Evaluating the Quality of NMR Structures by Local Density of Protons

Yih-En Andrew Ban,[1] Johannes Rudolph,[1]* Pei Zhou,[1] and Herbert Edelsbrunner[2]
[1]*Department of Biochemistry, Duke University Medical Center, Durham, North Carolina*
[2]*Department of Computer Science, Duke University, NC and Raindrop Geomagic, Research Triangle Park, North Carolina*

ABSTRACT **Evaluating the quality of experimentally determined protein structural models is an essential step toward identifying potential errors and guiding further structural refinement. Herein, we report the use of proton local density as a sensitive measure to assess the quality of nuclear magnetic resonance (NMR) structures. Using 256 high-resolution crystal structures with protons added and optimized, we show that the local density of different proton types display distinct distributions. These distributions can be characterized by statistical moments and are used to establish local density Z-scores for evaluating both global and local packing for individual protons. Analysis of 546 crystal structures at various resolutions shows that the local density Z-scores increase as the structural resolution decreases and correlate well with the ClashScore (Word et al. J Mol Biol 1999;285(4):1711–1733) generated by all atom contact analysis. Local density Z-scores for NMR structures exhibit a significantly wider range of values than for X-ray structures and demonstrate a combination of potentially problematic inflation and compression. Water-refined NMR structures show improved packing quality. Our analysis of a high-quality structural ensemble of ubiquitin refined against order parameters shows proton density distributions that correlate nearly perfectly with our standards derived from crystal structures, further validating our approach. We present an automated analysis and visualization tool for proton packing to evaluate the quality of NMR structures. Proteins 2006;62:852–864.**

Key words: Voronoi diagram; packing; statistical moment; Z-score; visualization; accuracy; inflation; compression; protein structure

INTRODUCTION

Solution structures of proteins determined by nuclear magnetic resonance (NMR) rely predominantly on the distance constraints derived from nuclear Overhauser enhancement and exchange spectroscopy (NOESY). A dense network of these distance constraints, along with dihedral angle constraints derived from vicinal scalar couplings and more recently individual bond orientations derived from residual dipolar couplings (RDCs), has enabled high-quality structural models of proteins to be determined by NMR. However, it is generally acknowledged that individual models of an NMR solution structure are less accurate compared with the corresponding model derived from X-ray crystallography. This indicates a need to improve models derived by NMR, including the development of structure evaluation tools to identify problems and measure improvements.

Numerous approaches have been used to evaluate the quality of NMR structures. These approaches can be roughly divided into two categories. First, specific to NMR, there are methods to check the precision or self-consistency of the constraints used in the structure calculations (e.g., NOE, RDC).[1–5] Second, there are generic quality measures that check the accuracy of the structure by measuring geometric deviations from assumed reality (e.g., PROCHECK,[6,7] WHAT IF,[8] MolProbity[9,10]). Although these quality measures have been initially developed and tuned for structures determined by X-ray crystallography, they have been recently applied with success to models generated by NMR spectroscopy. However, the NMR models are subject to errors different from those determined by X-ray crystallography. In particular, whereas the directly observed electron density in X-ray diffraction defines the overall volume of a protein, systematic errors in NOE distance calibrations, although small in scale, can lead to overpacked or underpacked NMR models. Overpacked NMR structures are usually associated with local van der Waals (vdW) clashes and can be readily detected by software such as MolProbity and WHAT IF. However, the deleterious effect of "inflation" in NMR structures is much harder to detect. Inflated or underpacked structural models can maintain correct backbone conformation and good local geometry. As a result, software designed to measure deviations from ideal geometry

The Supplementary Material referred to in this article can be found at http://www.interscience.wiley.com/jpages/0887-3585/suppmat/

Grant sponsor: National Science Foundation; Grant number: CCR-0086013; Grant sponsor: National Institutes of Health; Grant number: AI-055588; Grant sponsor: Whitehead Institute.

*Correspondence to: Johannes Rudolph, Department of Biochemistry, Mailstop 3813, LSRC, C125, Duke University Medical Center, Durham, NC 27710. E-mail: rudolph@biochem.duke.edu; Pei Zhou. E-mail: peizhou@biochem.duke.edu; or Herbert Edelsbrunner. E-mail: edels@cs.duke.edu

Received 8 July 2005; Revised 12 September 2005; Accepted 23 September 2005

Published online 9 December 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20811

(PROCHECK and WHAT IF) or atomic clashes (MolProbity) do not usually detect this distortion.

Underpacking reflects an overall increase of the occupied space for a protein model, suggesting that volume could serve as a suitable basis for measuring or improving the quality of NMR models. Indeed, the radius of gyration, which reflects the hydrated volume of a spherical protein, can be used as a constraint during NMR refinement.[11,12] Although this technique can minimize global inflation, it does not by itself provide a measure of local perturbations or detect simultaneous over- and underpacking for a model. Herein, we propose a quality measure based on the volumetric occupancy of individual atoms. Volumetric measurements of atoms and residues in proteins were first computed by Richards using the Voronoi diagram.[13–15] To date, the Voronoi diagram is still the most rigorous method of computing volumes in protein models, providing a stable framework to investigate packing.[16–19] Voronoi diagrams have also been used as a quality measure for protein models determined by X-ray crystallography.[20] As shown in Figure 1(a) (in two dimensions), a Voronoi cell for a specific atom is the space closest to that atom as defined entirely by its neighbors. Thus, all of the space can be filled with Voronoi cells of known volume [in three dimensions, Fig. 1(b)], each of which is specifically assigned to individual atoms of the protein.

In our approach, directed at protein models determined by NMR, we probe the local density of protons, a measure of how efficiently a proton uses its allocation of space in the Voronoi diagram. Specifically, we focus on backbone amide and interior protons that best reflect the core packing of a protein. Although the importance of protons in protein packing has been well documented,[21,22] prior volume studies and quality measures using volume have focused on heavy atoms. We use a set of high-quality X-ray models to characterize and establish standards for the local density of selected types of protons. Local density Z-scores, based on the statistical deviation from standard distributions of protons, increase as the resolution of crystal structures decreases and correlate with ClashScore,[21] a well-established quality measure for protein structure. Analyses of both database-wide and individual models show problematic inflation and compression for NMR structures across a wide range of mean pair-wise root-

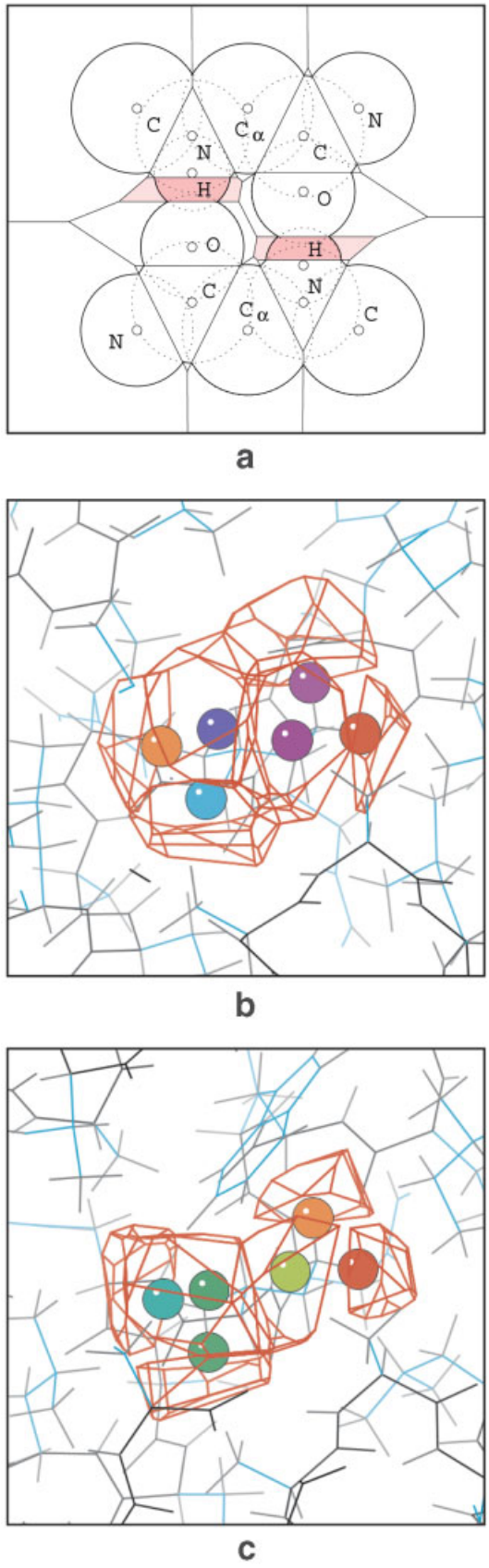

Fig. 1. Power diagram and local density. **a:** The power diagram in two dimensions. The configuration of atoms approximates a two-dimensional projection of the backbone unit of a β-sheet. The two backbone amide protons are colored in dark pink and their power cells are highlighted in light pink. Notice that only the fraction of the proton within its power cell is colored, illustrating the portion of the proton used to calculate its local density. Because of the use of the (non-Euclidean) power distance, portions of space contained within the atom, even the atom's center, may be assigned to a different atom's cell. The cells on the "boundary" of the diagram are unbounded. **b, c:** The power diagram in three dimensions. The same isoleucine in the MRF-2 protein is shown from the first-generation NMR structure 1BMY (ILE 72) in panel (b) and the second-generation NMR structure 1IG6 (ILE 60) in panel (c). There is an offset of 12 in the residue numbering between the structures. Atoms are colored based on their per-atom local density Z-score from magenta (inflated, at most −2.0 standard deviations) to green (about 0 standard deviations) to red (compressed, at least 2.0 standard deviations); see Methods for details.

mean-square deviations (MPRMSDs). Further refined NMR structures fare better, with significant improvements in packing quality seen for water-refined solution structures. In addition, analysis of an ensemble of a highly refined ubiquitin structure shows proton local density distributions that correlate nearly perfectly with our standards. Finally, we present an automated analysis and visualization tool for proton packing to evaluate the quality of NMR structures.

METHODS

Voronoi Diagram and Local Density

For our purposes, we use the space-filling representation of a protein, which abstracts an atom to a ball with a defined center and vdW radius. We then use the power or (weighted) Voronoi diagram to allocate space to atoms in the protein structure. The power diagram is defined by the power distance, which generalizes Euclidean distance to measuring distance from a point to a ball. The specific form is chosen to preserve the linearity of bisectors between balls. For a point x and an atom $b_i = (z_i, r_i)$ with center z_i and radius r_i, the power distance is defined as,

$$\pi_i(x) = \|x - z_i\|^2 - r_i^2 ,$$

where $\| \ \|$ is the *Euclidean norm*. For an atom b_i, the *power cell* includes all points in space at least as close to b_i as to any other atom,

$$V_i = \{x \in \mathbf{R}^3 | \pi_i(x) \leq \pi_j(x), \forall j\} .$$

There are two notable observations about a power cell: (1) each cell is a convex polyhedron that may be unbounded or empty, and (2) portions of space contained within the atom, sometimes even the atom's center, may be physically located in a different atom's cell [Fig. 1(a)]. The power diagram of a collection of atoms B is the collection of power cells V_i with the polygons, edges, and vertices shared by the cells. The local density of an atom b_i is the ratio of the atom's volume lying within its power cell to the volume of its power cell,

$$\rho(b_i) = \frac{\mathrm{Vol}(b_i \cap V_i)}{\mathrm{Vol}(V_i)} .$$

Notice that the atomic volume contribution is not from the full atom, but from the portion of the atom inside its power cell [as demonstrated by Fig. 1(a)]. The local density varies from 0.0, as in the case of an infinite power cell, to 1.0, when an atom completely fills its power cell. As an example, the densest possible sphere packing, the face-centered cubic lattice, has a local density of 0.7405.[23–25] We note that the balls in this densest packing are nonoverlapping and have equal radii, properties not enjoyed by the atoms in the vdW diagram of a protein. The power diagram and local density calculations are performed by extensions to the Ciel program, originally designed to construct protein–protein interface surfaces.[26] The vdW radii of atoms are those originally defined by Word et al.[21] Our method is dependent on this choice of atomic radii. Changing the atomic radii will change the power diagram and therefore the local density measurements, and an analysis of which would require a reparameterization of the standards. However, the differences between most atomic radii sets are reasonably small, so although there may be small differences in the resulting measurements and Z-scores, the overall quantitative results are expected to hold.

Data Generation

X-ray databases are derived from the Richardson Top500[27] and Richardson SCOP (J.S. Richardson and D.C. Richardson, personal communication). The Top500 dataset contains 500 structures chosen for high resolution, nonhomology, and wild-type sequence. The SCOP dataset of 1,785 structures contains a representative of structures from families in the structural classification of proteins (SCOP) databank.[28] Structures from X-ray databases were filtered manually to ensure a well-packed core by selecting proteins with 55 or more residues in length that have an elliptical shape (no single chain, "multidomain" structures) and do not contain buried cofactors or ligands. The resulting filtered Top500 and SCOP databases are denoted Pack256 and Test546, respectively. Pack256 has 256 structures with sizes ranging from 56 to 389 residues and resolutions ranging from 0.78 to 1.8 Å. Test546 has 546 structures with sizes ranging from 58 to 1,287 residues and resolutions ranging from 0.62 to 3.5 Å.

NMR databases are derived from the Protein Data Bank[29] (PDB) and DRESS.[30] NMR284 consists of 284 NMR structures derived from the PDB (edition @ 2005/01/27, 22:30p) using the following search criteria: (1) no DNA/RNA/glycoprotein/carbohydrate, (2) deposition after 2000/01/01, (3) minimum chain length = 60, (4) maximum chains = 1, (5) ensembles only, (6) mean backbone pairwise RMSD ≤ 1.1 Å, and (7) manual filtering by the same criteria applied to the X-ray databases. DRESS16 consists of 16 NMR structures from the DRESS dataset[30] (before and after water refinement for a total of 32 structures) filtered by hand via the same criteria as the X-ray databases with additional constraint of ensemble models only and 100 or more residues in length. Note that unstructured tails are removed from NMR structures by identifying the first and last secondary structure element, and the length criteria described above is evaluated after clipping. Additionally, we have analyzed a recently refined ensemble of high-resolution ubiquitin structures (PDB ID is 1XQQ[31]) and the 1.8 Å X-ray crystal structure (PDB ID is 1UBQ[32]).

Structure preparation

Structures were prepared for calculations by adding and optimizing proton positions via the Reduce program[22] with default options and a penalty value of 200. Note that methyl groups are not allowed to rotate. Secondary structure assignments were performed by the STRIDE program.[33] MPRMSD calculations for NMR ensembles were computed by applying rigid least-squares alignment[34] on backbone heavy atoms C, CA, and N. Lists for the PDB IDs of all structures are available in Supplementary Material 1.

TABLE I. The Five Proton Categories Chosen for Local Density Analysis[†]

Type	Protons from	Cutoff (local density)	Cutoff (no. of samples)[a]
α-Helix[b]	Backbone	0.15	6
β-Sheet[c]	Backbone	0.15	10
Methyl (CH_3)	VAL, LEU, ILE	0.15	30
Methylene (CH_2)	LEU, ILE, PHE, TYR	0.15	10
Aromatic ring[d]	PHE, TYR	0.04	12

[†]Cutoffs for local density and number of samples per protein are indicated.
[a]If the structure is an NMR ensemble, the cutoff is evaluated by comparing against the average number of protons per model.
[b]The four protons involved in the first turn of an alpha helix were not included because they do not form the same backbone hydrogen bonds as the rest of the helix.
[c]The parallel and anti-parallel beta sheet distributions were nearly indistinguishable, so these two categories were merged into one pool.
[d]Only the ε and δ protons on the aromatic rings of PHE and TYR are taken.

Scoring

Local density distributions are computed through kernel density estimates using a Gaussian kernel by the *stats* package in the R language[35] v1.9.1. The bandwidth of the kernel is computed by direct-plugin method using the KernSmooth package for R.[36] All results are plotted using R. We characterize distributions by four types of moments. The rth sample moment of a sample with size n taken about some point a is;

$$m_r = \frac{1}{n}\sum_{k=1}^{n}(x_k - a)^r .$$

Raw moments are taken about zero, whereas central moments are moments taken about the mean, μ. To calculate the moments, we use the following unbiased estimators.[37]

The mean (1st raw moment) $\mu = m_1 = 1/n\ \Sigma_{k=1}^{n}\ (x_k - 0)^1$ describes the average, and the variance (2nd central moment) $m_2 = (1/n-1)\ \Sigma_{k=1}^{n}\ (x_k - \mu)^2$ describes the spread of the distribution. The 3rd central moment $m_3 = (n/(n-1)(n-2))\ \Sigma_{k=1}^{n}\ (x_k - \mu)^3$ describes the symmetry of the distribution, with a tail to the left or right of the distribution resulting in a negative or positive 3rd central moment, respectively. The 4th central moment $m_4 = (n^2\ \Sigma_{k=1}^{n}\ (x_k - \mu)^4) - 3(2n^2 - 5n + 3))\mu_2^2/(n^3 - 4n^2 + 6n - 3)$ where $\mu_2 = \langle (x - \mu) \rangle$ is the population variance, complements the variance by being more sensitive to infrequent and extreme deviations. A sharper distribution results in a smaller 4th central moment, whereas a rounded distribution results in a larger 4th central moment. In practice, we use the sample variance m_2 in place of μ_2.

Using these moments, we develop a scoring methodology to evaluate local density distributions for proton types in single proteins. Each moment, m_i, is computed for all proton categories across all proteins of the reference dataset (Pack256) to determine its average, μ_i, and standard deviation, σ_i. To avoid skewed distributions from proteins with insufficient samples for a particular proton type, we implement the hard cutoffs listed in Table I. For the secondary structure categories, these cutoffs are derived by taking the average number of samples per protein and subtracting one standard deviation. For side-chain categories, this same cutoff is too restrictive and we instead require an equivalent of five amino acids for nonaromatics and three amino acids for aromatics. We evaluate a sample distribution by comparing each of its four moments against these average values. For a moment m_i from a sample distribution, the Z-score is defined as

$$z(m_i) = \frac{m_i - \mu_i}{\sigma_i},$$

where μ_i and σ_i are the average and standard deviation of the i-th moment, as defined above. In summary, a proton category receives four Z-scores, one for each of its moments. To give an overall score to a protein, we define an overall Z-score for each moment m_i as a linear combination of the Z-scores for m_i from each proton category,

$$z_{overall}(m_i) = \sum_c \frac{n_c}{n_t} z(m_i)$$

where c is a proton category, n_c is the number of samples in category c, and n_t is the total number of samples for all categories. For a given protein, not all proton categories may necessarily be present. If the original (biased) forms of the sample moments are used, then this expression essentially evaluates a single proton's local density by the average and standard deviation for a moment of that proton's standard local density distribution. Note that $z_{overall}$ may potentially suffer from cancellation because of positive and negative Z-score terms. However, empirical observations using our standard datasets show problems caused by cancellation are minimal. Finally, remember that any overall score is a compromise that roughly estimates the quality of the protein and inevitably results in some loss of information regarding the details of the measurements.

The previously described scoring methodology applies to single proteins. This method uses an average and standard deviation for each moment, resulting in a mean of means, which is equivalent to weighting each of the observations

(single local densities) from a protein by the number of observations the protein contributes to the total random sample. To construct an analogous scoring method for individual protons, we begin by constructing weighted local density distributions for each proton type that weight an observation by the total number of observations a protein contributes to the total random sample. The mean and standard deviation for the weighted distribution are calculated (the equivalent of fitting to a Gaussian using the method of maximum likelihood), and used to Z-score a given observation (single local density). Tables in Supplementary Material 2 show these values.

Visualization

A kinemage[38] is generated to visualize the Z-score on the protein structure through the KiNG program.[10,39] Z-scores are shown as small colored balls located on the center of the proton. Color ranges for Z-scores follow the standard kinemage palette ranging from extremes of magenta (-2.0) to red (2.0) in increments of 0.4. Atoms with Z-scores beyond ± 2.0 are assigned the extremal colors of red or magenta. A ball is not drawn for atoms not satisfying the local density cutoffs. We consider standard deviations beyond ± 1.9 to be sufficiently compressed or inflated to warrant further review.

Box-and-whisker plots are generated using R with the default parameters. Briefly, the median is specified by the line within the box. The top and bottom of the box are the hinges, which are approximations of the 1st and 3rd quartiles. The whiskers extend out to the most extreme sample in the bin, but not beyond 1.5 times the length of the box. Any samples outside of this range are drawn individually as points. The width of the boxes is proportional to the square roots of the number of samples in the bin.

Individual protein reports are generated using R. The mean and standard deviation of the weighted distribution are taken to specify a Gaussian distribution (as described above) and used as reference curves (dashed) in the plots.

Although we have excluded structures with buried cofactors and ligands from the analyses in this work for clarity and dataset consistency, our standards and calculations are applicable to these cases and handled properly by the automated calculation on our Web site.

RESULTS AND DISCUSSION

It is well known that the radius of gyration of a globular protein, which describes the volumetric occupancy of the protein, is largely proportional to its molecular weight. This suggests that the average volumetric occupancy for individual atoms inside the protein is largely invariant. However, different types of atoms likely occupy different space because of their different chemical natures. To test this hypothesis, we have quantified packing by measuring the local density of different types of protons for Pack256, a set of high-quality X-ray structures. We chose protons as they form most of the noncovalent contacts associated with packing. In contrast, packing of heavy atoms is typically dominated by covalent bonds to other atoms and therefore is less sensitive as a measure for local density.

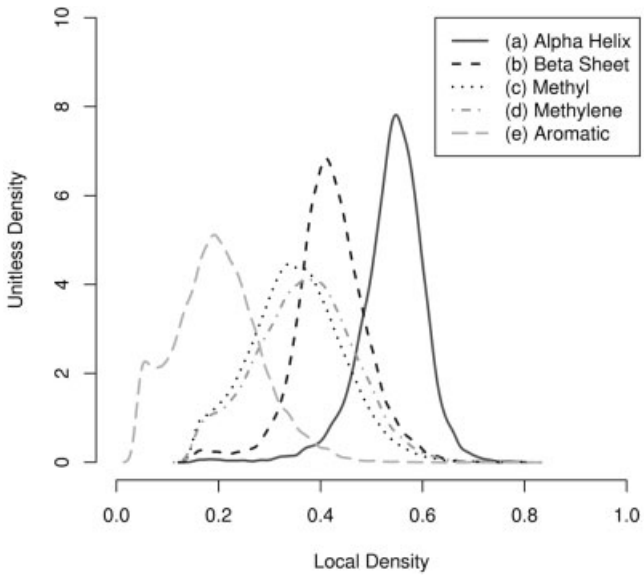

Fig. 2. Local density distributions of the following type. **a:** α-Helix backbone amide protons, excluding the first turn of the helix; 10,163 samples, mean = 0.540. **b:** β-Sheet backbone amide protons; 8,980 samples, mean = 0.419. **c:** Methyl protons from VAL, LEU, ILE; 53,235 samples, mean = 0.352. **d:** Methylene protons from LEU, ILE, PHE, TYR; 19,439 samples, mean = 0.368. **e:** δ and ϵ aromatic ring protons from PHE, TYR; 13,343 samples, mean = 0.198.

To emphasize core packing, we restrict our density measurements to buried protons in two ways. First, we discard all surface-exposed protons that have infinite or very large Voronoi cells resulting in low local densities (Table I). Second, we select protons expected to be in the buried core including methyls, methylenes, and aromatics from hydrophobic residues and the backbone protons from α-helices and β-sheets (Table I). Although we initially computed and characterized the local densities and distributions for all possible groupings of protons (single atom type, single residue, etc.), only protons from these five categories had sufficient signal strength and representation of samples to give distinct distributions (Fig. 2). The protons in these categories were grouped on the basis of similarity. For example, the local density distributions for parallel and antiparallel β-sheets were essentially indistinguishable, as were the methyl groups of valine, leucine, and isoleucine. The backbone protons from the 3-10 helix, π-helix, turns and coils, and the methanyl, hydroxyl, amine, and thiol side-chain protons were not chosen because of either noisy or weak signals. Note that many of these functional groups belong to charged or polar residues and are generally closer to the surface of the protein.

Score Standards

We use the high-quality Pack256 dataset to set standards for scoring (Table II). Evaluating Pack256 using these standards then serves as an elementary check of self-consistency. Additionally, the spread in scores for these high-quality structures indicates where to set thresholds for structures that have nonstandard packing and therefore require further evaluation. To describe and

TABLE II. Standard Values for the Four Moments†

	Helix	*Sheet*	*Methyl*	*Methylene*	*Aromatic*
m_1					
μ	0.5409	0.4208	0.3509	0.3667	0.1955
σ	0.0189	0.0167	0.0133	0.0169	0.0152
m_2					
μ	0.0037	0.0052	0.0083	0.0091	0.0070
σ	0.0021	0.0022	0.0014	0.0019	0.0018
m_3					
μ	−0.0003	−0.0002	0.0001	0.0000	0.0002
σ	0.0004	0.0004	0.0003	0.0003	0.0004
m_4					
μ	0.0001	0.0001	0.0002	0.0002	0.0002
σ	0.0002	0.0001	0.0001	0.0001	0.0002

†μ is the average and σ is the standard deviation of the moment computed across all proteins of the Pack256 dataset.

evaluate each of the proteins in the Pack256 dataset, we characterize the packing behavior of each proton type as described by the four moments of their local density distribution. As detailed in the Methods, the standard values in Table II are used to score the moments for each category of protons of a protein's local density distributions, resulting in a 4-vector Z-score. The protein's overall distribution is also evaluated with a 4-vector overall Z-score constructed from component Z-scores, and is dependent on the secondary structure and amino acid composition of the protein. Figure 3(a) shows boxplots for the Z-scores of the overall local density distributions versus the resolution of the X-ray structures. The stability of all four moments reflects the similarity in proton packing and consistent high quality of the structures in Pack256. The Z-scores for all individual proton categories generally show similar behavior with the secondary structure protons having a tighter range of Z-scores than the side-chain protons (data not shown). This tighter range arises from the more uniform packing enforced by their hydrogen bonds. Also, an upper bound of only one standard deviation in the 3rd central moment Z-score for the backbone amide protons of α-helices is seen (data not shown), probably arising from the already high average local density (0.54) of these protons (Fig. 2).

There are only two structures (1TGX and 1BBZ) with Z-scores beyond ±3 standard deviations for the mean and 3rd central moment, and +3 standard deviations for the variance and 4th central moment. 1BBZ has an unusual flat shape with a visually less-well-packed core. 1TGX has a small β barrel structure resulting in a high amount of contortion in the β-sheets. In general, however, the overall regularity of the Pack256 dataset allows us to establish an alarm threshold that is consistent for all moment Z-scores across all proton categories. A warning signal is given off for Z-scores beyond ±3 standard deviations for the mean and 3rd central moment, and +3 standard deviations for the variance and 4th central moment.

We briefly note here several observations relevant to evaluating our standard. First, methyl group rotation was not allowed during assignment of proton positions during structure preparation by Reduce. Improvements in Z-scores are negligible when allowing methyl group rotation, and only seen for a few isolated cases by 0.1 to 0.2 standard deviations, which is within error. Second, assignment of secondary structure by STRIDE to classify backbone amide protons works well. Misclassification does occasionally occur, but these protons are small in number relative to the entire population. A specific example is when STRIDE marks a helix-like turn before a true helix as part of the helix. In this case, the protons involved in this helix-like turn and the true (non-hydrogen bonding) first turn of the true helix are included in our local density calculation and found to be inflated. Thus, it is important to look at the details of the scoring on the structure using a generated kinemage, which we describe later in the text. Finally, it is possible that Pack256 may not sample all packing environments for each proton type. Because we have seen essentially identical standards from both a smaller dataset, the Richardson Top100[21,27] as well as the full Top500, this suggests that the sampling size is sufficiently large.

X-Ray Test

As an initial evaluation of our local density scores, we used Test546, a larger dataset of X-ray structures. Interestingly, the medians of the overall mean Z-scores as well as their range versus resolution are relatively constant, similar to the signature seen in Pack256 [Fig. 3(b)]. This is likely the result of the fixed allocation of space used in fitting the observed electron density. In contrast, the Z-scores for the variance, 3rd central moment, and 4th central moment all exhibit an increase versus resolution, beginning from about 1.6 Å [Fig. 3(b)]. Because the mean is constant, the increase in the 3rd central moment Z-score from roughly zero to positive suggests that more residual protons exist in a compressed environment, in agreement with previous observations that there is an increase in atomic clashes at lower resolutions.[21] The increase in the variance and 4th central moments at lower resolution reflects the increased overall spread of the local densities as more protons are found in compressed environments. Individual proton categories now begin to have more distinct signatures. β-Sheet protons have the most consistent packing versus resolution, particularly in the 3rd central moment, whereas methyl protons have the steepest increase in 3rd central moment Z-scores versus resolution (data not shown).

Comparison With ClashScore

To further evaluate our local density scores, we compared the overall Z-scores with the quality measure provided in the Reduce/Probe program suite (MolProbity) using the Test546 dataset. The ClashScore computed with this suite is defined as the number of non-hydrogen bond overlaps of at least 0.4 Å per thousand atoms and has been shown to increase as X-ray resolution decreases.[21] As expected, there is no significant correlation between the mean Z-score and ClashScore (Fig. 4). Although the presence of atomic clashes may affect the local density, they

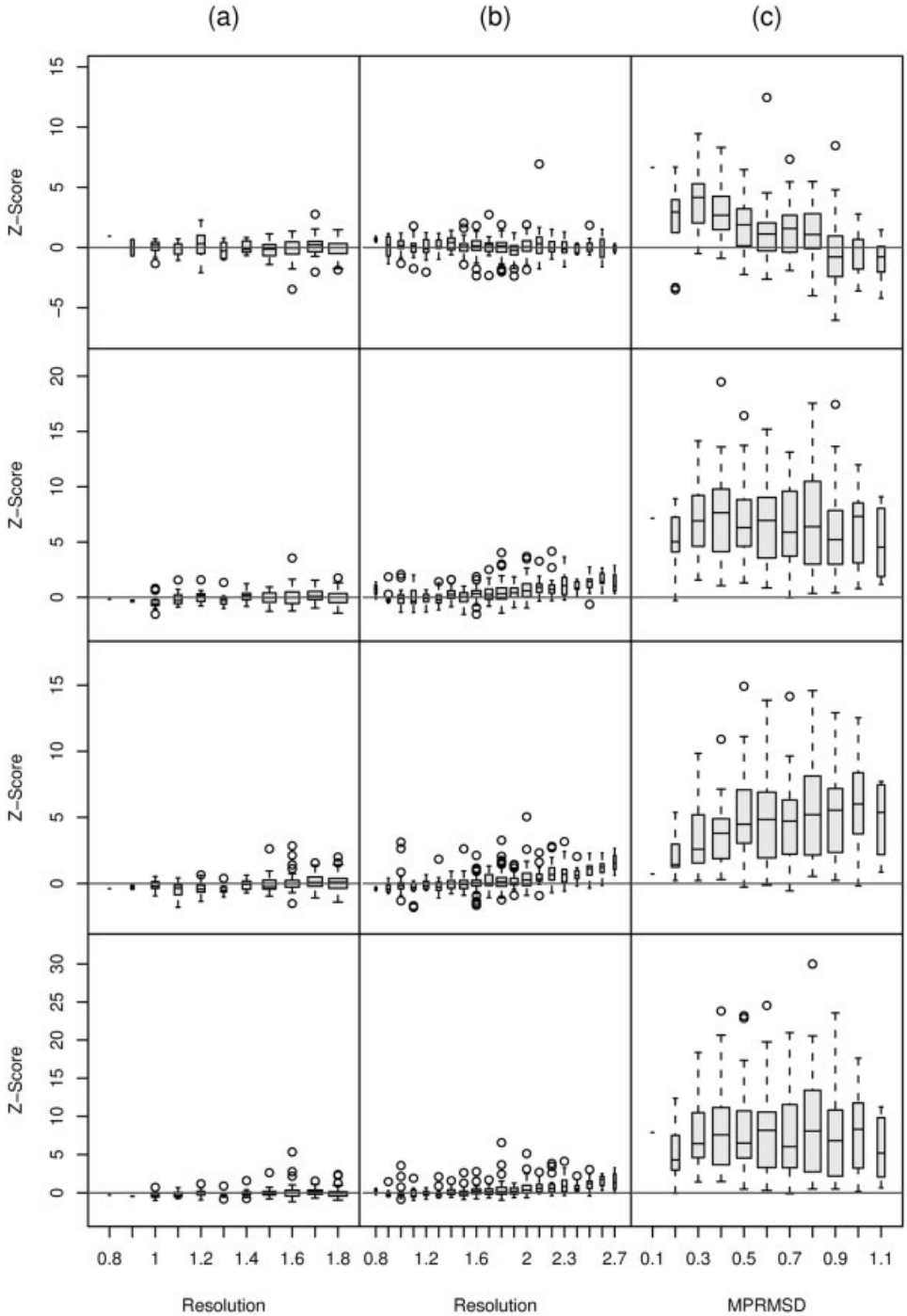

Fig. 3. Boxplots are drawn so that the width of the boxes is proportional to the square roots of the number of proteins in the bin. The rows from top to bottom are the mean, variance, 3rd central moment, and 4th central moment. **a:** Overall local density distribution Z-score versus X-ray resolution for the Pack256 dataset. **b:** Test546 overall Z-scores versus X-ray resolution. One structure below 0.7 Å and 10 structures above 2.7 Å are not shown because of insufficient samples for bins. **c:** NMR284 overall Z-scores versus MPRMSD.

are not a measure of average volume. Instead, the effect of clashes manifests itself as perturbations in volume, as shown by the correlation in the variance, 3rd, and 4th central moment Z-scores with ClashScore. As seen in the increasingly positive Z-scores of the 3rd central moment, these perturbations in volume are partially attributable to the increased local densities of the clashing atoms. Note, however, that high local density is not always a result of clashing atoms, illustrating the subtle difference between the traditional connection of overpacking and clashing atoms in proteins, and that of compression in the local density measure. As an extreme example, we have ob-

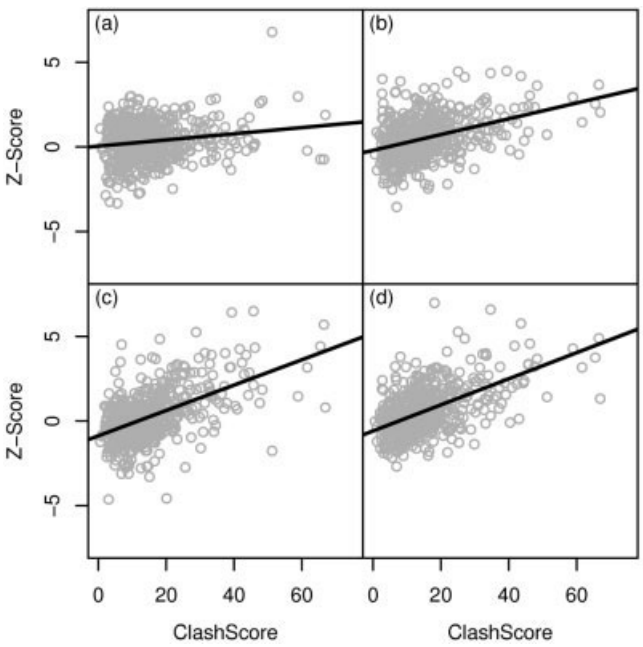

Fig. 4. Test546 overall Z-scores versus ClashScore. One structure with a ClashScore of 135 is not shown. Fits to a linear model are shown by the black line. **a:** Mean Z-score, slope = 0.018, R^2 = 0.034. **b:** Variance Z-score, slope = 0.047, R^2 = 0.186. **c:** Third central moment Z-score, slope = 0.075, R^2 = 0.326. **d:** Fourth central moment Z-score, slope = 0.077, R^2 = 0.347.

served a few cases, particularly in very high-resolution X-ray structures, where protons are highly compressed with atomic Z-scores of 2.4 or greater but are very well packed with no atomic clashes. Examples include a backbone amide proton in 1CEQ involved in the turn of a β-sheet and two backbone amide protons in 1QQ4 involved in twisted β-sheets. Because there are no distinct clashes around these protons, these may be cases of true compression and may be of interest for further structural study. However, there are very few of these protons and overall populations typically are not skewed by these unusually compressed protons. It is of theoretical interest to point out that although small clashes can cause the local density to increase, extreme clashes can actually cause a decrease in the local density because of the reduction in volume contribution from the clashing proton. In practice, this has not been observed because of the high energy barrier of these extreme vdW clashes.

NMR Test

Having established a correlation between our local density measure and quality of protein structures determined by X-ray crystallography, we next investigated packing in NMR284, a filtered random sample of recent NMR structure ensembles chosen from the PDB. The precision of NMR structures can be measured in terms of MPRMSD, wherein a higher MPRMSD reflects a larger uncertainty in the model.[40] The overall Z-scores of the local density distributions versus MPRMSD in NMR284 show much larger spreads for all proton categories than those seen in the X-ray datasets [Fig. 3(c)]. At low MPRMSD, we observe compression in the mean Z-scores with structures significantly above the threshold of three standard deviations. Most of this compression is the result of side-chain protons, whereas the backbone protons score more moderately (data not shown). As the MPRMSD increases, the mean Z-score decreases and approaches more moderate values whereas the high 3rd central moment Z-scores reveal an increasing asymmetry toward compression. This suggests that as MPRMSD increases, the large residual populations of protons that remain compressed must be accompanied by the remaining protons becoming more inflated. This simultaneous compression and inflation causes a wide spread of local density values, as indicated by the constantly high variance and 4th central moment Z-scores. Although MPRMSD is a different measure than X-ray resolution, these results indicate a significant departure from the stable mean Z-score we saw previously in X-ray structures. NMR structure ensembles with high MPRMSD appear to have regions of both excessive compression and inflation. NMR structure ensembles with low MPRMSD have a significant tendency toward compression, and may have been overly constrained in the refinement process, resulting in structures that are too tight. Importantly, the ensemble for an NMR structure generally scores better than the individual models (data not shown). Comparing with the standards derived from the X-ray data, these findings suggest a quantitative method for evaluating the quality of an NMR structure.

Local Density Profile

Because individual protein structures and different proton types in a protein can vary in their local density distributions, we have created the local density profile, an evaluation sheet summarizing the local density measurements for a protein (Fig. 5). These evaluations compare query structures simultaneously against our standard distribution and are accessible via our Web server (http://biogeometry.cs.duke.edu/research/localdensity/). There are six panels per page, a summary panel with the calculated moments and overall Z-scores, and five cells representing the five proton categories with estimates of the local density distributions, the Gaussian fit of the weighted Pack256 standard distributions (see Methods), and the Z-scores. A proton category plot is left empty if there are not enough protons or if the category is absent altogether. Bar plots are used to help visualize Z-scores for each proton category, and the solid portions of the box bounding the bar plots indicate the three standard deviation threshold. To minimize overlaps in the bars, variance and 4th central moment bars are plotted so that positive values point down and negative values (better than the standard) point up. Figure 5 shows the evaluation sheet for two structures of MRF-2, a member of the AT-rich interaction domain family of DNA binding proteins.[41,42] A careful comparison of the first generation, 1BMY,[42] and second generation, 1IG6,[41] NMR structures reveals a significant improvement in the second-generation NMR structure.

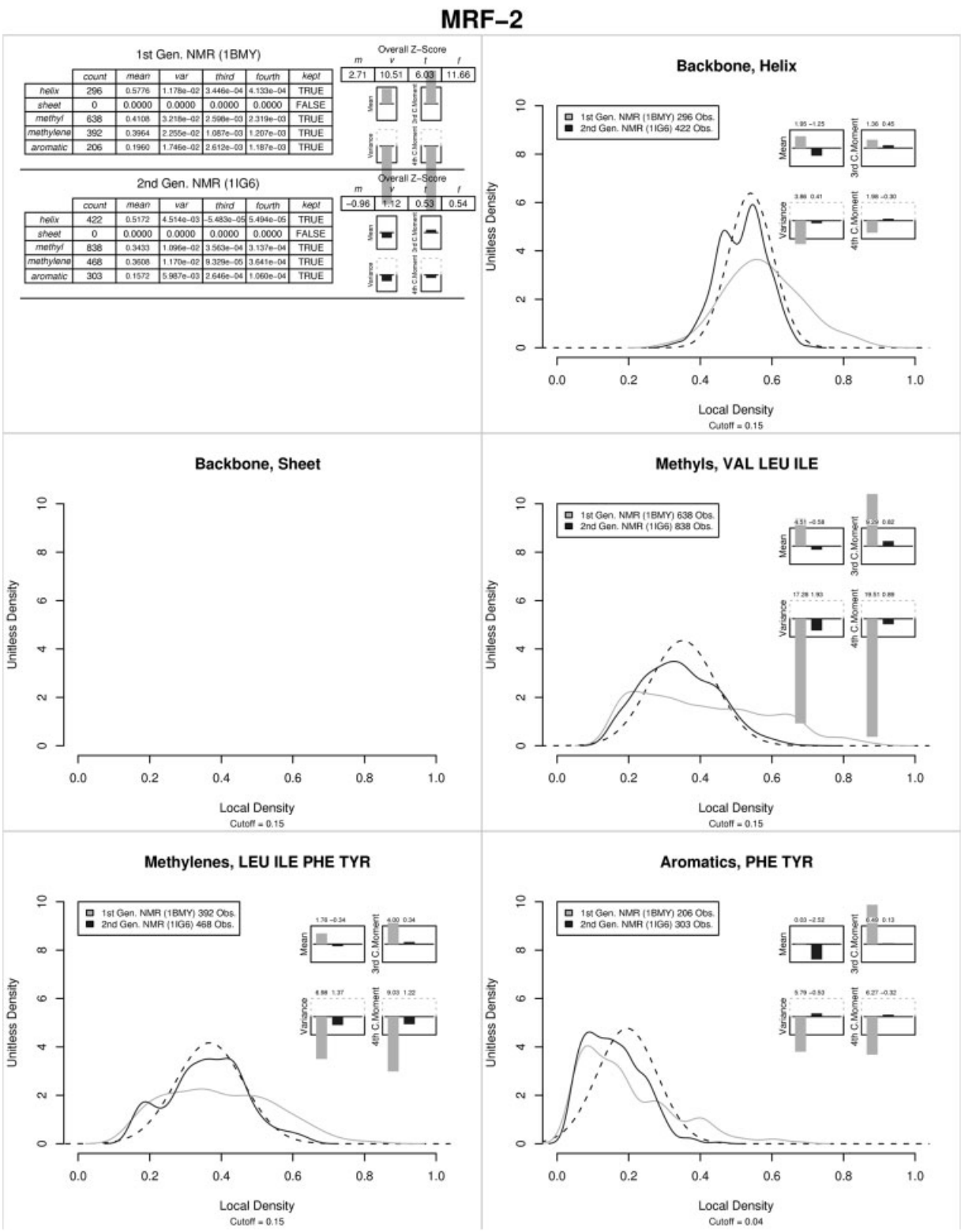

Fig. 5. Evaluation sheet of local density distributions and scores comparing 1BMY and 1IG6. The upper left panel summarizes the raw statistics of our local density evaluation in a table. The remaining five panels display the local density distributions for each of the proton categories. The X-ray standard derived from Pack256 is shown as a dotted line, and up to three different proteins are shown by solid colored lines. Insert boxes in each panel graphically depict the Z-scores, with bars extending outside solid borders of the box if the Z-scores are greater than our warning thresholds defined in the text.

TABLE III. DRESS16 Statistics†

	No. of proteins improved	Average improvement	Best improvement	Worst deterioration
Mean	9	−0.5	−2.8	3.6
Variance	12	−3.4	−11.7	3.9
3rd central moment	14	−3.9	−10.8	0.5
4th central moment	13	−5.6	−16.7	3.6

†Improvement/deterioration are in units of standard deviation and calculated as the difference between the absolute refined Z-score and the absolute original Z-score.

The local density distributions for the first-generation structure are all much too wide, as shown by the large variance Z-scores. The second-generation structure scores well in all proton categories with all moment Z-scores below the threshold of three standard deviations. In fact, all but one moment Z-score is below two standard deviations. The helix protons of 1BMY appear to be compressed, with mean and 3rd central moment Z-scores having the same sign, as well as a variance Z-score 0.86 standard deviations above threshold. In contrast, 1IG6 has a tighter local density distribution, with small variance and 4th central moment Z-scores. The mean Z-score indicates a slight amount of inflation which is seen in the local density distribution as a smaller, second peak. Finding the reason for this second peak requires examining details of the atomic local density Z-scores in the actual structure (see below). The methyl proton local density distribution for 1BMY is very flat with no discernible peak. The large plateau in the distribution extending past 0.5 (local density) indicates that most of the methyl protons are overly compressed and is reflected in the large positive mean and 3rd central moment Z-scores. Notice that the variances and 4th central moments are the largest of all the proton categories; so large that they are in fact indicators of both extreme compression (the small tail around 0.8 in the distribution) and inflation (the large portion of the distribution around 0.2). The methyl protons of 1IG6 are packed more normally, indicated by the well-shaped local density distribution and all the Z-scores well below warning threshold. Similar phenomena were seen with the methylene proton local density distributions. The methylene distribution for 1BMY is again flat and shows indicators of both extreme compression and inflation. The 1IG6 distribution is more well shaped and is a better match to the standard, although there appears to be a few inflated protons, as indicate by the small bump in the left of the distribution. The local density distribution of the aromatic protons for both 1BMY and 1IG6 shows significant perturbations from the standard. Even though the mean Z-score of 1BMY seems to be centered well, the high variance and 4th central moments show that there is both inflation and compression. The positive 3rd central moment Z-score captures the set of residually compressed protons in the tail to the right of the distribution. The 1IG6 aromatic proton distribution is much tighter, with 3rd central moment, variance, and 4th central moment Z-scores relatively close to zero. However, the mean Z-score indicates that the mass of the distribution has shifted left. Although the mean Z-score (−2.52) is within the three standard distribution threshold, this shows that there is still some slight inflation in the aromatic protons of 1IG6.

Visualizing changes in local density and volume in protein structures with the typical representations (lines, spheres, molecular surfaces, etc.) is a difficult task. As a further tool to interpret the local density distributions and help see the volume perturbations, our local density profile is accompanied by a visualization of the Z-scores of atomic local density (see Methods) and the Voronoi cells superimposed on the protein structure. This helps the user easily identify the regions and specific protons that fall outside the thresholds. Atomic local density Z-scores are represented by colored balls (see Methods for the color scale) and Voronoi cells are drawn in red. Both Z-score balls and Voronoi cells are selectable via intervals of score ranges. [Fig. 1(b)] is a subset of the visualization provided for the previously described 1BMY and 1IG6 structures, depicting the same isoleucine in both structures. The more extreme colored balls representing the atomic local density Z-scores show that there is both inflation and compression for the isoleucine in 1BMY. The purple balls represent protons that exist in the far left portion of the local density distribution for the 1BMY methyl protons in Figure 5, whereas the red ball lies in the far right of the distribution. The more moderately colored balls, tending toward green, for the same isoleucine in 1IG6 show a packing closer to the standard. This is in agreement with the tighter and better-shaped local density distribution for 1IG6 methyl protons in Figure 5. Using this visualization, we can see that the previously noted second peak in the local density distribution of the backbone amide helix protons of 1IG6 is attributable to consistent inflation of both the N- and C-terminal helices in most of the models in the ensemble (data not shown).

Refined NMR Structures

Recently a set of refinement procedures using explicit solvent has been proposed to generate higher-quality NMR structures.[43–47] To evaluate the effects of these refinement procedures, we measured the local density on DRESS16, 16 structure ensembles from the DRESS database[30] in their original and refined forms. The refinement procedure dramatically improves the Z-scores in most cases, sometimes by up to 13 standard deviations, suggesting a significant improvement in these NMR structures (Table III). In contrast, very few structures deteriorate in quality, and when deterioration occurs the scores only increase by

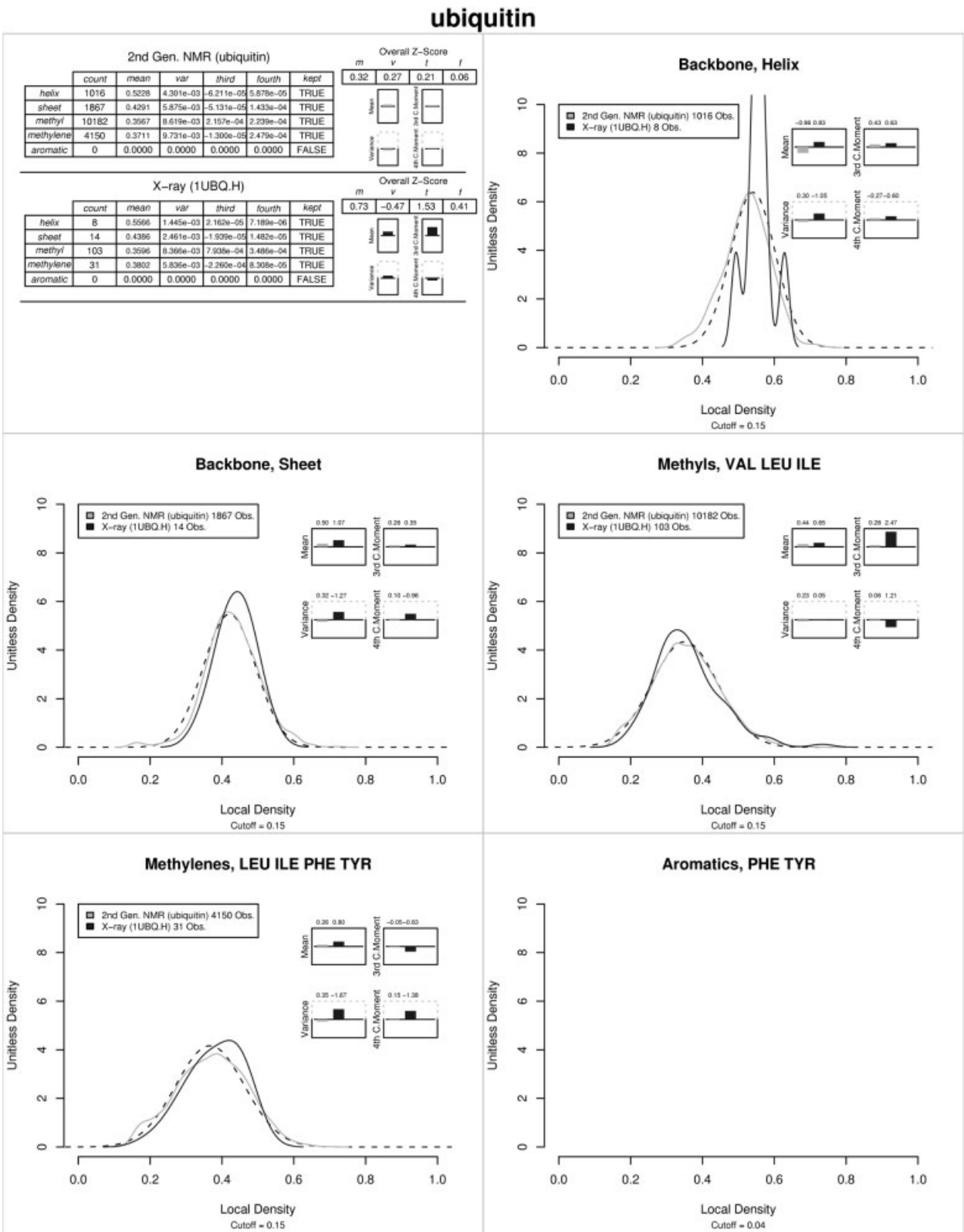

Fig. 6. Local density profiles of ubiquitin structure 1XQQ (128 model NMR ensemble) and 1UBQ (X-ray).

1 to 2 standard deviations at worst. In terms of moving within the three standard deviation threshold, the 3rd central moment improves the most upon refinement, with the Z-scores of 13 structures decreasing and moving within the threshold (data not shown), suggesting a significant resolution of atomic clashes. Improvement in Z-scores is

seen for backbone protons even though they already had the most reasonable Z-scores before refinement. For example, β-sheet protons, which tended to be slightly inflated in prerefined form, moved toward more moderate and slightly compressed values. As expected, side-chain protons showed the most significant improvements, sometimes as much as 10 standard deviations. Despite these improvements, however, we still observe certain packing deviations, particularly in the methyl protons where 13 of the 16 refined proteins still have variance Z-scores outside of the threshold. These results suggest that these water refinement procedures improve the quality of NMR structures, but further fine-tuning of the nonpolar interactions essential to core packing may be desirable.

We have also begun to investigate the benefits of other refinement procedures on the packing quality of NMR structures. For example, we have compared the NMR structures of several proteins as generated with or without refinement using RDC[48–50] data. When these restraints are included in the refinement, minor improvements are sometimes seen in the backbone, but no detectable improvement is seen in side-chain packing (data not shown). This is not surprising as almost all of the RDCs routinely used in the structural refinement are derived from protein backbone resonances.

Finally, we took advantage of the recent determination of a high-quality 128-structure ensemble of ubiquitin[31] (1XQQ) to further probe the possibility of inherent differences between structures produced by X-ray and NMR. In particular, as the purpose of the work by Lindorff-Larsen et al. was to couple NMR data and molecular dynamics simulations to capture the dynamics of a protein, our analysis of the local density distributions for this particular structure ensemble should provide insight into whether the dynamic nature of NMR causes a systematic perturbation in the local density distributions. As seen in the Z-scores, we find that the ubiquitin ensemble scores well, with local density distributions for overall and individual proton types essentially identical to the X-ray standards (Fig. 6). Although there is more variation in the overall Z-scores for the individual models of the ensemble, there are no models beyond the three standard deviation threshold in the overall mean Z-score (Supplementary Material 3). Most models actually score within a two standard deviation threshold. For comparison, we also show the distributions and scores for the 1.8 Å X-ray crystal structure of ubiquitin (1UBQ),[32] which are similar to those of the ensemble. Note that the helix protons of the X-ray structure have a splintered local density distribution because of the low number of samples. Although 1XQQ is only one example of this type of structure ensemble, the consistently good results versus our standard and compared against the corresponding X-ray structure indicate that packing standards derived from static X-ray structures are suitable for use as quality measures for dynamic NMR models. In addition, we note that 1XQQ is among the best NMR models we have analyzed, and postulate that using refinement protocols similar in style to Lindorff-Larsen et al. would perhaps improve the overall packing in NMR models.

CONCLUSION

In this work, we present the local density of protons as a useful and sensitive quality measure geared toward NMR structures. Standard local density distributions are calculated from Pack256, a set of high-resolution X-ray structures. We show that the local density distributions of different proton types are distinct and characterize them by four moments. These moments and distributions are used to establish both global Z-scores for scoring proteins and atomic level Z-scores of individual protons for visualization. The local density Z-scores are validated on X-ray datasets and applied to NMR datasets, and shown to be useful in describing both the global behavior of packing in the protein and the local behavior of packing for individual protons. These Z-scores are shown to correlate with both resolution and all-atom contact analysis ClashScores. For NMR structures, the local density shows widespread compression at low MPRMSD, moving toward a combination of inflation and compression at high MPRMSD. Analyses of water-refined NMR structures show the positive influence of the water refinement, resulting in significantly improved structures. Analysis of a high-quality ubiquitin ensemble refined against order parameters demonstrates that the dynamic nature of NMR does not cause a perturbation in the local density distributions. Therefore, using standards derived from static X-ray structures to score NMR ensembles seems to be reasonable. Toward a useful application for evaluating the local density for NMR structures, we have created a local density profile indicating "per-protein" Z-scores and local density distributions. Additionally, we generate a kinemage for visualizing the "per-atom" Z-scores superimposed on the protein structures. We provide a Web site at http://biogeometry.cs.duke.edu/research/localdensity/ for automated generation of these reports. We envision future incorporation of our local density method as a pseudo-potential during refinement of NMR structures, a technically challenging task because decisions need to be made on when and where to incorporate it into the refinement process.

ACKNOWLEDGMENTS

The authors thank Jane Richardson, David Richardson, and Jeremy Block for helpful discussions.

REFERENCES

1. Clore GM, Garrett DS. R-factor, free R, and cross-validation for dipolar coupling refinement of NMR structures. J Am Chem Soc 1999;121(39):9008–9012.
2. Cornilescu G, Marquardt JL, Ottiger M, Bax A. Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. J Am Chem Soc 1998;120(27):6836–6837.
3. Huang YJ, Powers R, Montelione GT. Protein NMR recall, precision, and F-measure scores (RPF scores): structure quality assessment measures based on information retrieval statistics. J Am Chem Soc 2005;127(6):1665–1674.
4. Nabuurs SB, Spronk CA, Krieger E, Maassen H, Vriend G, Vuister GW. Quantitative evaluation of experimental NMR restraints. J Am Chem Soc 2003;125(39):12026–12034.

5. Snyder DA, Montelione GT. Clustering algorithms for identifying core atom sets and for assessing the precision of protein structure ensembles. Proteins 2005;59(4):673–686.
6. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283–291.
7. Laskowski RA, Rullman JAC, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 1996;8(4):477–486.
8. Hooft RWW, Vriend G, Sander C, Abola EE. Errors in protein structures. Nature 1996;381(6580):272.
9. Lovell SC, Davis IW, Arendall WB 3rd, et al. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. Proteins 2003; 50(3):437–450.
10. Davis IW, Murray LW, Richardson JS, Richardson DC. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. Nucleic Acids Res 2004;32(Web Server issue):W615–W619.
11. Huang X, Powers R. Validity of using the radius of gyration as a restraint in NMR protein structure determination. J Am Chem Soc 2001;123(16):3834–3835.
12. Kuszewski J, Gronenborn AM, Clore GM. Improving the packing and accuracy of NMR structures with a pseudopotential for the radius of gyration. J Am Chem Soc 1999;121(10):2337–2338.
13. Richards FM. The interpretation of protein structures: total volume, group volume distributions and packing density. J Mol Biol 1974;82(1):1–14.
14. Voronoi G. Nouvelles applications des paramètres continus à la théorie des formes quadratiques, premier mémoire, sur quelques propriètès es formes quadratiques positives parfaites. J Reine Angew Math 1907;133:97–178.
15. Voronoi G. Nouvelles applications des paramètres continus à la théorie des formes quadratiques, deuxième mémoire, recherche sur les parallelloèdres primitifs. J Reine Angew Math 1908;134: 198–287.
16. Gerstein M, Tsai J, Levitt M. The volume of atoms on the protein surface: calculated from simulation, using Voronoi polyhedra. J Mol Biol 1995;249(5):955–966.
17. Lo Conte L, Chothia C, Janin J. The atomic structure of protein-protein recognition sites. J Mol Biol 1999;285(5):2177–2198.
18. Tsai J, Gerstein M. Calculations of protein volumes: sensitivity analysis and parameter database. Bioinformatics 2002;18(7):985–995.
19. Tsai J, Taylor R, Chothia C, Gerstein M. The packing density in proteins: standard radii and volumes. J Mol Biol 1999;290(1):253–266.
20. Pontius J, Richelle J, Wodak SJ. Deviations from standard atomic volumes as a quality measure for protein crystal structures. J Mol Biol 1996;264(1):121–136.
21. Word JM, Lovell SC, LaBean TH, et al. Visualizing and quantifying molecular goodness-of-fit: small-probe contact dots with explicit hydrogen atoms. J Mol Biol 1999;285(4):1711–1733.
22. Word JM, Lovell SC, Richardson JS, Richardson DC. Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. J Mol Biol 1999;285(4):1735–1747.
23. Hales TC. A proof of the Kepler conjecture. Ann Math 2005;162(3): 1063–1183.
24. Hales TC. Sphere packings, I. Discrete Comput Geometry 1997; 17(1):1–51.
25. Hales TC. Sphere packings, II. Discrete Comput Geometry 1997; 18(2):135–149.
26. Ban Y-EA, Edelsbrunner H, Rudolph J. Interface surfaces for protein-protein complexes. Proceedings of the eighth annual international conference on computational molecular biology. San Diego, CA: ACM Press; 2004.
27. Richardson DC, Richardson JS. http://kinemage.biochem.duke.edu.
28. Murzin AG, Brenner SE, Hubbard T, Chothia C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol 1995;247(4):536–540.
29. Berman HM, Westbrook J, Feng Z, et al. The protein data bank. Nucleic Acids Res 2000;28(1):235–242.
30. Nabuurs SB, Nederveen AJ, Vranken W, et al. DRESS: a database of REfined solution NMR structures. Proteins 2004;55(3):483–486.
31. Lindorff-Larsen K, Best RB, Depristo MA, Dobson CM, Vendruscolo M. Simultaneous determination of protein structure and dynamics. Nature 2005;433(7022):128–132.
32. Vijay-Kumar S, Bugg CE, Cook WJ. Structure of ubiquitin refined at 1.8 Å resolution. J Mol Biol 1987;194(3):531–544.
33. Frishman D, Argos P. Knowledge-based protein secondary structure assignment. Proteins 1995;23(4):566–579.
34. Besl PJ, McKay HD. A method for registration of 3-D shapes. IEEE Trans Pattern Anal Machine Intell 1992;14(2):239–256.
35. R: A language and environment for statistical computing. Vienna, Austria; 2005.
36. Wand MP, Ripley BD. KernSmooth 2.22, R package; 2004.
37. Cramér H. Mathematical methods of statistics. Princeton: Princeton University Press; 1946.
38. Richardson DC, Richardson JS. The kinemage: a tool for scientific communication. Protein Sci 1992;1(1):3–9.
39. Davis IW. KiNG. http://kinemage.biochem.duke.edu.
40. Clore GM, Gronenborn AM. New methods of structure refinement for macromolecular structure determination by NMR. Proc Natl Acad Sci USA 1998;95(11):5891–5898.
41. Lin D, Tsui V, Case D, Yuan Y-C, Chen Y. Human Mrf-2 Domain, NMR, 11 Structures, PDB code 1IG6, to be published. 2001.
42. Yuan Y-C, Whitson RH, Liu Q, Itakura K, Chen Y. A novel DNA-binding motif shares structural homology to DNA replication and repair nucleases and polymerases. Nat Struct Biol 1998;5(11):959–964.
43. Linge JP, Habek M, Rieping W, Nilges M. ARIA: automated NOE assignment and NMR structure calculation. Bioinformatics 2003; 19(2):315–316.
44. Linge JP, Nilges M. Influence of non-bonded parameters on the quality of NMR structures: a new force field for NMR structure calculation. J Biomol NMR 1999;13(1):51–59.
45. Linge JP, Williams MA, Spronk CA, Bonvin AM, Nilges M. Refinement of protein structures in explicit solvent. Proteins 2003;50(3):496–506.
46. Spronk CA, Linge JP, Hilbers CW, Vuister GW. Improving the quality of protein structures derived by NMR spectroscopy. J Biomol NMR 2002;22(3):281–289.
47. Xia B, Tsui V, Case DA, Dyson HJ, Wright PE. Comparison of protein solution structures refined by molecular dynamics simulation in vacuum, with a generalized Born model, and with explicit water. J Biomol NMR 2002;22(4):317–331.
48. Bax A. Weak alignment offers new NMR opportunities to study protein structure and dynamics. Protein Sci 2003;12(1):1–16.
49. Lipsitz RS, Tjandra N. Residual dipolar couplings in NMR structure analysis. Ann Rev Biophys Biomol Struct 2004;33:387–413.
50. Prestegard JH, Bougault CM, Kishore AI. Residual dipolar couplings in structure determination of biomolecules. Chem Rev 2004;104(8):3519–3540.