OXFORD

Structural bioinformatics

Sphinx: merging knowledge-based and *ab initio* approaches to improve protein loop prediction

Claire Marks[1], Jaroslaw Nowak[1], Stefan Klostermann[2], Guy Georges[3], James Dunbar[3], Jiye Shi[4], Sebastian Kelm[4] and Charlotte M. Deane[1,*]

[1]Department of Statistics, University of Oxford, Oxford, UK, [2]Pharma Research and Early Development, Informatics and [3]Pharma Research and Early Development, Large Molecule Research, Roche Innovation Center Munich, Penzberg, DE, Germany and [4]Department of Informatics, UCB Pharma, Slough, UK

*To whom correspondence should be addressed.
Associate Editor: Alfonso Valencia

Abstract

Motivation: Loops are often vital for protein function, however, their irregular structures make them difficult to model accurately. Current loop modelling algorithms can mostly be divided into two categories: knowledge-based, where databases of fragments are searched to find suitable conformations and *ab initio*, where conformations are generated computationally. Existing knowledge-based methods only use fragments that are the same length as the target, even though loops of slightly different lengths may adopt similar conformations. Here, we present a novel method, Sphinx, which combines *ab initio* techniques with the potential extra structural information contained within loops of a different length to improve structure prediction.

Results: We show that Sphinx is able to generate high-accuracy predictions and decoy sets enriched with near-native loop conformations, performing better than the *ab initio* algorithm on which it is based. In addition, it is able to provide predictions for every target, unlike some knowledge-based methods. Sphinx can be used successfully for the difficult problem of antibody H3 prediction, outperforming RosettaAntibody, one of the leading H3-specific *ab initio* methods, both in accuracy and speed.

Availability and Implementation: Sphinx is available at http://opig.stats.ox.ac.uk/webapps/sphinx.

Contact: deane@stats.ox.ac.uk

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

Loops, the irregular regions of a protein that connect secondary structure elements, often play a vital functional role. Structural information is critical to fully understanding a protein's properties – however, obtaining structures experimentally is time-consuming and difficult, especially for loop regions, which can be missing from experimentally solved structures (Petoukhov *et al.*, 2002). Due to the variability of loop structures and sequences between homologues, predicting loop conformations remains challenging – loops are usually the least accurate regions of a protein model (Baker and Sali, 2001; Gront *et al.*, 2012).

Traditionally, methods for predicting protein loop structures are divided into two categories, knowledge-based or *ab initio*, depending on how they generate possible conformations (decoys). Knowledge-based methods rely upon databases of previously observed protein structure fragments. Loop structures are selected according to certain criteria such as fragment length (i.e. they must be the same length as the target loop), fragment-target sequence similarity and how closely the anchor geometry of the fragment matches that of the target loop. Methods of this type are extremely fast and can be very accurate when the structure of the target loop is similar to one previously observed. However, there is not currently

enough structural data to cover the conformational space, especially for long loops (Fernandez-Fuentes and Fiser, 2006). When a similar loop structure has not been observed previously, knowledge-based methods either give poor predictions or fail to return a prediction at all. Examples of this type of algorithm include FREAD (Choi and Deane, 2010; Deane and Blundell, 2001), SuperLooper (Hildebrand *et al.*, 2009), LoopWeaver (Holtby *et al.*, 2013), LIP (Michalsky *et al.*, 2003) and LoopIng (Messih *et al.*, 2015).

Ab initio methods do not rely on previously observed structures; instead, decoys are produced computationally. *Ab initio* methods work by exploring the possible conformational space, for example by randomly sampling the ϕ and ψ dihedral angles of the loop. Depending on how the loops are built, the continuity of the protein backbone may need to be enforced through the implementation of a closure algorithm, such as cyclic coordinate descent (CCD) (Canutescu and Dunbrack, 2003), random tweak (Shenkin *et al.*, 1987) or KIC (Mandell *et al.*, 2009). Like knowledge-based methods, *ab initio* algorithms also have their limitations: they are computationally expensive, since many decoys must be generated to sample the conformational space sufficiently; and their prediction accuracy decreases with loop length (as the number of degrees of freedom increases). *Ab initio* methods include PLOP (Jacobson *et al.*, 2004), MODELLER (Fiser *et al.*, 2000), Loopy (Xiang *et al.*, 2002), LoopBuilder (Soto *et al.*, 2007), LEAP (Liang *et al.*, 2014), and the loop modelling routine within Rosetta (Stein and Kortemme, 2013).

The idea of a hybrid loop modelling algorithm, combining knowledge-based and *ab initio* approaches, has been explored before. CODA (Deane and Blundell, 2001) generates decoys using a knowledge-based method and an *ab initio* method separately, then combines the two decoy sets and makes a consensus prediction. Fasnacht *et al.* (2014), Martin *et al.* (1989), and Whitelegg and Rees (2000)) have used similar approaches and applied it to modelling the complementarity determining regions (CDRs) of antibodies – initial conformations are selected from a database of structures, and the middle section is then remodelled using *ab initio* techniques. In all of these examples, however, the *ab initio* and knowledge-based sections of the algorithms are kept distinct from one another. An alternative approach using Rosetta is described by Rohl *et al.*, 2004 – this used a Monte Carlo-based fragment assembly method, in conjunction with a minimization protocol. The fragments used by this algorithm are very short (three residues for loops under 15 residues in length) and therefore any structural similarity is only considered for short segments of the target loop, not all of it.

Recent research into the canonical conformations of antibody CDRs has shown that loops of different lengths can adopt similar structures (Nowak *et al.*, 2016), indicating that loops with a different number of residues could still provide useful information when modelling a particular target loop. However, no existing methods are able to use this knowledge, since they either do not use previously known structures, must only use length-matched fragments, or build conformations by assembling very short fragments. Some algorithms, such as Frag'r'us (Bonet *et al.*, 2014), allow the user to remodel loop structures using fragments of a different length to the original loop, however, the fragment must still be of the same length as the desired target loop. Here, we introduce a new loop modelling method, Sphinx, that can use this extra source of structural information, by integrating aspects of both knowledge-based and *ab initio* approaches in a novel way. Sphinx begins the loop modelling process by identifying potential fragments that are slightly shorter than the loop to be predicted. By combining the structural information

from such fragments with *ab initio* techniques, the length is adjusted and decoys of the correct length are generated. In this way, we are able to use the additional information present in different length loops. Sphinx is not limited to predicting what has already been observed, like a traditional knowledge-based method, nor does it ignore the available structural information, as would be the case using a purely *ab initio* approach. Our algorithm gives a prediction in every case (unlike some knowledge-based methods) and generates decoy sets that are enriched with near-native conformations. This leads to high-accuracy predictions, demonstrated by its ability to outperform existing software in the challenging problem of in antibody CDR H3 prediction.

2 Methods

The Sphinx algorithm is a combination of FREAD (Choi and Deane, 2010; Deane and Blundell, 2001), a knowledge-based method and our own *ab initio* algorithm. The inputs to Sphinx are a protein structure or model (in PDB format) and the location and sequence of the loop to be modelled. The algorithm includes four main steps: database search, loop building, loop closure and decoy selection. A flowchart of the algorithm is given in Figure 1.
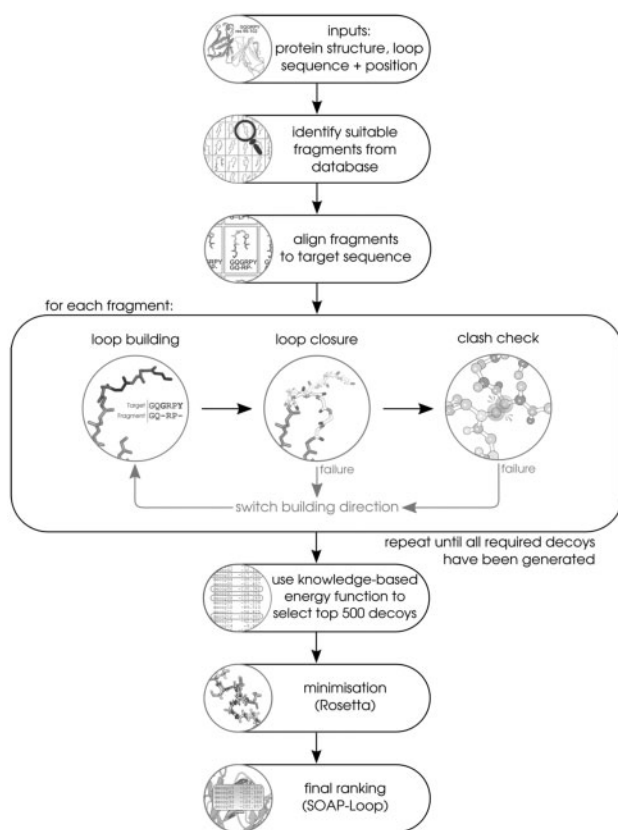
2.1 Database search

The first step of the Sphinx algorithm is the search of a fragment database, for fragments that are a different length to the target, but that may share the same conformation. To create this database, the loop structures of all PDB entries with non-identical sequences [according to PISCES (Wang and Dunbrack, 2003), using the structure with the best resolution for each sequence], plus three residues on each side, were extracted and split into all possible fragments of 3–30 residues in length. Loop regions were located using DSSP (Joosten *et al.*, 2011) and were defined as regions connecting secondary structures of at least three residues in length. This led to a database containing the loop structures of 65 108 protein chains (created October 2015).

Fragments from the database are selected as potential starting points for prediction depending on three criteria:

1. length – the fragment must be shorter than the target loop, since preliminary tests indicated that fragments longer than the target were not able to produce decoys of the same quality. This may be because shorter loops are represented more in the database.
2. the anchor geometry of the fragment (the two residues on each side of the loop) must match that of the target. As is the case in the FREAD algorithm (Choi and Deane, 2010; Deane and Blundell, 2001), the distances between the Cα atoms of the anchor residues (the two residues on each side of the loop) are used to select fragments with similar anchor geometry.
3. sequence similarity – the sequences of fragments that meet the first two criteria are aligned to the target sequence using the Needleman–Wunsch algorithm (Needleman and Wunsch, 1970). These alignments are scored using an environment-specific substitution (ESS) score (a score which takes into account the dihedral angles of the fragment residues) (Deane and Blundell, 2001; Hill *et al.*, 2011), as implemented in FREAD.

The best fragments according to their ESS scores are then passed to the loop building stage of the algorithm, where their length (n) will be adjusted using *ab initio* techniques. Based on exploratory testing of Sphinx (using the general loop set described in Section 2.6), the

Fig. 1. Flowchart describing the Sphinx algorithm. As input, the algorithm requires the protein structure on which the loop is to be modelled, and the loop sequence and location. The algorithm begins by searching through a database of loop fragments. Suitable fragments, that are shorter than the target loop, are used to help build a set of 100 decoys. The number of fragments used depends on the length (*n*) of the target – $12.5(n-2)$ for even lengths and $12.5(n-1)$ for odd lengths. For any residues aligned between the target and the fragment, structural information is copied across. Any residues not aligned to fragment residues are built *ab initio*. Each loop structure undergoes a closure step and a clash check – if closure fails or there are clashes, the next decoy is built from the alternate anchor residue ('switching'). Once all decoys have been generated, a loop-specific energy function is used to select the top 500 decoys, which subsequently undergo minimization (using Rosetta) and final ranking (with SOAP-Loop). A larger, colour version of this figure is provided in the Supplementary Information

number and length of fragments to be used depend on the following rules:

$$\text{minimum target length} = \begin{cases} n-3 & \text{if } n \leq 12 \\ n-4 & \text{if } n > 12 \end{cases}$$

$$\text{number of fragments} = \begin{cases} 12.5(n-2) & \text{if } n \text{ is even} \\ 12.5(n-1) & \text{if } n \text{ is odd} \end{cases}$$

Throughout this work, we ignore fragments that come from sequence-identical chains to the target protein, as in a real modelling situation these structures would not be available.

2.2 Loop building
The structure and sequence alignment information of each selected fragment is then used to build a set of 100 complete loop decoys of the correct length. Each decoy is built, one residue at a time, onto one of the anchor residues, using the algorithm described by Parsons *et al.*, 2005.

For target residues that are paired with a fragment residue in the sequence alignment, the structural data (bond lengths, bond angles and dihedral angles) used to build the residue are calculated from the corresponding fragment residue. For target residues that are not aligned to a fragment residue, the structural data are randomly selected from a set of purpose-built distributions. Bond lengths and angles are drawn at random from the Gaussian distributions (parameters given by Engh and Huber, 1991), and dihedral (ϕ/ψ) angles are selected from a set of residue-specific Ramachandran distributions, built using the same set of loop structures that were used to create the fragment database (see previous section).

2.3 Loop closure
Since the termini of the loop must be connected to the anchor residues to form a continuous backbone, the structure generated must be 'closed'; for example, if the loop was built from the N-anchor (the anchor nearest to the N terminus of the protein), then the structure must be altered so that the other end of the loop attaches to the C-anchor. The loop closure method used by Sphinx is CCD (Canutescu and Dunbrack, 2003), which works by iteratively changing the ϕ/ψ angles of the loop, gradually minimizing the distance between the free end of the loop and the anchor.

For Sphinx, we have written a modified version of the CCD algorithm (as suggested by Canutescu and Dunbrack, 2003), where constraints are put on the angle changes so that only the allowed regions of the Ramachandran plot are occupied (see Supplementary Information for details). Thus, all decoys generated by Sphinx will be physically feasible.

It may be preferential for a loop to be built from one particular anchor residue instead of the other. Therefore, if loop closure fails, or the resulting loop structure causes clashes between atoms in the structure, then the loop decoy is discarded and the next loop decoy is built from the alternate anchor residue. This approach, known as 'switching' (Chys and Chacon, 2013) allows the preferred direction of loop building to dominate.

Once 100 decoys have been made from a fragment (or the number of failures reaches a cut-off value of 3000), the algorithm moves onto the next fragment until all selected fragments have been used.

2.4 Decoy selection
Once a complete set of decoys has been generated (100 per fragment), the number is reduced to 500 using our own loop-specific, knowledge-based energy function. This function is based on the RAPDF (Samudrala and Moult, 1998) and compares the distances between atoms of real structures to those in a decoy. Our version considers only backbone and C_β atoms, meaning that the time-consuming process of side-chain addition does not have to be carried out. The data for this function were calculated from a non-redundant set of structures (with a maximum sequence identity of 40% and a resolution cut-off of 2 Å), considering only those pairwise distances involving loop atoms. The function uses six bins for the data (one bin for distances between 0 and 3 Å, followed by five 1 Å bins up to a maximum of 8 Å).

Once the top 500 decoys have been selected, side-chains are added using SCWRL4 (Krivov *et al.*, 2009), they are minimized using the loop refinement protocol within Rosetta (see Supplementary Information) and subsequently ranked using the SOAP-Loop potential

in MODELLER (Dong *et al.*, 2013), a statistical potential developed specifically for loop modelling. SOAP-Loop is fast and gave the best results when tested against other ranking methods (see Supplementary Information).

2.5 Sphinx-H3

The H3-specific version of Sphinx is different to the general version in three ways; the database of fragments searched, the dihedral data used and the number of fragments from the database to be used in loop building.

The H3-specific database was created by extracting the structures of all the H3 loops that were present in the Structural Antibody Database (SAbDab) (Dunbar *et al.*, 2014), which were then split into all their possible fragments (in the same way as the general fragment database). We define the H3 loop region according to Chothia and Lesk (1987). In the Chothia numbering scheme, the H3 is found on the heavy chain between residues 95 and 102 inclusive. At the time, the database was made (April 2015), this led to a database containing all possible fragments of 3043 H3 structures, from 1848 different PDB entries.

As there are currently not enough H3 structures available to produce residue-specific Ramachandran distributions, we constructed the distributions using a novel resampling method that combines general protein loop and H3 data (see Supplementary Information for details). The resampling method makes the Ramachandran distributions less sparsely populated while maintaining the unique nature of the H3 data.

To verify the ability of Sphinx to predict H3 loop structures, we first ran it on a target set of H3 loops (see Supplementary Information for details). Following this test, we allowed Sphinx to use twice as many fragments in loop building than were used in the case of general protein loops.

2.6 Datasets

Three datasets are used in this paper: a general protein loop set and two sets of antibody H3 loops. From the general protein loops that were extracted to make the fragment database, a series of loops were selected with no more than 40% sequence identity, resolution below 2 Å, no B-factor above 30, and no hydrogen bonds to the surrounding crystal. From these loops, a dataset of 36 targets was selected, with lengths ranging from 6 to 18 residues. The first set of antibody H3 targets was taken from Sivasubramanian *et al.* (2009). The target 2ai0 was removed from the original set since the structure is incomplete and was previously incorrectly designated as a 6-residue loop. Loops that interact with other proteins in the crystal were also removed, leading to a set of 39 targets which range in length from 4 to 22 residues. The second set of H3 targets are the ten targets from the second round of the Antibody Modelling Assessment II. Complete lists of targets are given in the Supporting Information.

3 Results

Sphinx begins its prediction by searching for fragments of a slightly different length but possibly a similar conformation to the target loop, as indicated by sequence similarity and anchor geometry. By using this knowledge-based approach in tandem with *ab initio* techniques, these fragments can be adjusted to be the correct length. In all cases, fragments from protein chains that are identical in sequence to the target chain, and fragments of the same length as the target that could be found using a normal knowledge-based method, are ignored. All RMSDs are calculated after first superimposing the

rest of the protein structure, or in the case of the antibody models, the framework residues of the heavy chain [as defined in the second Antibody Modelling Assessment (Teplyakov *et al.*, 2014)]. All reported RMSDs use all backbone atoms of the loop, except for the Anitbody Modelling Assessment II target set, which uses only the carbonyl carbon and oxygen atoms of residues 95–100*x* (using the Chothia numbering scheme).

3.1 Comparison to knowledge-based and *ab initio* methods

Figure 2 shows the accuracy of the methods on general loops. The methods shown are Sphinx and the algorithms on which it is based: the knowledge-based method FREAD (Choi and Deane, 2010; Deane and Blundell, 2001) and our own *ab initio* method. FREAD is a knowledge-based method, which searches a database for suitable fragments of the same length as the target. We ran FREAD using the same fragment database as used by Sphinx (Section 2), which on average returned seven decoys per target. The *ab initio* method can be thought of as running Sphinx without any fragment information – i.e. all bond lengths, angles and dihedral angles are chosen randomly from the relevant distributions. To ensure a fair comparison, the *ab initio* method was run so as to generate the same number of decoys for each target as Sphinx (around 11 000 on
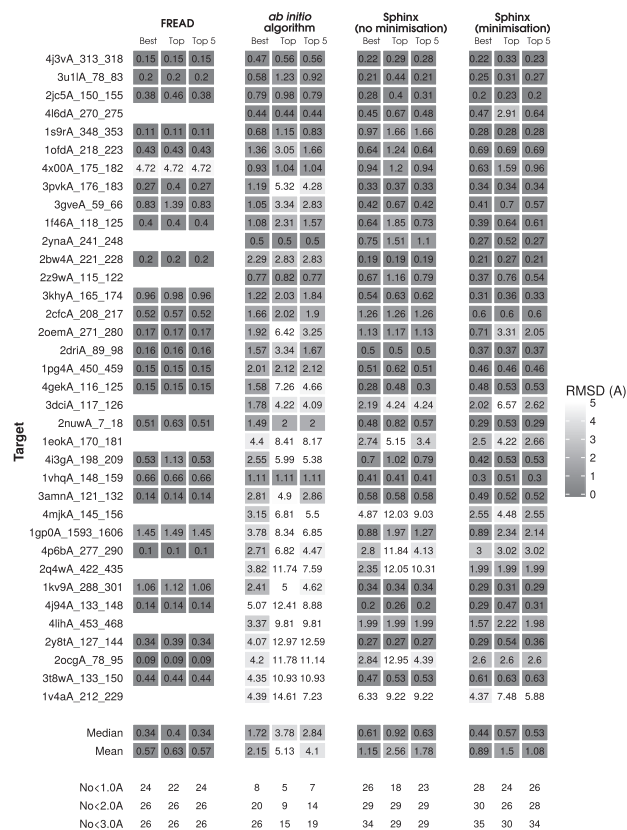
Target	FREAD Best	FREAD Top	FREAD Top 5	ab initio algorithm Best	ab initio algorithm Top	ab initio algorithm Top 5	Sphinx (no minimisation) Best	Sphinx (no minimisation) Top	Sphinx (no minimisation) Top 5	Sphinx (minimisation) Best	Sphinx (minimisation) Top	Sphinx (minimisation) Top 5
4j3vA_313_318	0.15	0.15	0.15	0.47	0.56	0.56	0.22	0.29	0.28	0.22	0.33	0.23
3u1lA_78_83	0.2	0.2	0.2	0.58	1.23	0.92	0.21	0.44	0.21	0.25	0.31	0.27
2jc5A_150_155	0.38	0.46	0.38	0.79	0.98	0.79	0.28	0.4	0.31	0.2	0.23	0.2
4l6dA_270_275				0.44	0.44	0.44	0.45	0.67	0.48	0.47	2.91	0.64
1s9rA_348_353	0.11	0.11	0.11	0.68	1.15	0.83	0.97	1.66	1.66	0.28	0.28	0.28
1ofdA_218_223	0.43	0.43	0.43	1.36	3.05	1.66	0.64	1.24	0.64	0.69	0.69	0.69
4x00A_175_182	4.72	4.72	4.72	0.93	1.04	1.04	0.94	1.2	0.94	0.63	1.59	0.96
3pvkA_176_183	0.27	0.4	0.27	1.19	5.32	4.28	0.33	0.37	0.33	0.34	0.34	0.34
3gveA_59_66	0.83	1.39	0.83	1.05	3.34	2.83	0.42	0.67	0.42	0.41	0.7	0.57
1f46A_118_125	0.4	0.4	0.4	1.08	2.31	1.57	0.64	1.85	0.73	0.39	0.64	0.61
2ynaA_241_248				0.5	0.5	0.5	0.75	1.51	1.1	0.27	0.52	0.27
2bw4A_221_228	0.2	0.2	0.2	2.29	2.83	2.83	0.19	0.19	0.19	0.21	0.27	0.21
2z9wA_115_122				0.77	0.82	0.77	0.67	1.16	0.79	0.37	0.76	0.54
3khyA_165_174	0.96	0.98	0.96	1.22	2.03	1.84	0.54	0.63	0.62	0.31	0.36	0.33
2cfcA_208_217	0.52	0.57	0.52	1.66	2.02	1.9	1.26	1.26	1.26	0.6	0.6	0.6
2oemA_271_280	0.17	0.17	0.17	1.92	6.42	3.25	1.13	1.17	1.13	0.71	3.31	2.05
2drlA_89_98	0.16	0.16	0.16	1.57	3.34	1.67	0.5	0.5	0.5	0.37	0.37	0.37
1pg4A_450_459	0.15	0.15	0.15	2.01	2.12	2.12	0.51	0.62	0.51	0.46	0.46	0.46
4gekA_116_125	0.15	0.15	0.15	1.58	7.26	4.66	0.28	0.48	0.3	0.48	0.53	0.53
3dclA_117_126				1.78	4.22	4.09	2.19	4.24	4.24	2.02	6.57	2.62
2nuwA_7_18	0.51	0.63	0.51	1.49	2	2	0.48	0.82	0.57	0.29	0.53	0.29
1eokA_170_181				4.4	8.41	8.17	2.74	5.15	3.4	2.5	4.22	2.66
4l3gA_198_209	0.53	1.13	0.53	2.55	5.99	5.38	0.7	1.02	0.79	0.42	0.53	0.53
1vhqA_148_159	0.66	0.66	0.66	1.11	1.11	1.11	0.41	0.41	0.41	0.3	0.51	0.3
3amnA_121_132	0.14	0.14	0.14	2.81	4.9	2.86	0.58	0.58	0.58	0.49	0.52	0.52
4mjkA_145_156				3.15	6.81	5.5	4.87	12.03	9.03	2.55	4.48	2.55
1gp0A_1593_1606	1.45	1.49	1.45	3.78	8.34	6.85	0.88	1.97	1.27	0.89	2.34	2.14
4p6bA_277_290	0.1	0.1	0.1	2.71	6.82	4.47	2.8	11.84	4.13	3	3.02	3.02
2q4wA_422_435				3.82	11.74	7.59	2.35	12.05	10.31	1.99	1.99	1.99
1kv9A_288_301	1.06	1.12	1.06	2.41	5	4.62	0.34	0.34	0.34	0.29	0.31	0.29
4j94A_133_148	0.14	0.14	0.14	5.07	12.41	8.88	0.2	0.26	0.2	0.29	0.47	0.31
4llhA_453_468				3.37	9.81	9.81	1.99	1.99	1.99	1.57	2.22	1.98
2y8tA_127_144	0.34	0.39	0.34	4.07	12.97	12.59	0.27	0.27	0.27	0.29	0.54	0.36
2ocgA_78_95	0.09	0.09	0.09	4.2	11.78	11.14	2.84	12.95	4.39	2.6	2.6	2.6
3t8wA_133_150	0.44	0.44	0.44	4.35	10.93	10.93	0.47	0.53	0.53	0.61	0.63	0.63
1v4aA_212_229				4.39	14.61	7.23	6.33	9.22	9.22	4.37	7.48	5.88
Median	0.34	0.4	0.34	1.72	3.78	2.84	0.61	0.92	0.63	0.44	0.57	0.53
Mean	0.57	0.63	0.57	2.15	5.13	4.1	1.15	2.56	1.78	0.89	1.5	1.08
No<1.0A	24	22	24	8	5	7	26	18	23	28	24	26
No<2.0A	26	26	26	20	9	14	29	29	29	30	26	28
No<3.0A	26	26	26	26	15	19	34	29	29	35	30	34

RMSD (A) 5 4 3 2 1 0

Fig. 2. Performance of Sphinx compared to its constituent knowledge-based and *ab initio* parts, for general protein loop types. Targets are listed in order of increasing length. 'Best' refers to the decoy in the set that is closest to the native structure, when considering the 500 decoys returned by the loop-specific energy function (for Sphinx and the *ab initio* method) or all the returned decoys (for FREAD). 'Top' means the decoy that is predicted to be the best, and 'Top 5' denotes the RMSD of the best decoy amongst the five top-ranked. A colour version of this figure can be found in the Supplementary Information
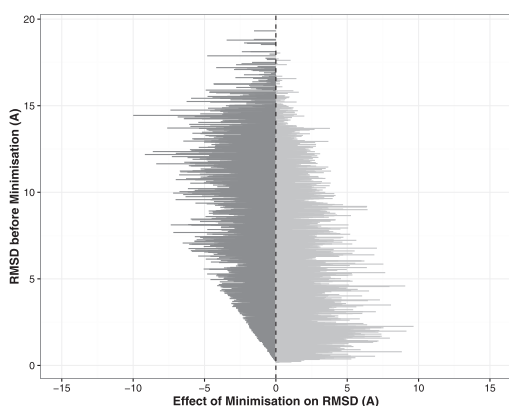
average), and its decoys were ranked in the same way. The results in Figure 2 consider only the 500 decoys selected by our loop specific energy function and the top decoys are selected with SOAP-Loop, either before or after minimization with Rosetta.

The effect of the Rosetta minimization protocol on decoy RMSD is shown in Figure 3. It can be seen that the refinement procedure does not consistently move the decoy structures closer to the native conformation. Overall, 59.1% of decoys have lower RMSDs after minimization, with an average improvement of 0.97 Å. Decoys that before minimization had a high RMSD are more likely to be improved by the minimization protocol; 60% of decoys that begin with an RMSD of over 1.5 Å are improved (with an average reduction in RMSD of 1.12 Å), while 55% of those with RMSDs below 1.5 Å are moved closer to the native structure (with an average RMSD improvement of 0.31 Å).

However, even though the minimization does not reliably improve RMSDs, overall the results are improved by its inclusion (Fig. 2). This appears to be due to SOAP-Loop's enhanced ability to select near-native decoys – the mean RMSD of the top-ranked decoy is reduced from 2.56 to 1.50 Å, and the mean RMSD of the best decoy in the top five decreased from 1.78to 1.08 Å. In addition, the number of sub-angstrom top-ranked predictions is increased from 18 to 24.

Figure 2 shows that FREAD is more accurate than Sphinx on average, but there are some targets (9) for which FREAD failed to find suitable fragments in the database. Sphinx, on the other hand, returned a prediction in every case. In addition, for 20 of the 36 targets Sphinx was able to generate a better decoy than FREAD, and for 28 targets, Sphinx produced a decoy with an RMSD of below 1 Å; four more than FREAD. Sphinx also gives more top-ranked decoys with sub-angstrom RMSDs, with 24 compared to 22 for FREAD.

Sphinx, by including information from fragments, is able to considerably outperform its base *ab initio* algorithm. When considering the top 500 decoys, Sphinx is able to generate more accurate conformations – the average RMSD of the best decoy produced was 0.89 Å, compared to 2.15 Å for the *ab initio* algorithm, and overall the lowest-RMSD decoy generated by Sphinx was better than that of the *ab initio* algorithm in 33 cases. The *ab initio* algorithm could

only produce a conformation with an RMSD below 1 Å for 8 targets; Sphinx gave 28.

The improved ability of Sphinx to produce near-native loop conformations is further illustrated by the RMSD distributions of the decoy sets produced (Fig. 4). There is a greater proportion of low-RMSD decoys in the decoy set generated by Sphinx, indicating that the inclusion of fragment information helps to direct the conformational search towards near-native structures.

The RMSDs of the top-ranked decoys (as predicted by SOAP-Loop) are significantly better when using Sphinx compared to the base *ab initio* method, with an average RMSD of 1.50 Å compared to 5.13 Å for the *ab initio* algorithm. The five top-ranked decoys contained 26 and 7 sub-angstrom conformations for Sphinx and the *ab initio* method, respectively. The improved accuracy is especially noticeable for longer loops – for the 16 targets of 12 or more residues, the *ab initio* method gave no sub-angstrom predictions, while Sphinx was able to produce 7.

To compare Sphinx's performance to other *ab initio* algorithms, we used the loop modelling algorithm within the MODELLER software (Fiser *et al.*, 2000) as well as RAPPER (DePristo *et al.*, 2003; de Bakker *et al.*, 2003) to model the loops in the general target set. In accordance with their respective papers, for MODELLER we generated 500 decoys for each target, and for RAPPER we generated 1000. As well as using each method's scoring system to rank the decoys, for RAPPER we also used the Sphinx ranking protocol (see Section 2.4). The results of this comparison are shown in Figure 5. Sphinx is better than the other two methods by every measure. The decoy sets produced by Sphinx are of a higher quality, with an average best RMSD of 0.89 Å compared to 4.34, 2.35 and 2.29 Å for MODELLER, RAPPER and RAPPER using our ranking system, respectively. Sub-ångström decoys were generated by Sphinx for most of the targets (28), while this was rare for the other methods. Considering the top-ranked decoy for each target, Sphinx's predictions are considerably more accurate, with an average RMSD of 1.50 Å compared to 6.03 Å, 6.34 Å and 4.49 Å for the other three algorithms. The use of Sphinx's ranking procedure on the decoys generated by RAPPER improved the latter's performance, for example, decreasing the average top RMSD by 1.85 Å, but Sphinx's
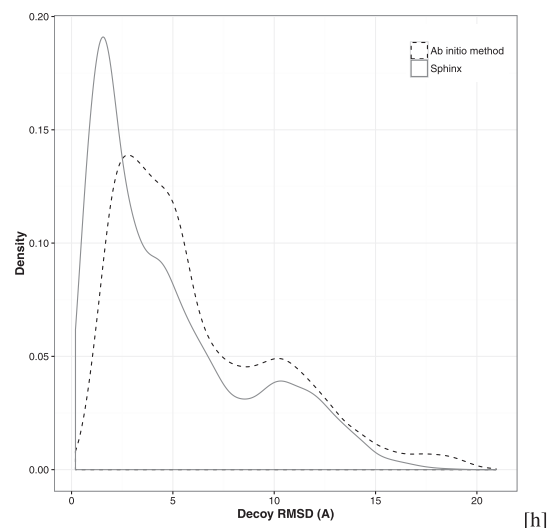

Fig. 3. The effect of Rosetta minimization on decoy RMSD, for the general loop target set. The change in RMSD of each decoy is represented as a horizontal line – the position of the line on the y axis indicates the RMSD before minimization, while the length of the line represents the change in RMSD. Decoys with improved RMSDs are on the left of the figure in dark grey while those whose RMSDs were made worse by the minimization step are on the right

[h]

Fig. 4. RMSD distributions for the 500-decoy sets produced by Sphinx and its base *ab initio* method. Decoy sets produced by Sphinx contain proportionally more near-native decoys. Data shown is for all targets in the general loop dataset

	Sphinx			MODELLER			RAPPER (1)			RAPPER (2)		
	Best	Top	Top 5	Best	Top	Top 5	Best	Top	Top 5	Best	Top	Top 5
Median	0.44	0.57	0.53	2.92	4.32	3.68	1.77	6.27	4.39	1.38	2.17	1.55
Mean	0.89	1.5	1.08	4.34	6.03	5.28	2.35	6.34	5.02	2.29	4.49	3.08
No < 1Å	28	24	26	5	4	4	6	0	0	9	5	6
No < 2Å	30	26	28	12	6	8	20	3	5	19	17	19
No < 3Å	35	30	34	18	11	16	28	9	14	25	19	22

Fig. 5. Comparison of Sphinx to other *ab initio* methods: the loop modelling algorithm of MODELLER and RAPPER. RAPPER results are shown using both the method's own ranking system (1) and the ranking procedure used by Sphinx (2)

A

	Sphinx			Rosetta		
	Best	Top	Top 5	Best	Top	Top 5
Median	0.71	2.31	1.06	0.75	3.25	1.48
Mean	0.94	2.5	1.52	1.07	3.38	2.35
No<1.0A	27	9	17	25	8	17
No<2.0A	37	17	30	35	13	20
No<3.0A	38	25	34	37	18	26

B

	Crystal Structures (with minimisation)			Model Structures (no minimisation)			Model Structures (with minimisation)		
	Best	Top	Top 5	Best	Top	Top 5	Best	Top	Top 5
Median	0.71	2.31	1.05	1.51	2.18	1.79	1.18	3.12	2.53
Mean	0.94	2.5	1.52	1.73	3.21	2.58	1.41	3.26	2.6
No<1.0A	27	9	17	10	2	5	11	4	5
No<2.0A	37	17	30	28	18	22	34	10	14
No<3.0A	38	25	34	35	24	27	37	18	27

Fig. 6. Sphinx-H3 prediction accuracy. (**A**) Comparison of Sphinx and Rosetta antibody H3 loop prediction. (**B**) Performance of Sphinx antibody H3 loop prediction in a non-native environment

enhanced ability to sample near-native conformations means that its accuracy is still higher. It should be noted, however, that since members of the general loop set were used to set parameters used by Sphinx, this test may overstate its performance.

3.2 Antibody H3 prediction

One type of protein for which loop modelling is key is antibodies – a class of protein produced by the immune system to bind to and initiate the removal of foreign substances. Their exceptional ability to bind to a huge variety of antigens, with high specificity and affinity, makes them ideally suited for use as therapeutics (Chames *et al.*, 2009).

The region of the antibody that both contributes the most to antibody binding and introduces the most diversity is the H3 loop (the third CDR of the heavy chain) – it is found in the centre of the binding site, forms the most contacts with the antigen (Alzari *et al.*, 1988; MacCallum *et al.*, 1996) and has the greatest effect on the energetics of binding (Kunik and Ofran, 2013). Knowledge of H3 structures is therefore extremely useful, enabling predictions to be made about antibody binding properties and hence their suitability as drugs (Clark *et al.*, 2006; Diskin *et al.*, 2011; Kiyoshi *et al.*, 2014; Kuroda *et al.*, 2012; Krawczyk *et al.*, 2013; Lewis *et al.*, 2014; Lippow *et al.*, 2007; Thakkar *et al.*, 2014).

Unfortunately, H3 structure prediction is a particularly difficult problem, as demonstrated by the results of a recent blind prediction study (Teplyakov *et al.*, 2014), where the other five CDRs (L1, L2, L3, H1 and H2) were predicted with an average RMSD of 1.2, 0.5, 1.6, 1.1 and 1.0 Å respectively, while the H3 loop was predicted with an average RMSD of 3.5 Å. The difficulty arises due to the diversity of H3 structures, which is caused by gene recombination and junctional diversification (Jeske *et al.*, 1984; Tonegawa, 1983). H3 lengths are extremely variable and tend to be longer than general protein loops (Li, 2013), meaning that searching the entire possible conformational space is challenging. The use of H3-specific methods can help address this problem, however, there are currently less than 2500 antibody structures in the PDB (Dunbar *et al.*, 2014), each containing only one H3 loop, or a few copies of the same loop. Given the sequence and length diversity of H3, this covers only a marginal fraction of possible conformations. The challenging nature of the problem, however, makes H3 an ideal loop on which to test the performance of loop modelling algorithms (Sellers *et al.*, 2010).

We produced an H3-specific version of our algorithm, Sphinx-H3 (see Section 2 for details) and used it to predict the structures of 39 H3 loops, taken from the set used to test RosettaAntibody (Sivasubramanian *et al.*, 2009). To compare our results to other software, we also generated predictions using the H3 modelling routine within RosettaAntibody; a leading H3-specific *ab initio* loop prediction algorithm. This uses the next-generation kinematic modelling (KIC) approach described by Stein and Kortemme (2013) – the full command used is given in the Supplementary Information. The decoys obtained using Rosetta were ranked using its internal

scoring function. We generated 500 decoys per target, like in Stein and Kortemme (2013), and the same number from Sphinx that undergo minimization and final ranking. This allows us to compare execution times. Predictions were generated with the rest of the antibody in its native conformation, in the same manner as the second round of the second Antibody Modelling Assessment (Teplyakov *et al.*, 2014). Figure 6A shows the results of this comparison.

The ability of Sphinx-H3 to generate near-native H3 conformations is comparable to RosettaAntibody: the average RMSD of the best decoy generated is 0.94 Å (median 0.71 Å) for Sphinx-H3 compared to 1.07 Å (median 0.75 Å) for RosettaAntibody. After decoy ranking, Sphinx-H3 outperforms RosettaAntibody, producing an average top-ranked RMSD of 2.5 Å (median 2.31 Å) compared to 3.38 Å (median 3.25 Å), and giving predictions below 2 Å for 17 targets (compared to 13 for RosettaAntibody). On a target-by-target basis, Sphinx gave a more accurate prediction than RosettaAntibody 24 times, increasing to 28 when considering the best of the five top-ranked decoys.

In addition to achieving better accuracies, Sphinx is also much faster than RosettaAntibody. RosettaAntibody takes approximately 1 h to generate a single loop conformation (and hence up to 500 hours to generate the whole decoy set); Sphinx took (on average) 14 h per target to generate, minimize and rank all decoys, with the majority of the time being spent on minimization (approximately 12 h).

H3 loops can be classified as either 'kinked' or 'extended' depending on the conformation of their C-termini, with the majority being kinked (Weitzner *et al.*, 2015). Using the definition of Weitzner *et al.* (2015), we tested how often Sphinx correctly predicts a kinked or extended conformation. For the 30 kinked loops in the H3 target set, the top-ranked prediction of Sphinx was also kinked in 24 cases. This compares favourably to Rosetta, which only produced a top-ranked prediction of the correct conformation for 14 of the kinked targets. For both algorithms, five of the nine extended loops were predicted with the correct conformation.

We then used Sphinx to predict the H3 loops on antibody models rather than crystal structures, to see how accuracy is affected (Figure 6B). In the model environment, anchor residues are in a non-native conformation, making prediction more difficult. We generated model structures of the 39 antibodies in the set using ABodyBuilder (Leem *et al.*, 2016), a template-based fragment assembly method that builds antibody models using information from already known antibody structures. As expected, the results are not as good as when using the native environment – the average top-ranked RMSD increases from 2.5 (median 2.31 Å) to 3.26 Å (median 3.12 Å), and fewer sub-angstrom predictions are made. Interestingly, however, unlike the previous tests where minimization improved results, here

it appears to decrease accuracy. The average top-ranked RMSD and the average RMSD of the best decoy in the five top-ranked stay approximately the same, but the number of targets predicted with an RMSD below 2 Å decreases from 18 to 10. In 2014, the second Antibody Modelling Assessment (AMA-II) was held; a CASP-style blind prediction test specifically for antibody structure prediction. The second round of this competition focused entirely on H3 prediction - participants were given the crystal structures of ten antibodies with the H3 loops removed, and asked to model the missing residues (allowing them to submit five possible models per target). We have used Sphinx to predict these H3 structures to further compare our results with those of other groups (all results are given in the Supplementary Information). We used the same protocol as that used for the previous H3 loop set, except that fragments from PDB structures deposited later than March 2013 were ignored (to ensure a fair comparison). We report our results using the same RMSD measure as used in AMA-II – using only the carbonyl C and O atoms of residues 95–100x.

On average, the best of the five models returned by Sphinx has an RMSD of 1.41 Å. This compares well to the other groups, who achieved RMSDs of 1.86, 2.09, 1.97, 1.25, 2.41 and 1.12 Å (for Accelrys, CCG, the Gray group, the Shirai group, Macromoltek and Schrödinger, respectively). Only the Shirai group and Schrödinger achieved better results than Sphinx according to this measure.

When looking at the average RMSD across all models, Sphinx outperforms all the other methods with an RMSD of 2.17 Å compared to 2.89, 3.08, 3.22, 2.43, 3.08 and 2.54 Å. Sphinx is therefore more consistent than the other methods. For two targets (Ab04 and Ab10), the best model from Sphinx had a lower or the same RMSD as the best produced from all other methods. In addition, Sphinx was able to produce a sub-angstrom prediction for six of the ten targets, which is the most achieved by any method (equal to the number generated by Schrödinger).

4 Conclusions

We have developed a loop modelling algorithm, Sphinx, which combines both knowledge-based and *ab initio* methodologies in a novel way, such that structural information contained within fragments of a different length to the target loop can be used. This is important, since it has been shown that loops of different lengths can adopt similar conformations (Nowak *et al.*, 2016), and other methods do not make use of this information. Sphinx is capable of generating high-quality decoy sets that are enriched with near-native conformations and in addition produces more accurate predictions than can be achieved using a straightforward *ab initio* algorithm. Sphinx is also able to produce a prediction for every target, unlike some knowledge-based methods, which can fail if a suitable fragment of the correct length cannot be found.

The H3-specific version of our algorithm, Sphinx-H3, produces promising results, achieving comparable or better results than some of the leading H3-specific *ab initio* algorithms (Teplyakov *et al.*, 2014). The kinked/extended conformation of the H3 targets was correctly predicted in the majority of cases – even though no prior assumptions or predictions were made about the loops, and no restraints were imposed on the decoys to force a kinked/extended structure. Sphinx is also fast – for example, in comparison to RosettaAntibody, Sphinx gives increased or similar accuracy in under 3% of the computation time. Sphinx is therefore a more viable option for use in applications such as structure-based virtual screening.

We have demonstrated that, except potentially in the case of modelling H3 in the non-native environment, the inclusion of a minimization procedure improves prediction accuracies, even though decoy RMSDs may not be consistently improved. This appears to be because the ability of the ranking method to correctly select near-native decoys is enhanced. However, this does not occur when predicting onto model structures – this could be because although the models onto which the decoys are built are good, they do not have atomic accuracy, and hence errors exist – especially in the conformations of the side chain. As such, the loop minimization is being performed in the non-native environment, which may affect results adversely. Preliminary results show that ranking is improved by minimizing the entire structure, not just the loop region.

The results achieved by our method indicate that fragments that are a slightly different length to the target loop are indeed a valuable source of structural information. Using a hybrid approach allows Sphinx to use the structural data that is available without being restricted to only returning loop conformations that have previously been observed. Also, since we have demonstrated that this approach can be applied successfully to prediction of the difficult H3 loop, we believe that the algorithm can be easily adapted for other specific protein loop types and should continue to perform well. For example, it could be applied to membrane protein loops, or the other five antibody CDR loops, since recent exploration into length-independent canonical conformations show that different length CDR loops can adopt the same conformation (Nowak *et al.*, 2016).

While Sphinx is the first algorithm of its kind, the concept behind it could easily be applied to other loop modelling algorithms, allowing them to use the information from structurally similar but length non-identical loops. For instance, in the case of antibody H3 prediction, RosettaAntibody could be combined with a knowledge-based method [such as FREAD or H3Loopred (Messih *et al.*, 2014)] to offer an alternative.

From the results presented in this paper, we believe that in a real loop modelling situation a knowledge-based method such as FREAD should be tried first (since such methods are usually very fast). However, if a suitable fragment cannot be found, then the next approach should be to use a hybrid algorithm such as Sphinx. Instead of reverting to a pure *ab initio* algorithm that ignores previously observed structures, this allows more of the available structural data to be used, leading to more accurate predictions.

As the results presented here demonstrate that, using the extra structural information contained within different length loops, Sphinx produces high-quality decoy sets and accurate predictions in a fraction of the time of some existing algorithms. As the PDB expands and more structural data becomes available, Sphinx's performance should improve further.

Funding

Conflict of Interest: none declared.

References

Alzari,P.M. *et al.* (1998) Three-dimensional structure of antibodies. *Annu. Rev. Immunol.*, **6**, 555–580.

Baker,D.H. and Sali,A. (2001) Protein structure prediction and structural genomics. *Science*, **294**, 93–96.

Bonet,J. *et al.* (2014) Frag'r'Us: knowledge-based sampling of protein backbone conformations for de novo structure-based protein design. *Bioinformatics*, **30**, 1935–1936.

Canutescu,A.A. and Dunbrack,R.L.,Jr (2003) Cyclic coordinate descent: a robotics algorithm for protein loop closure. *Protein Sci.*, **12**, 963–972.

Chames,P. *et al.* (2009) Therapeutic antibodies: successes, limitations and hopes for the future. *Brit. J. Pharmacol.*, **157**, 220–233.

Choi,Y. and Deane,C.M. (2010) FREAD revisited: accurate loop structure prediction using a database search algorithm. *Proteins*, **78**, 1431–1440.

Chothia,C. and Lesk,A. M. (1987) Canonical structures for the hypervariable regions of immunoglobulins. *J. Mol. Biol.*, **196**, 901–917.

Chys,P. and Chacon,P. (2013) Random coordinate descent with spinormatrices and geometric filters for efficient loop closure. *J. Chem. Theor Comput.*, **9**, 1821–1829.

Clark,L.A. *et al.* (2006) Affinity enhancement of an in vivo matured therapeutic antibody using structure-based computational design. *Protein Sci.*, **15**, 949–960.

Deane,C.M. and Blundell,T.L. (2001) CODA: a combined algorithm for predicting the structurally variable regions of protein models. *Protein Sci.*, **10**, 599–612.

de Bakker,P.I.W. *et al.* (2003) Ab initio construction of polypeptide fragments: accuracy of loop decoy discrimination by an all-atom statistical potential and the AMBER force field with the generalized Born solvation model. *Proteins*, **51**, 21–40.

DePristo,M.A. *et al.* (2003) Ab initio construction of polypeptide fragments: efficient generation of accurate, representative ensembles. *Proteins*, **51**, 41–55.

Diskin,R. *et al.* (2011) Increasing the potency and breadth of an HIV antibody by using structure-based rational design. *Science*, **334**, 1289–1294.

Dong,G.Q. *et al.* (2013) Optimized atomic statistical potentials: assessment of protein interfaces and loops. *Bioinformatics*, **29**, 3158–3166.

Dunbar,J. *et al.* (2014) SAbDab: the structural antibody database. *Nucleic Acids Res.*, **42**, D1140–D1146.

Engh,R.A. and Huber,R. (1991) Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr. A*, **47**, 392–400.

Fasnacht,M. *et al.* (2014) Automated antibody structure prediction using Accelrys tools: results and best practices. *Proteins*, **82**, 1583–1598.

Fernandez-Fuentes,N. and Fiser,A. (2006) *BMC Struct. Biol.*, **6**, 15.

Fiser,A. *et al.* (2000) Modeling of loops in protein structures. *Protein Sci.*, **9**, 1753–1773.

Gront,D. *et al.* (2012) Optimization of protein models. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **2**, 479–493.

Hildebrand,P.W. *et al.* (2009) SuperLooper—a prediction server for the modeling of loops in globular and membrane proteins. *Nucleic Acids Res.*, **37**, W571–W574.

Hill,J.R. *et al.* (2011) Environment specific substitution tables improve membrane protein alignment. *Bioinformatics*, **27**, 15–23.

Holtby,D. *et al.* (2013) LoopWeaver: loop modeling by the weighted scaling of verified proteins. *J. Comput. Biol.*, **20**, 212–223.

Jacobson,M.P. *et al.* (2004) A hierarchical approach to all-atom protein loop prediction. *Proteins*, **55**, 351–367.

Jeske,D.J. *et al.* (1984) Junctional diversity is essential to antibody activity. *J. Immunol.*, **133**, 1090–1092.

Joosten,R.P. *et al.* (2011) A series of PDB related databases for everyday needs. *Nucleic Acids Res.*, **39**, D411–D419.

Kiyoshi,M. *et al.* (2014) Affinity improvement of a therapeutic antibody by structure-based computational design: generation of electrostatic interactions in the transition state stabilizes the antibody-antigen complex. *PLoS ONE*, **9**, 1–9.

Krawczyk,K. *et al.* (2013) Antibody i-Patch prediction of the antibody binding site improves rigid local antibody-antigen docking. *Protein Eng. Des. Sel.*, **26**, 621–629.

Krivov,G.G. *et al.* (2009) Improved prediction of protein side-chain conformations with SCWRL4. *Proteins*, **77**, 778–795.

Kunik,V. and Ofran,Y. (2013) The indistinguishability of epitopes from protein surface is explained by the distinct binding preferences of each of the six antigen-binding loops. *Proteins*, **26**, 599–609.

Kuroda,D. *et al.* (2012) Computer-aided antibody design. *Protein Eng. Des. Sel.*, **25**, 507–521.

Leem,J. *et al.* (2016) ABodyBuilder: antibody structure prediction with data-driven accuracy estimation. *mAbs*, **8**, 1259–1268.

Lewis,S.M. *et al.* (2014) Generation of bispecific IgG antibodies by structure-based design of an orthogonal Fab interface. *Nat. Biotechnol.*, **32**, 191–198.

Li,Y. (2013) Conformational sampling in template-free protein loop structure modeling. *Comput. Struct. Biotechnol. J*, **5**, e201302003.

Liang,S. *et al.* (2014) LEAP: highly accurate prediction of protein loop conformations by integrating coarse-grained sampling and optimized energy scores with all-atom refinement of backbone and side chains. *J. Comput. Chem.*, **35**, 335–341.

Lippow,S. *et al.* (2007) Computational design of antibody-affinity improvement beyond in vivo maturation. *Nat. Biotechnol.*, **25**, 1171–1176.

MacCallum,R.M. *et al.* (1996) Antibody-antigen interactions: contact analysis and binding site topography. *J. Mol. Biol.*, **262**, 732–745.

Mandell,D.J. *et al.* (2009) Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling. *Nat. Methods*, **6**, 551–552.

Martin,A.C. *et al.* (1989) Modeling antibody hypervariable loops: a combined algorithm. *Proc. Natl. Acad. Sci. USA*, **86**, 9268–9272.

Messih,M.A. *et al.* (2014) Improving the accuracy of the structure prediction of the third hypervariable loop of the heavy chains of antibodies. *Bioinformatics*, **30**, 2733–2740.

Messih,M.A. *et al.* (2015) A template-based tool for predicting the structure of protein loops. *Bioinformatics*, **31**, 3767–3772.

Michalsky,E. *et al.* (2003) Loops in Proteins (LIP)—a comprehensive loop database for homology modelling. *Protein Eng.*, **16**, 979–985.

Needleman,S.B. and Wunsch,C.D. (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.*, **48**, 433–453.

Nowak,J. *et al.* (2016) Length-independent structural similarities enrich the antibody CDR canonical class model. *mAbs*, **8**, 751–760.

Parsons,J. *et al.* (2005) Practical conversion from torsion space to Cartesian space for in silico protein synthesis. *J. Comput. Chem.*, **26**, 1063–1068.

Petoukhov,M.V. *et al.* (2002) Addition of missing loops and domains to protein models by X-ray solution scattering. *Biophys. J*, **83**, 3113–3125.

Rohl,C. *et al.* (2004) Modeling structurally variable regions in homologous proteins with Rosetta. *Proteins*, **55**, 656–677.

Samudrala,R. and Moult,J. (1998) An all-atom distance-dependent conditional probability discriminatory function for protein structure prediction. *J. Mol. Biol.*, **275**, 895–916.

Sellers,B.D. *et al.* (2010) Antibodies as a model system for comparative model refinement. *Proteins*, **78**, 2490–2505.

Shenkin,P.S. *et al.* (1987) Predicting antibody hypervariable loop conformations. I. Ensembles of random conformations for ringlike structures. *Biopolymers*, **26**, 2053–2085.

Sivasubramanian,A. *et al.* (2009) Toward high-resolution homology modeling of antibody Fv regions and application to antibody-antigen docking. *Proteins*, **74**, 497–514.

Soto,C.S. *et al.* (2007) Loop modeling: sampling, filtering and scoring. *Proteins*, **70**, 834–843.

Stein,A., and Kortemme,T. (2013) Improvements to robotics-inspired conformational sampling in Rosetta. *PLoS ONE*, **8**, e63090.

Teplyakov,A. *et al.* (2014) Antibody modeling assessment II: structures and models. *Proteins*, **82**, 1563–1582.

Thakkar,S. *et al.* (2014) Affinity improvement of a therapeutic antibody to methamphetamine and amphetamine through structure-based antibody engineering. *Sci. Rep.*, **4**, 3673.

Tonegawa,S. (1983) Somatic generation of antibody diversity. *Nature*, **302**, 575–581.

Wang,G. and Dunbrack,R.L.,Jr (2003) PISCES: a protein sequence culling server. *Bioinformatics*, **19**, 1589–1591.

Weitzner,B.D. *et al.* (2015) The origin of CDR H3 structural diversity. *Structure*, **23**, 302–311.

Whitelegg,N.R.J. and Rees,A.R. (2000) WAM: an improved algorithm for modelling antibodies on the web. *Protein Eng. Des. Sel.*, **13**, 819–824.

Xiang,Z. *et al.* (2002) Evaluating conformational free energies: the colony energy and its application to the problem of loop prediction. *Proc. Natl. Acad. Sci. USA*, **99**, 7432–7437.