Predicting protein residue–residue contacts using deep networks and boosting

Jesse Eickholt¹ and Jianlin Cheng¹,²,³,*

¹Department of Computer Science, ²Informatics Institute and ³C. Bond Life Science Center, University of Missouri, Columbia, MO 65211, USA

Associate Editor: Anna Tramontano

ABSTRACT

Motivation: Protein residue–residue contacts continue to play a larger and larger role in protein tertiary structure modeling and evaluation. Yet, while the importance of contact information increases, the performance of sequence-based contact predictors has improved slowly. New approaches and methods are needed to spur further development and progress in the field.

Results: Here we present DNCON, a new sequence-based residue–residue contact predictor using deep networks and boosting techniques. Making use of graphical processing units and CUDA parallel computing technology, we are able to train large boosted ensembles of residue–residue contact predictors achieving state-of-the-art performance.

Availability: The web server of the prediction method (DNCON) is available at http://iris.net.missouri.edu/dncon/.

Contact: chengji@missouri.edu

Supplementary information: Supplementary data are available at Bioinformatics online.

Received on June 28, 2012; revised on September 27, 2012; accepted on September 30, 2012

1 INTRODUCTION

The prediction of protein residue–residue contacts is seen by many as an important intermediate step for gaining traction on the challenging task of tertiary structure prediction. This idea has been spurred further recently by encouraging results that demonstrate that predicted contact information can indeed be used to improve tertiary structure prediction and effectively transform some unfolded structures into their folded counterpart (Wu et al., 2011). The addition of a contact guided structure modeling category in the Critical Assessment of Techniques for Protein Structure Prediction (CASP) on a rolling basis has also aided in sparking interesting in residue–residue contact prediction. Beyond the scope of tertiary structure prediction, protein residue–residue contacts have been used in drug design (Kligler et al., 2009), model evaluation (Wang et al., 2011) and model ranking and selection (Miller and Eisenberg, 2008; Tress and Valencia, 2010).

Existing methods for residue–residue contact prediction can be broadly categorized as sequence based or template/structure based. Sequence-based methods attempt to predict contacts from the primary sequence or information that can be derived directly from the sequence. A number of these sequence-based methods use various machine learning methods such as support vector machines (Cheng and Baldi, 2007; Wu and Zhang, 2008), neural networks (Fariselli et al., 2001; Hamilton et al., 2004; Polliastri and Baldi, 2002; Tegge et al., 2009; Walsh et al., 2009; Xue et al., 2009), hidden Markov models (Bjorkholm et al., 2009), Markov logic networks (Lippi and Frasconi, 2009), random forests (Li et al., 2011) and deep architectures (Di Lena et al., 2012) to make residue–residue contact predictions. Other sequence-based approaches have used evolutionary information contained in multiple sequence alignments (MSAs) to identify possible contacts (Gobel et al., 1994; Olmea and Valencia, 1997; Vicatos et al., 2005). Methods using MSAs were among the first sequence-based approaches tried but suffered from low accuracies caused by indirect or transitive correlations. More recent developments by Jones et al. (2012) using large MSAs and sparse covariance matrices have been better able to identify contacting residues from the alignments and resulted in significant improvements in accuracy. Template-/structure-based methods operate by extracting contact information from structural data. For template-based methods, this structural data comes in the form of templates (i.e. homologous proteins with known structure). Once templates have been found and aligned, residue–residue contacts are predicted using the contacts found in the structures (Wu and Zhang, 2008).

Given the relatively high quality of the tertiary structure models generated by template-based techniques, residue–residue contact data is most useful when dealing with hard targets (i.e. those for which a structural template does not exist or hard to identify by sequence alone). For hard targets, the conformational search is much larger and overall model quality is usually much lower. Thus, there is great interest in high quality sequence-based residue–residue contact predictors that do not rely on template data. Such contact predictors would be able to provide additional information when generating models for hard targets. Unfortunately, recent assessments of state-of-the-art sequence-based contact predictors routinely report average accuracies in the 20–30% range, indicating a need for further development and new methods and ideas (Ezkurdia et al., 2009; Izarzeguza et al., 2007; Monastyrsky et al., 2011).

Here we present a new sequence-based residue–residue contact predictor using deep networks (DNs) and boosting. Our method differs from other implementations of deep architectures owing to its boosted nature, overall network architecture and training procedure. More specifically, for training, we initially use an unsupervised approach to learn patterns in the data and initialize...
parameters and then fine tune them with back propagation. Furthermore, by using the computational power of graphical processing units (GPUs) and CUDA, we were able to train large boosted ensembles of DN classifiers achieving state-of-the-art performance.

2 METHODS

2.1 Datasets and evaluation metrics

Several datasets were used to train and evaluate our residue-residue contact predictor. The primary dataset, DNCON, was formed by an advanced search of the Protein Data Bank filtering the results by 30% sequence similarity and a resolution of 0–2 Å (Berman et al., 2000). The results from this initial search were then filtered by sequence length and disorder content, retaining sequences that were 30–300 residues in length and contained fewer than 20% disordered residues (i.e. coordinates were missing for fewer than 20% of the residues in the experimentally determined structure). The resulting set of proteins was then merged with the training set from SVMcon and then filtered by three existing datasets: D329, SVMCON_TEST and CASP9, which were used as evaluation sets. The filtering process ensured that the pairwise sequence identity between the merged dataset and any sequence in the evaluation sets was ≤25%. The end result of the search and filter process was our primary dataset, DNCON, consisting of 1426 proteins. This dataset was then randomly split into two sets: DNCON_TRAIN consisting of 1230 proteins and DNCON_TEST consisting of 196 proteins. Supplementary Figure S1 illustrates the entire dataset generation and filtering process.

The evaluation datasets used included D329, a set of 329 proteins used to evaluate Proc_S3 (Li et al., 2011); SVMCON_TEST, a set of 48 short to medium length proteins used to evaluate SVMcon (Cheng and Baldi, 2007); CASP9, a set of 111 targets used during the ninth Critical Assessment of Techniques for Protein Structure Prediction (Moult et al., 2011); CASP9_HARD, a subset of 16 targets taken from the CASP9 set that are solely composed of free modeling (FM) or free modeling/template-based modeling (FM/TBM) domains; and DNCON-TEST. Owing to the filtering process used in the creation of our dataset, all evaluation datasets are independent (i.e. ≤25% sequence identity) to the training set.

In this study, two amino acid residues are said to be in contact if the distance between their Cα atoms (Cα for glycine) in the experimental structure is ≤8 Å. Short-range contacts are defined as residues in contact whose separation in the sequence is ≥6 and <12. Likewise, medium-range contacts are residues in contact whose separation in sequence is ≥12 but <24 and long-range contacts are defined as having separation in the sequence ≥24 residues. These definitions are in agreement with recent studies and CASP residue-residue contact assessments (Eickholt et al., 2011; Ezekurdia et al., 2009; Grana et al., 2005; Izarzugaza et al., 2007; Li et al., 2011; Monastyrskyy et al., 2011; Tegee et al., 2009).

A common evaluation metric for residue-residue contact predictions is the accuracy of the top L/5 or L/10 predictions where L is the length of the protein in residues. In this context, accuracy (Acc) is defined as the number of correctly predicted residue-residue contacts divided by the total number of contact predictions evaluated. We also considered the coverage (Cov) of residue-residue contact predictions, which is defined as the number of correctly predicted contacts divided by the number of true contacts. As predicting short-, medium- and long-range contacts have varying degrees of difficulty, it is common to separate predicted contacts by sequence separation and then calculate the accuracy and coverage of the top L and L/5 predictions for each range. Note that this evaluation is done on a per target basis and irrespective of the domain architecture. Estimates for the standard error for accuracy and coverage were calculated using the sample mean and sample variance of the per-target accuracies and coverages.

2.2 Restricted Boltzmann machine and deep belief network

The general framework used for classifying residue-residue contacts was a combination of restricted Boltzmann machines (RBMs) trained to form DNs. A RBM is a two-layer network that can be used to model a distribution of binary vectors. In this model, a layer of stochastic binary nodes representing feature detectors are connected via symmetric weights to stochastic binary nodes that take on the values of the vectors to be modeled (Hinton, 2002; Smolensky, 1986). Conceptually, the layer of stochastic nodes corresponding to the feature detectors can be viewed as the ‘hidden’ or ‘latent’ data and the other layer of nodes as the ‘visible’ data. The energy of a particular configuration of this network can be defined by

\[
E(h,v) = - \sum_i h_i v_i - \sum_j c_j h_j - \sum_{ij} h_i v_j w_{ij}
\]

where \(v_i\) and \(h_j\) are the states of the \(i\)th and \(j\)th units, \(c_j\) is the bias for the \(j\)th visible node, \(w_{ij}\) is the weight of the connection between the \(i\)th visible and \(j\)th hidden nodes. A probability can then be assigned to a configuration of visible data by

\[
\rho(v) = \frac{e^{-E(v,h)}}{Z}
\]

where \(Z\) is a normalizing constant and the sum is over all possible configurations of \(h\).

Training consists of adjusting the weights of the model such that real data (e.g. training data) has a higher probability than arbitrarily chosen configurations of visible nodes. This can be done using a process called contrastive divergence, which adjusts the weights in a manner that seeks to minimize an approximation to a difference of Kullback–Leibler divergences (Hinton, 2002). The use of contrastive divergence learning as opposed to maximum likelihood learning has to do with a problematic term in the gradient of average log likelihood function, which is exponential in nature and difficult to approximate. With contrastive divergence, the terminal probability terms cancel out. Full details are provided in Hinton’s presentation on training products of experts (2002). In this work, the weights in the \(n\)-th epoch of training were updated as follows:

\[
\Delta^{(n)}w_{ij} = \epsilon (\langle v_i p_i^{(m)} >_{data} - < p_i^{(1)} >_{recon} )w_{ij} + \nu w_{ij}^{(n-1)}
\]

\[
\Delta^{(n)}b_i = \epsilon (\langle v_i >_{data} - < p_i^{(1)} >_{recon} ) + \nu b_i^{(n-1)}
\]

\[
\Delta^{(n)}b_j = \epsilon (\langle p_j^{(m)} >_{data} - < p_j^{(1)} >_{recon} ) + \nu b_j^{(n-1)}
\]

In these equations, the angle brackets represent averages over the batch. The subscripts ‘data’ and ‘recon’ are descriptors that illustrate that the first average is taken over the data and the second average over reconstructions of the data after one round of Gibbs sampling. For \(< >_{data}\), \(p_i^{(m)}\) is the probability that the \(j\)th hidden unit will be activated when driven by the data and calculated as

\[
p_i^{(0)} = \sigma (\sum_j w_{ij} v_j + b_i)
\]

where \(\sigma(\bullet)\) is the sigmoid function. For \(< >_{recon}\), \(p_i^{(1)}\) is the probability that the \(j\)th visible unit will be activated and is calculated as

\[
p_i^{(1)} = \sigma (\sum_j h_j w_{ij} + a_j)
\]

where \(a_j\) is a binary value set to 1 with probability \(p_j^{(m)}\). The final value to be computed is \(p_i^{(1)}\), which is the probability that the \(j\)th hidden unit will
be activated when driven by the probabilities of the reconstructed visible nodes and calculated as

\[ p^{(l)}(i) = \sigma \left( \sum p^{(l-1)}(w_j + b_j) \right) \]

In Equations 3, 4 and 5, \( \sigma \) is the learning rate, \( w \) is the weight cost and \( v \) the momentum. These parameters and update equations were used in accordance with recent findings on how to train RBMs in practice (Hinton, 2010). In our study, the learning rate \( \sigma \) was set to 0.01 for \( w \) and 0.1 for the biases, and the weight cost \( v \) was set to 0.0002. The momentum \( v \) was initially set to 0.5 and after 5 epochs of training increased to 0.9. Training a RBM was done using batches of 100 training examples over 20 epochs. During training, the average free energy of the training data was compared with that of a small holdout set taken from DNCON_TEST to confirm that the RBM was not over-fitting the training data. After training the RBM, the probability of selecting the \( i \)-th example from the training pool and \( y_i \in \{0, 1\} \) be the class of the \( i \)-th example. Also, let \( W(i) \) be the probability of selecting the \( i \)-th example from the training pool in the \( t \)-th round of boosting and call the DN classifier trained in round \( t \) to be \( m^t(\bullet) \), which outputs a value between 0 and 1. Now, after each round of boosting, \( W(i) \) is updated via \( e_t, a_t \) and \( h_t(\bullet) \) in the following manner.

\[ h_t(i) = \begin{cases} 0 & \text{if } m_t(x_i) < 0.5 \\ 1 & \text{if } m_t(x_i) \geq 0.5 \end{cases} \]

\[ e_t = \sum_{h_t(x(y))} W(i) \]

\[ a_t = \frac{1}{2} \left( \frac{1 - e_t}{e_t} \right) \]

\[ W_{t+1}(i) = \frac{W(i) e^{-a_t} \text{ if } h_t(x_i) = y_i}{e^{a_t} \text{ if } h_t(x_i) \neq y_i} \]

After 35 rounds of boosting, the final prediction for an input \( x_i \) is given by \( H(x_i) \) and is a value between 0 and 1.

\[ H(x_i) = \frac{\sum_{t=50}^{200} 0.5e^{-a_t} h_t(x_i)}{\sum_{t=0}^{35} a_t} \]

Supplementary Figure S4 shows a boosted ensemble. Additionally, we found that after several rounds of boosting, the weights of a number of particularly difficult training examples became disproportionately large and dominated the selected training data. Models trained on these sets did not generalize well and prohibited boosting beyond 10 rounds. This phenomenon is not new and has been studied elsewhere (Vezhnevets and Barinova, 2007). One solution was to reinitialize the weights on all the training examples after 7 rounds, which in effect joins bagging with several rounds of boosting. Another solution was to reinitialize the weights with a trimming procedure, which removed up to 30% of the hard training cases. This was achieved by removing (i.e. trimming) those training examples that were 5 times more likely to be selected than by chance.

2.3 Prediction of medium-/long-range contacts

For the prediction of medium- and long-range contacts, we trained multiple ensembles of DNs. The inputs for each DN included sequence-specific values for the residues in two windows centered around the residue–residue contact pair in question, several pairwise potentials, global features and values characterizing the sequence between the contact pair (see Section 2.7 for details). The target was a single binary value that represented the pair being in contact or not. For the contact pair (see Section 2.7 for details). The target was a single binary value that represented the pair being in contact or not. For the contact pair (see Section 2.7 for details). The target was a single binary value that represented the pair being in contact or not.

To facilitate working with large RBMs and DNs, we implemented the training and classification procedures in terms of matrix operations and used CUDAMat (Mnih, 2009), a python library that provides fast matrix calculations on CUDA-enabled GPUs. CUDA is a parallel computing platform that provides high-level access to the computing cores of certain graphics processing units (http://www.nvidia.com/object/cuda_home.html). Using CUDAMat and GPUs allowed us to train DN classifiers with on the order of 1 million parameters in under an hour.

2.4 Consensus predictions for medium-/long-range contacts

In addition to training ensembles for various fixed-length windows, we also averaged the prediction scores for contacts across ensembles. In all, there were 490 classifiers with 35 coming from each possible combination of window size (7, 9, 11, 13, 15, 17 or 19) and sampling scheme (reweighted or reweighted with trimming).

2.5 Prediction of short-range contacts

For the prediction of short-range contacts, we trained one ensemble of DNs. The input for each DN was a window 12 residues in length and the target was all short-range contacts whose residue pairs were contained in the window. In all, 400 features were used for each window and the target contained 21 predictions. The overall architecture of the DN was 400-500-500-250-21 (Supplementary Fig. S5). Each layer was trained as a RBM using the previous described process, and the entire network was
predict residue–residue contact predictors in CASP9 (Monastyrskyy et al., 2011). Table 1 shows accuracy and coverage of the top L/5 and top L predictions for ProC_S3, SVMcon and DNCON. The predictions for SVMcon and ProC_S3 were downloaded from the official CASP website (http://predictioncenter.org/download_area/ CASP9/predictions/). The evaluation dataset for this comparison was CASP9_HARD, a set of 16 proteins that were comprised solely of domains classified as FM or FM/TBM by CASP9 assessors. The FM and FM/TBM classification indicates that template-based information was scant or difficult to obtain for these targets. As seen in Table 1, DNCON performed well on these targets, achieving state-of-the-art performance for accuracy and converage of long-range contacts when considering the top L or L/5 contact predictions. Given the comparable performances of the methods, we also examined if the methods were identifying the same contacts or if they were in some sense complementary. We discovered that while there was some overlap between prediction sets (~18–30%), each method was identifying a number of unique true contacts among those selected (see Supplementary Tables S1–S3 for full details).

While evaluating residue–residue contacts on hard targets is arguably the best means of evaluation (i.e. it is on these types of targets that contact information may have the largest impact), the drawback is that these datasets are usually composed of a small number of targets. To increase the robustness of our evaluation, we also compared DNCON with SVMcon and ProC_S3
on larger datasets, namely SVMCON_TEST and D329. Table 2 presents the accuracy and coverage of the top $L/5$ and top $L$ predictions for SVMcon and DNCON on the SVMCON_TEST, the dataset used to evaluate SVMcon. Similarly, Table 3 presents the results for ProC_S3 and DNCON on the D329, a dataset used to evaluate ProC_S3. The predictions for SVMcon on the SVMCON_TEST dataset were obtained by downloading SVMcon and made locally and evaluated with our pipeline. Note that the values for ProC_S3 on the D329 dataset are those reported by Li et al. (2011) in their assessment of their method. Both of these evaluations show that the performance of DNCON is on par with the two state-of-the-art contact predictors (Table 3).

As an additional comparison between DNCON, ProC_S3 and SVMcon, we evaluated each method on all valid CASP9 targets. Again, the predictions for SVMcon and ProC_S3 were downloaded from the official CASP website. Note that this evaluation was done using the entire protein and meant to complement the assessment technique used by CASP assessors, which evaluates predictions on a per domain basis. Tables 4 and 5 show the results of the three methods when evaluated on long- and medium-range contacts. Once again, DNCON performs competitively against SVMcon and ProC_S3.

The final evaluation set used was DNCON_TEST, an evaluation set of 196 proteins from the dataset that we curated. Table 6 shows the performance DNCON on our evaluation set. We also calculated the sensitivity of DNCON on DNCON_TEST at the 95% specificity rate. For long-range contacts, 38% of the true long-range contacts can be recovered at the 95% specificity rate (i.e. 95% of non contacts are recovered). For medium-range contacts, the sensitivity is 44% at the 95% specificity level.

In addition to the evaluation on the entire DNCON_TEST dataset, we also evaluated our method on three subsets of DNCON_TEST. Using the CATH structure classification database (Cuff et al., 2011), we indentified and grouped 140 of the proteins in DNCON_TEST as mainly alpha, $\alpha$ (29), mainly beta, $\beta$ (111), and mixed (1).
beta (30) and alpha beta, alpha beta (81). These results are summarized in Table 7. Interestingly, our method appears to have more difficulty with mainly alpha proteins than with the mainly beta or alpha beta mix. Difficulty in predicting mainly alpha proteins has been noted elsewhere (Cheng and Baldi, 2007; Di Lena et al., 2012) and is a starting point for future study.

Finally, we assessed our contact predictor by evaluating its performance when considering neighborhoods. In this setting, a predicted contact is considered correct if there is a true residue-residue contact with ±delta residues for small values of delta (e.g. for delta = 1, a predicted contact [i,j] would be counted correct if there were a true contact at [i,j], [i±1,j], [j±1,i] or [i±1,j±1]). Table 8 states the accuracy of DNCON on the DNCON_TEST dataset for delta = 1 and delta = 2. These results demonstrate that while the predictions contain some noise, which prohibits residue-level precision, in general, the contact predictions are accurate and contain a strong signal. This strong signal, particularly at the short and medium range, could provide valuable local information, which could be propagated and incorporated into the prediction of longer range contacts. Recent work by Di Lena et al., has demonstrated that propagating neighboring contact predictions can indeed increase performance (2012). Given the flexibility of the DN architecture and its ability to handle a large number of input features, this type of local contact information could easily be included for longer range predictions. This is a line of investigation we are currently pursuing for a future work. As for long-range contact predictions, these too exhibit a strong relatively accurate contact signal, which may prove more useful than the residue-specific accuracies indicate, particularly for the purposes of guiding a search through the protein conformational space. An evaluation of SVMcon and Proc_S3 on DNCON_TEST using the neighborhood criteria found that Proc_S3 and DNCON perform comparably and both outperform SVMcon (data not shown).

3.2 Value of boosting and ensembles

To determine the value of the boosted ensembles and the consensus approach, we studied the performance of the predictions for various configurations of ensembles and across rounds of boosting. Figure 1 characterizes the affect of boosting on the accuracy of the top L and L/5 long-range predictions (see Supplementary Fig. S6 for effect on medium-range predictions). These particular figures are of the ensemble with windows of 13 residues in length and with the reweighted sampling scheme. They show the benefit of boosting and are typical of the affect seen in other ensembles.

To determine the effect of the window size on the method’s performance, we evaluated the performance using ensembles comprising DNs with only one window size and reweighting scheme. It is interesting to note that while all of the ensembles perform roughly the same, there is a marked difference in the performance of the individual ensembles and the consensus prediction for top L/5 predictions. Accuracies for the individual ensembles are in the range of 0.24–0.28 for the top L/5 long-range predictions, whereas the accuracy of their consensus is 0.34 (Supplementary Table S4). Similarly, for the top L/5 medium-range predictions, the accuracy jumps from the 0.32 to 0.34 range to 0.38.

4 CONCLUSION

In this work, we have presented DNCON, a new method for protein residue-residue prediction. The approach is based on two concepts, boosted ensembles and DNs, which are novel in

Table 7. Performance of DNCON on structurally defined subsets of the DNCON_TEST dataset

<table>
<thead>
<tr>
<th>Range</th>
<th>Type</th>
<th>Acc(L)</th>
<th>Acc(L/5)</th>
<th>Cov(L)</th>
<th>Cov(L/5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short</td>
<td>alpha</td>
<td>0.171</td>
<td>0.440</td>
<td>0.714</td>
<td>0.364</td>
</tr>
<tr>
<td></td>
<td>beta</td>
<td>0.285</td>
<td>0.612</td>
<td>0.748</td>
<td>0.318</td>
</tr>
<tr>
<td>Medium</td>
<td>alpha</td>
<td>0.201</td>
<td>0.512</td>
<td>0.723</td>
<td>0.350</td>
</tr>
<tr>
<td></td>
<td>beta</td>
<td>0.255</td>
<td>0.448</td>
<td>0.487</td>
<td>0.170</td>
</tr>
<tr>
<td></td>
<td>alpha beta</td>
<td>0.222</td>
<td>0.420</td>
<td>0.543</td>
<td>0.203</td>
</tr>
<tr>
<td>Long</td>
<td>alpha</td>
<td>0.111</td>
<td>0.156</td>
<td>0.112</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>0.222</td>
<td>0.322</td>
<td>0.147</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>alpha beta</td>
<td>0.234</td>
<td>0.384</td>
<td>0.169</td>
<td>0.053</td>
</tr>
</tbody>
</table>

Table 8. Performance of DNCON for short-, medium- and long-range contact prediction on the DNCON_TEST dataset when considering small neighborhoods

<table>
<thead>
<tr>
<th>Range</th>
<th>Acc(L/5)</th>
<th>Acc(L/2)</th>
<th>Acc(L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short (delta = 1)</td>
<td>0.789</td>
<td>0.657</td>
<td>0.532</td>
</tr>
<tr>
<td>Medium (delta = 1)</td>
<td>0.648</td>
<td>0.552</td>
<td>0.463</td>
</tr>
<tr>
<td>Long (delta = 1)</td>
<td>0.749</td>
<td>0.678</td>
<td>0.607</td>
</tr>
<tr>
<td>Long (delta = 2)</td>
<td>0.550</td>
<td>0.476</td>
<td>0.415</td>
</tr>
</tbody>
</table>

Fig. 1. Accuracy of the top L and L/5 long-range contact predictions for a boosted ensemble (13-win). The graph plots accuracy as a function of the number of rounds of boosting.
the context of residue–residue contact prediction. When compared with the current state-of-the-art, DNCON performs favorably, achieving state-of-the-art performance in the critical area of accuracy on top medium and long range contact predictions. When allowing for less than residue level precision, the performance of DNCON is even more impressive. Given the strong contact signal present for short- and medium-range contacts and the fast flexible architecture of DNs, in the future, we plan on modifying the DNs such that they can incorporate and propagate predicted short- to medium-range contacts when making longer range predictions. We also plan on refining the parameters and network architecture used to increase performance. The method is available as a web service at http://iris.rnet.missouri.edu/dncon/.

**Funding:** The work was partially supported by a National Library of Medicine Biomedical and Health Informatics Training fellowship (to J.E.) and a NIH NIGMS grant (R01GM093123 to J.C.).

**Conflict of Interest:** none declared.

**REFERENCES**


