: University of California; 2012.
23. Labute P. The generalized Born/volume integral implicit solvent model: estimation of the free energy of hydration using London dispersion instead of atomic surface area. *J Comput Chem* 2008;29:1693–1698.
24. Labute, P. Protonate3D: assignment of ionization states and hydrogen coordinates to macromolecular structures. *Proteins* 2008;75:187–205.
25. Corbeil C, Williams C, Labute P. Variability in docking success rates due to dataset preparation. *JCAMD* 2012;26:775–786