J Biomol NMR (2013) 56:95–112
DOI 10.1007/s10858-013-9725-y

ARTICLE

Nucleic acid helix structure determination from NMR proton chemical shifts

Ramon M. van der Werf · Marco Tessari · Sybren S. Wijmenga

Received: 1 February 2013 / Accepted: 27 March 2013 / Published online: 6 April 2013

Abstract We present a method for *de novo* derivation of the three-dimensional helix structure of nucleic acids using non-exchangeable proton chemical shifts as sole source of experimental restraints. The method is called *che*mical shift de nov*o* structure derivation *p*rotocol employing *s*ingular value decomposition (CHEOPS) and uses iterative singular value decomposition to optimize the structure in helix parameter space. The correct performance of CHEOPS and its range of application are established via an extensive set of structure derivations using either simulated or experimental chemical shifts as input. The simulated input data are used to assess in a defined manner the effect of errors or limitations in the input data on the derived structures. We find that the RNA helix parameters can be determined with high accuracy. We finally demonstrate via three deposited RNA structures that experimental proton chemical shifts suffice to derive RNA helix structures with high precision and accuracy. CHEOPS provides, subject to further development, new directions for high-resolution NMR structure determination of nucleic acids.

Keywords Proton chemical shift · Helix parameters · Structure · RNA · NMR

Software Available: NUCHEMICS and the python script, CHEOPS, are available upon request. 3DNA can be obtained via the website: http://adna.rutgers.edu.

Electronic supplementary material The online version of this article (doi:10.1007/s10858-013-9725-y) contains supplementary material, which is available to authorized users.

R. M. van der Werf · M. Tessari · S. S. Wijmenga (✉)
Department of Biophysical Chemistry, Institute of Molecules and Materials, Radboud University Nijmegen, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands
e-mail: S.Wijmenga@science.ru.nl

Introduction

Structure determination of nucleic acids from NMR data has been an intrinsically difficult problem for many years (Allain and Varani 1997; Gronenborn and Clore 1989; Kuszewski et al. 2001; Pardi et al. 1988; Varani et al. 1996; Wijmenga and van Buuren 1998). This is mainly due to the extended nature of helices, which conflicts with the short-range structure information contained in common NMR parameters, like NOEs and J-couplings, while severe spectral overlap combined with low proton densities (Mooren et al. 1994; Wijmenga et al. 1994; Wijmenga et al. 1995; Wijmenga and van Buuren 1998), spin-diffusion (Borgias and James 1990; Keepers and James 1984), and dynamics (Bailor et al. 2007; Shajani and Varani 2007) add to these difficulties. The introduction of heteronuclear residual dipolar couplings (RDCs) has helped to alleviate these problems, in part, thanks to the global structure information contained in these parameters (Mollova and Pardi 2000; Sibille et al. 2001). Unfortunately, commonly used heteronuclear RDCs only carry orientation but no translational information (Musselman et al. 2006). As a result of this incompleteness of NMR experimental restraints, NMR structures may still contain inaccuracies, e.g. in helical rise (Dickerson 1989; Olson et al. 2001) and non-bonded terms in the force field can significantly affect the derived structures (Kuszewski et al. 2001; Tolbert et al. 2010). Introduction of other more complete and structure-sensitive restraints is thus of relevance.

Chemical shifts are exquisitely sensitive probes of (bio) molecular structure. They provide detailed information on structure and electronic properties in solution, non-crystalline, and crystalline states (Cavalli et al. 2007; Cromsigt et al. 2001; Grzesiek and Sass 2009; Mielke and Krishnan 2009; Shen et al. 2008, 2009; Wijmenga et al. 1997;

Wishart et al. 2008). For proteins, chemical shifts are commonly employed to establish secondary structure and backbone torsion angles (Cornilescu et al. 1999; Mielke and Krishnan 2009). Recently, it has even been demonstrated that the three-dimensional structure of a protein can be derived, using chemical shifts as sole source of experimental restraints (Cavalli et al. 2007; Grzesiek and Sass 2009; Shen et al. 2008, 2009; Wishart et al. 2008). These methods make use of molecular replacement, which in turn is based on sequence similarity and empirical structure-chemical shift relationships established from analyses of large databases of protein structures and their related chemical shifts. For nucleic acids, chemical shifts have until now been employed less extensively to derive structure information. This may in part be due to the fact that precise and accurate structure-chemical-shift relationships for heteronuclear shifts are limited for nucleic acids (Lam and Chi 2010; Mielke and Krishnan 2009). However, on the positive side, in nucleic acids well-parameterized ring-current and magnetic-anisotropy effects govern proton conformational chemical shifts and via the associated semi-classical equations they can precisely and accurately be predicted given a three-dimensional structure (Cromsigt et al. 2001; Dejaegere et al. 1999; Wijmenga et al. 1997). For instance, the RMSD between predicted and observed proton chemical shifts in well-defined rigid RNA helices is found to be as small as 0.08 ppm (Cromsigt et al. 2001). This value approaches the prediction precision obtained from purely empirical relationships (Altona et al. 2000; Lam and Chi 2010; Mielke and Krishnan 2009). In contrast to heteronuclear RDCs, the semi-classical equations for ring current and magnetic anisotropy effects contain information on both translational and orientational helix parameters. Together with their high sensitivity to structure parameters, evident from the semi-classical equations, these characteristics suggest that proton chemical shift restraints may alleviate at least part of the problems associated with NMR RNA/DNA structure calculations.

Here we demonstrate how proton chemical shifts can be used to *de-novo* derive the NMR structures of nucleic acid helices at high resolution without the requirement of any additional experimental parameters in a number of instances. The method we employ is called CHEOPS (*che*mical shift de nov*o* structure derivation *p*rotocol employing *s*ingular value decomposition and backbone restrained molecular dynamics). Simulated input data is used to assess in a defined manner the effect of errors or limitations in input data on the derived structures, while the structure calculations on three published RNA structures employing experimental input chemical shifts demonstrate that CHEOPS also works in practice. It should be noted that CHEOPS can be extended to include also other structure restraints (e.g. J-couplings, RDC, NOEs etc.).

Materials and methods

Allowed helix space and 'standard' helix in terms of helix parameters

The helix parameters were analyzed using 3DNA v1.5 (Lu and Olson 2003) in 20 RNA helix structures solved by X-ray diffraction to a resolution better than 3 Å (157D, 1DQH, 1I9X, 1KFO, 1RNA, 1RXA, 1RXB, 1SDR, 1ZX7, 205D, 255D, 259D, 353D, 377D, 402D, 405D, 413D, 433D, 438D, 472D). Table S1 lists the average helix parameters and their standard deviations ($\sigma_{helixpar}$). The range of $\pm 3\sigma_{helixpar}$ around the average A-helix parameter value defines the allowed A-helix space. The allowed B-helix space (not used here) can likewise be derived from the values given in Table S1. It should be noted that by using a range of $\pm 3\sigma_{helixpar}$ the A-helix and B-helix spaces overlap in large part. The average helix parameters of the standard B-helix essentially reside within the A-helix space thus defined.

Evaluating CHEOPS using simulated chemical shifts

Standard A-helix as target structure

A standard fiber model A-helix (Olson et al. 2001), built using 3DNA (Lu and Olson 2003), was taken as target and called the *target structure*. The standard fiber-model A-helix parameters of Table S3 were used to build the target structure. In most cases, the *target structure* consisted of six canonical base pairs with primary sequence 5′-ACAGCU-3′:5′-AGCUGU-3′. In this sequence, essentially all possible combinations of base pair steps are present. A set of 93 ^{1}H chemical shifts, i.e. H1′, H2′, H3′, H4′, H5′, H5″, H2, H5, H6, and H8, was calculated with NUCHEMICS (Wijmenga et al. 1997) using the Giessner-Prettre parameters set (GP set)(Giessner-Prettre and Pullman 1987; Ribas-Prado and Giessner-Prettre 1981) from this *target structure*. These simulated chemical shifts were considered as 'observed shifts' in the test structure derivations using CHEOPS (see "Results", CHEOPS structure calculation protocol). In these test calculations, chemical shifts were calculated by means of NUCHEMICS using the GP parameter set, except otherwise stated. In the final test calculations presented here, as starting structures, 1,000 helices were generated with each helix parameter randomly selected in the interval $[-3\sigma_{helixpar}, 3\sigma_{helixpar}]$. Initially, also some test calculations were performed with a broader set of test helices, but they gave similar results (see "Results"). Further, the threshold (parameter λ) for zeroing small eigenvalues of the G-matrix was set to a value of 0.1 %. As discussed in supplementary material, λ cannot be set to zero because then the system-imprecision would

introduce large fluctuations into the solution **ΔH** vector. The parameter κ, that scales the solution vector, was set to 0.50. The maximum number of cycles in the iterative SVD was set to 150. Four main sets of test calculations were carried out with the convergence tolerance set to 0.05, 0.10, 0.15, and 0.30 ppm, respectively. Note that given the helix parameters the atomic structure model can be built (via 3DNA, see below). In each step of the SVD cycle an atomic structure is built from the current helix parameters so that a new set of chemical shifts can be calculated via NUCHEMICS and subsequently a new set of improved helix parameters be derived in a subsequent steps of the SVD cycle (see Fig. 2, "Results"). The final structures, thus obtained, are called converged SVD structures.

The converged SVD structures were checked for holonomic (Gelbin et al. 1996) and standard hydrogen bond distance (Arnott and Hukins 1973) violations in Xplor v3.851 (Brunger 1992). Structures with violations were not considered further. Because the converged SVD structures calculated with a complete set of chemical shifts rarely showed such violations, no 'backbone regularization' (see supplementary material) was applied to the converged SVD structures. The structures that pass all tests are called accepted structures.

Each accepted structure (and/or converged SVD structure) was compared to the *target structure* in terms of helix parameters as well as atomic coordinates, the latter via an all-atom RMSD (see below). The helix parameters provide a detailed description of each conformation element, i.e. 12 helix parameters fully define a base pair step conformation; the helix parameters can be analyzed, i.e. compared with their values in the target structure to obtain a detailed view of similarity/difference. The all-atom RMSD provides an overall and thus less detailed measure of (a) similarity within the ensemble of final structure(s) and (b) similarity of the ensemble of final structures with the target. The pairwise RMSD used ($RMSD_{pairwise}$) was the average all-atom pairwise RMSD of the atomic coordinates of the ensemble of accepted structures (and or converged SVD structures); it is a measure of the precision of the protocol as it gives the width of the bundle of accepted structures. A more detailed indication of the precision is given by the RMSD of the helix parameters. In addition, the average accepted structure (and/or converged SVD structure) was built using 3DNA based on the average helix parameters of the entire ensemble. The RMSD of this average structure to the *target structure*, $RMSD_{av2target}$, is a measure of the accuracy of the protocol. The deviations of the helix parameters of this average accepted structure (equal to the average helix parameters) from the target structure give a more detailed view of the accuracy.

To estimate the effect of uncertainty in parameterization of the ring current and magnetic anisotropy on the chemical shift derived structures the foregoing test calculations were repeated, using the more recent parameter set derived by Case (Case 1995) (DC set; see supplementary material for details).

Effect of the average number of chemical shift restraints per base pair The effect of a smaller number of chemical shifts restraints on the precision and accuracy of the derived structures was tested on the derivation of a A-helix (see "Materials and methods" for definition) of 10 canonical base pairs and primary sequence 5′-ACGUGCGUAC-3′:5′-UGCACGCAU-3′. This sequence was again chosen to contain all possible base pair steps. For this helix, a set of 154 chemical shifts of non-exchangeable protons was calculated with NUCHEMICS (Cromsigt et al. 2001; Wijmenga et al. 1997), using the GP-parameter set and taken as the 'observed' shifts in the calculations using simulated chemical shifts. In addition, from this set, 300 new sets of 'observed' shifts were generated as follows. First, 200 sets were obtained by randomly removing 4, 14, 24, 34, 44, 54, 64, 74 chemical shifts with 25 subsets per deletion. For each of these 200 sets, the structure derivation was performed starting with 20 random structures ($\pm 3\sigma_{helixpar}$). This corresponds to 500 structure calculations per random deletion subset (4,000 in total). The remaining 100 sets of chemical shifts were generated by randomly removing 84, 94, 104, 114, 124 chemical shifts with 20 subsets per deletion. For each of these 100 sets, the structure derivation was run starting with 250 random A-helices. This corresponds to 5,000 structure calculations per random deletion subset. In addition, a set of helices was calculated employing all 154 'observed' chemical shifts, starting from 500 random A-helices. In addition, three separate sets of experimentally 'easily' accessible chemical shifts were generated: a complete list of aromatic proton shifts complemented with H1′/2′ shifts (set 1) or H1′/4′ shifts (set 2) or H1′ shifts (set 3). These calculations started from 500 random A-helices. The parameters λ and κ were set as before and the chemical shift tolerance to 0.10 ppm. The converged SVD structures were checked for hydrogen-bond violations in the basepairs and subsequently backbone regularized (see supplementary material).

The accepted structures were analyzed as before. In addition, the quality of the accepted structures was assessed by means of a Validation Chemical Shift Q-factor (VCSQ-factor) defined as:

$$VCSQ = \sqrt{\frac{\sum_{i=1}^{N}\left(\delta_{conf,calc,i} - \delta_{conf,obs,i}\right)^2}{\sum_{j=1}^{N}\delta_{conf,obs,j}^2}} \quad (1)$$

Here, N equals the total number of non-exchangeable proton chemical shifts. The VCSQ-factor is a measure of how well the resulting structure explains the complete set of proton chemical shifts.

Non A-helix structures as target

In the above, we employed an ideal standard A-helix as target structure. To investigate how CHEOPS performs for targets with non-standard A-helix conformation we chose physically possible helical target structures that contain basepairs and basepair steps with that cover a range of conformations significantly deviating from that of a standard A-helix (see "Results" section for details). We performed the structure calculations in fashion similar to the one above. We started from 500 structures generated from helix parameters randomly spread around standard A-helix values in a range of $\pm 3\sigma_{\text{helixpar}}$ or larger when needed. The parameters λ and κ were set as before and the chemical shift tolerance to 0.10 ppm. We employed as experimental restraints the full set of non-exchangeable ^{1}H chemical shifts calculated using the GP-set from the target. No backbone regularization was performed.

Evaluating CHEOPS on nucleic acids of known structure using experimental chemical shifts

Three NMR RNA hairpin structures with canonical basepairs in the A-helix stem were retrieved from the PDB (http://www.pdb.org) and their matching proton chemical shifts from the BMRB (http://www.bmrb.wisc.edu): 1PJY (BMRB-entry:5834) and 2JTP (BMRB-entry:15417), and 1QES (BMRB-entry: 37865). The CHEOPS protocol used for simulated input data was employed, except for the following changes. (1) The input chemical shifts are now the real experimental chemical shifts. They were checked for outliers or incorrect calibrations by running the iterative SVD protocol using a small number ($\sim$5) of randomized starting helices and parameter settings as described below for the full protocol. Experimental chemical shifts that consistently deviated more than the chemical shift tolerance from the calculated values were provisionally marked as outliers. They were only removed when an acceptable reason for removal did exist (see "Results" for details). (2) The structure calculations started from a set of 500 random starting helix structures, generated within A-helix space with a variation of $\pm 3\sigma_{\text{helixpar}}$ around the standard A-helix values. (3) The parameter settings for the iterative SVD were adjusted compared to the CHEOPS protocol for simulation data. The λ-value was set to 4.25 %, a value optimal for providing stable SVD solutions given the error in the experimental chemical shift of a few percent (measurement error plus estimated back-calculation error). The scaling factor was set to 0.5 and the maximum number of cycles in the iterative SVD to 150. For UG_RNA (1QES) the helix parameters shear, stretch and opening of the UG base pairs were fixed to the values found in the pdb (Table S16c).

Results and discussion

CHEOPS structure calculation protocol

The formal basis for CHEOPS, the proposed method for nucleic acid helix structure calculation from chemical shifts, is presented below. We employ the helix parameter description (Dickerson 1989) to define the nucleic acid conformation. Helix parameters fully define nucleic acid conformation (12 per basepair step) and provide a quantitative framework for describing the conformation of nucleic acids, and thus for understanding the, sometimes subtle, variations in their three-dimensional structures (Olson et al. 2001; Olson et al. 2009). Moreover, three-dimensional atomic models can be reconstructed given a set of helix parameters (Lu and Olson 2003). Furthermore, the reduced space spanned by the helix parameters (12 per basepair) allows one to focus on the essential conformational features rather than having to consider all degrees of freedom, as would be the case when nucleic acid conformation is described in terms of atomic coordinates. In the CHEOPS structure derivation, we use the program 3DNA and its up-to-date helix parameter analysis and atomic-model rebuilding facilities (Lu and Olson 2003), while the proton chemical shifts are calculated given a three-dimensional structure using NUCHEMICS (Cromsigt et al. 2001; Wijmenga et al. 1997) and are thus based on ring current and magnetic anisotropy as discussed below.

The conformational chemical shift of a proton k in a helix ($\delta_{conf,k}$) can be expressed as a function g of the complete set of helix parameters that describe the (helix) structure,

$$\delta_{conf,k} = g\big(S_{x,1}, S_{y,1}, S_{z,1}, \kappa_1, \omega_1, \sigma_1, dx_1, dy_1, D_{z,1}, \eta_1, \theta_1, \Omega_1, \ldots, S_{x,n}, S_{y,n}, S_{z,n}, \kappa_n, \omega_n, \sigma_n, dx_n, dy_n, D_{z,n}, \eta_n, \theta_n, \Omega_n\big) \quad (2)$$

The change in proton chemical shift, $\Delta\delta_{conf,k}$, from its current value $\delta^o_{conf,k}$ resulting from a change of one or more helix parameters can be calculated by a Taylor expansion of the function g. For small changes such a series can be truncated to first order leading to Eq. (3).

$$\Delta\delta_{conf,k} \approx \sum_{basepair} \sum_{helixpar} \left(\frac{\partial \delta_{conf,k}}{\partial H_{basepair,helixpar}} \right)_{H=H^o_{basepair,helixpar}} \times \Delta H_{basepair,helixpar} \quad (3)$$

Here, $\Delta H_{basepair,helixpar}$ is the change in helix parameter $H_{basepair,helixpar}$ from its current value $H^o_{basepair,helixpar}$. The partial derivatives in Eq. (3) can be evaluated numerically by reconstructing via 3DNA an atomic model from the helix parameters and employing NUCHEMICS (Cromsigt et al. 2001; Wijmenga et al. 1997) and thus the semi-classical

equations for the ring current and magnetic anisotropy effect that govern the proton conformational chemical shifts of nucleic acids. In compact notation, Eq. (4) provides an expression for the change in proton chemical shifts upon a (small) change in the initial helix structure:

$$\Delta\mathbf{C} = \mathbf{G} \cdot \Delta\mathbf{H} \tag{4}$$

Here, $\Delta\mathbf{C}$ (= $\mathbf{C–C_o}$) and $\Delta\mathbf{H}$ (= $\mathbf{H–H_o}$) are column vectors containing the changes in conformational proton chemical shifts and helix parameters, respectively, while $\mathbf{G}$ is the matrix whose elements contain the partial derivatives of the chemical shift versus helix parameters and is evaluated numerically (see above and supplementary material). In Eq. (4), the helix parameters in $\Delta\mathbf{H}$ and $\mathbf{G}$ are expressed in units of σ_{helixpar} (Table S1). The position of the flanking residues (i − 1) and (i + 1) and to a smaller degree (i − 2) and (i + 2) affect the proton chemical shifts of residue i. Consequently, the G-matrix has a block-diagonal form. The aim is to solve Eq. (4) given a list of chemical shift differences ($\Delta\mathbf{C}$):

$$\Delta\mathbf{H} = \mathbf{G}^{-1} \cdot \Delta\mathbf{C} \tag{5}$$

When $\Delta\mathbf{C}$ is the difference between calculated and observed shifts, $\Delta\mathbf{H}$ contains changes required in helix parameters to obtain the structure consistent with observed chemical shifts. The required inverse of the G-matrix in Eq. (5) can be obtained via singular value decomposition (SVD). The new helix parameters, $\mathbf{H_{new}}$, are subsequently calculated from $\Delta\mathbf{H}$,

$$\mathbf{H}_{new} = \mathbf{H_0} + \kappa \cdot \Delta\mathbf{H} \tag{6}$$

Here, $\mathbf{H_0}$ is the vector that contains the previous helix parameters. The factor κ is a user chosen parameter (default value = 0.50). Because proton chemical shifts and helix parameters are often close to but not completely linearly related (Fig. 1a–d), an iterative procedure is therefore followed to find the final solution. A small value for the factor κ assures that, in the iterative procedure, subsequent solution vectors $\Delta\mathbf{H}$ follow the potential curvature of the helix parameter-chemical shift space. In each step (cycle) in the iterative procedure an atomic model of the new helix is rebuild from $\mathbf{H_{new}}$ using 3DNA (Lu and Olson 2003), from which new chemical shifts are calculated using NUCHEMICS (Cromsigt et al. 2001; Wijmenga et al. 1997). A new cycle can then start. The scheme continues until calculated and observed shifts are equal within a chemical shift tolerance set by the user. The obtained structures are called converged SVD structures. As a final step, the converged SVD structures undergo backbone regularization, when needed. For this, a short restrained molecular dynamics (rMD) run is executed in Xplor (Brunger 1992), during which the aromatic planes are kept fixed in Cartesian space and loose standard A-helix distance and/or sugar-backbone torsion angle restraints are employed. Moreover, the sugar conformation is thus kept in its N-puckered state. In this way, the global structure and thus helix parameters remain unchanged during backbone regularization (see supplementary material for further details). The above forms the basis for the calculation of a nucleic acid helix structure from chemical shifts presented here. Figure 2 shows the complete protocol of CHEOPS.

A few aspects should be stressed. (a) The CHEOPS protocol starts from a set of 1,000 random starting structures that well cover allowed A-helix conformational space: a space defined by helices that are within $\pm 3\sigma_{\text{helixpar}}$ (Table S1) from average A-helix parameter values, as follows from an analysis of a representative set of helices. It should be noted that by using a range of $\pm 3\sigma_{\text{helixpar}}$ the A-helix and B-helix spaces overlap in large part. For instance, the helix parameters of the standard B-helix essentially all reside within the A-helix space thus defined. The B-helix deviates significantly from an A-helix in the helix parameters only in h-rise, inclination, X-dis, and h-twist (Table S1). The σ_{helixpar} values of these helix parameters are such that with a space of $\pm 3\sigma_{\text{helixpar}}$ away from mean A-helix values, also conformations will be sampled that have B-helix like basepair step conformations. Thus, the mean helix parameters with B-helix nature all fall well within $\pm 3\sigma_{\text{helixpar}}$ from their standard A-helix values, except for the mean X-dis; the mean X-dis of a B-helix falls on the edge, although many of the X-dis B-helix values fall within $\pm 3\sigma_{\text{helixpar}}$. Other types of helices, e.g. B′- and A′-helices, are also part of the conformational space that is sampled. Thus, the sampled space includes significant deviations from standard values, deviations that could be caused by interactions with proteins or ions. (b) In addition, the sequence of the 6 basepair helix is chosen to include essentially all possible basepair steps combinations. In this way, the tests indicate for all possible basepair steps how well proton chemical shifts define the helix starting from a large number of different positions in helix space, a space that encompasses different types of helices. (c) Furthermore, by starting from a large number of positions in helix space, no bias is introduced in the set of final accepted structures. There are two situations when bias may otherwise arise. (1) If the number chemical shifts is smaller than the number of helix parameters the SVD solution may depend on the starting structure (see supplementary material for a more extensive description of this point). On the other hand, if the number of chemical shifts is larger than the number of helix parameters, the SVD solution forms a least-squares solution and does not depend on the starting structure. (2) Parabolic dependences between chemical shifts and helix parameter (Fig. 1e, f) might potentially introduce a bias in the set of converged of

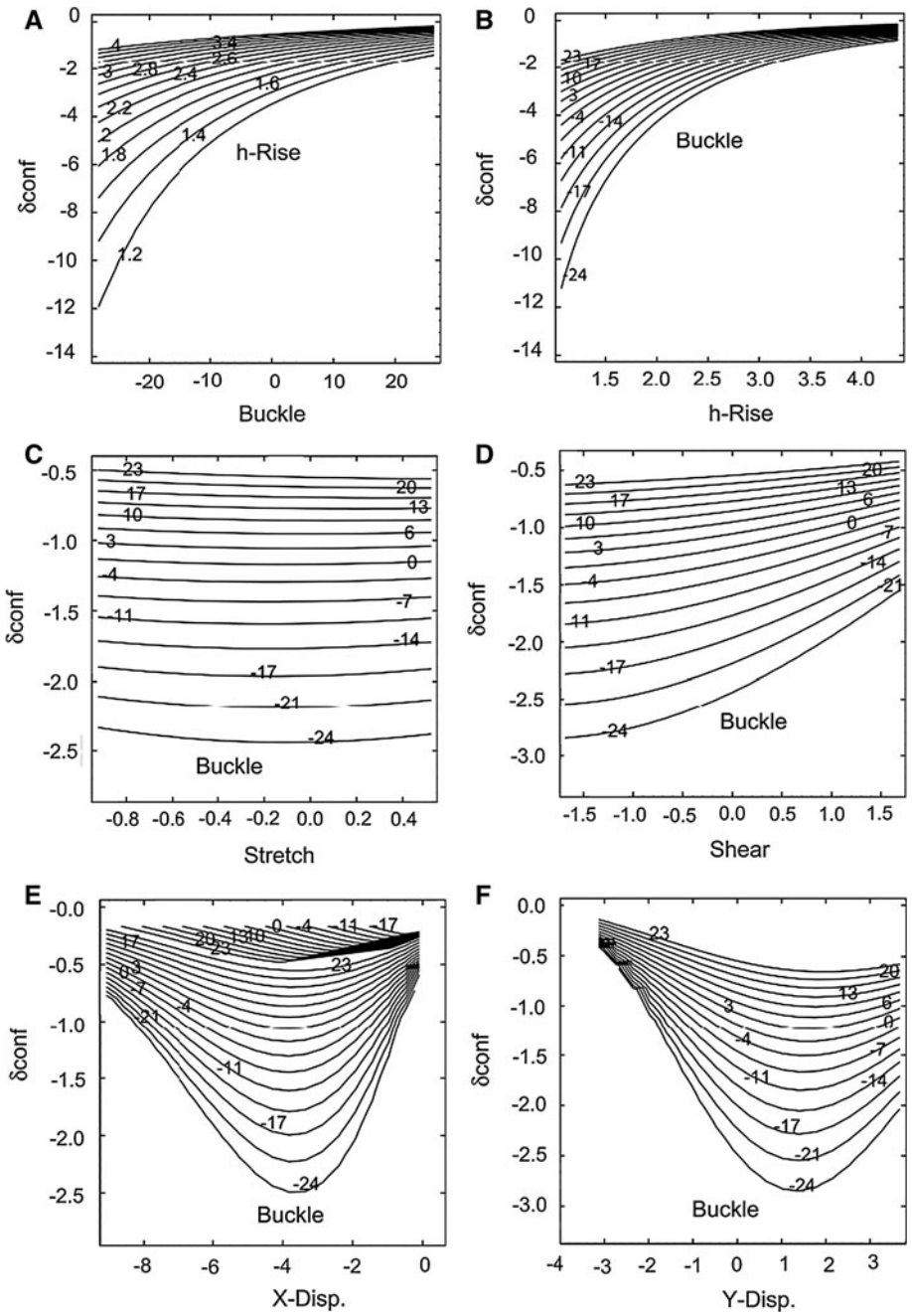


Fig. 1 Conformational chemical shift of the H5 proton, δ_{confH5}, of residue C4 in a standard anti-parallel A-helix with 6 base pairs and primary sequence 5′-GGACGA-3′:5′-UCGUCC-3′, as function of the helix parameters of basepair 4 (C4:G9). The helix parameters of basepair 4 were varied in the interval $[-4\sigma_{helixpar}, 4\sigma_{helixpar}]$, while keeping the other values fixed at the standard A-helix values (Table S1). For each combination of helix parameters an atomic model was built with the rebuilding feature in 3DNA (Lu and Olson 2003) and its proton chemical shifts calculated with NUCHEMICS (Wijmenga et al. 1997). The *number* in the *contour lines* indicates the value of the second helix parameter that is varied. **a**, **b** Typical example of the usual monotonic dependence of chemical shift upon change in helix parameter is shown; **a** δ_{confH5} versus Buckle (h-rise); **b** δ_{confH5} versus h-Rise (Buckle). **c**, **d** Example of helix parameters, that barely affect the chemical shift; **c** δ_{confH5} versus Stretch (Buckle); **d** δ_{confH5} versus Shear (Buckle). **e**, **f** Example is shown of helix parameters with a parabolic dependence on chemical shift; **e** δ_{confH5} versus X-disp (Buckle); **f** δ_{confH5} versus Y-disp (Buckle). A complete analysis using the partial derivatives of the proton chemical shift with respect to the helix parameters shows that the sensitivity of chemical shifts to structural changes follows the trend: H5 ≈ H2 > H6 ≈ H2′ ≈ H8 > H1′ ≈ H3′ > H5′ ≈ H4′ > H5″. Alternatively, the sensitivity of helix parameters to changes in the chemical shift is given by: *buckle* > *Y-displacement* ≈ *tip* > *propeller* *twist* > *inclination* ≈ *rise* > *helix-twist* ≈ *X-displacement* > *stagger* > *opening* > *shear* > *stretch*. Additional plots with other helix parameters combinations are shown Figure S1

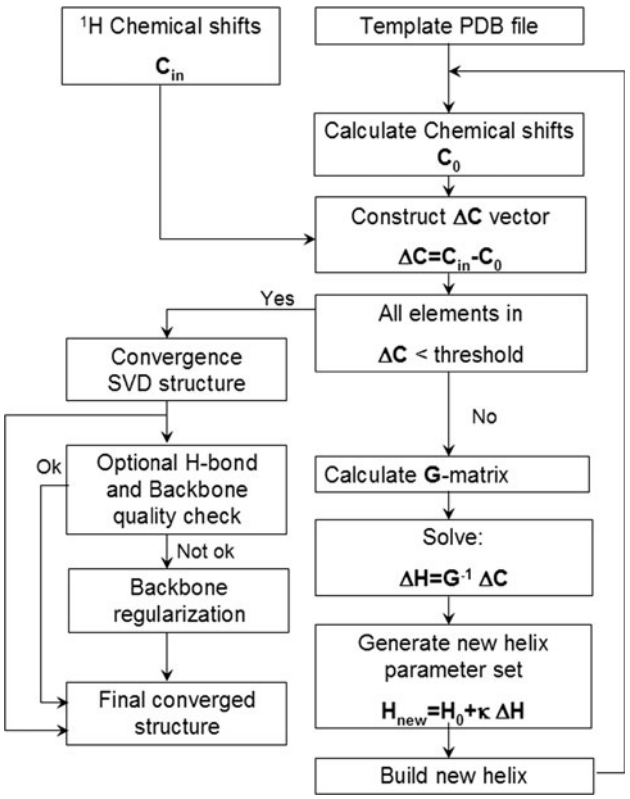


Fig. 2 Flowchart of the protocol for the derivation of helical nucleic acid structures from chemical shifts. The template PDB file is read and chemical shifts are calculated from this structure. The calculated chemical shifts are subtracted from the restraint chemical shifts, thereby composing vector **ΔC**. Subsequently, convergence is judged, i.e. are all chemical shifts within a user-set threshold limit? If 'Yes', the converged SVD structure is checked from compliance backbone and h-bond parameters, its backbone may be regularized when required by means of a brief restrained molecular dynamics runs during which the bases are kept fixed and subsequently the final converged structure is obtained. If 'No', the **ΔH** vector is solved from the G-matrix and the chemical shift deviation (ΔC) is solved and a new set of helix parameters calculated $\mathbf{H_{new}}$. The **ΔH** vector is scaled by κ in this calculation, to prevent too large changes and thereby restraining the solution to the regime, where the approximation of a linear relationship between chemical shifts and helix parameters holds. From the new helix parameter set, H_{new}, a new atomic model is built, and the calculation enters its new cycle

SVD solutions, because the iterative SVD might not be able to cross the minimum of such curves. However, the set of random starting structures would sample both sides of the *X/Y-displacement* curve as shown in Fig. 1e, f, thereby removing any bias. However, we find that our iterative calculation protocol crosses the minimum seen for individual chemical shifts like H5 (see Figure S2). In other words, this issue does not appear to pose a problem. The likely reason is that each helix parameter depends on a number of chemical shifts of which most show a monotonic and approximately linear dependence.

Evaluating CHEOPS using simulated chemical shifts

Standard A-helix as target structure

The performance of CHEOPS with regard to precision and accuracy, parameterization of ring current and magnetic anisotropy equations, and number of chemical shifts was evaluated as described below. To assess these aspects in a defined manner for essentially all A-helices, simulated chemical shifts were employed and A-helices that cover largely all combinations of canonical basepairs. The target structure employed was in the fiber model standard A-helix conformation (Table S3). Note further that in the structure calculations, the starting helix structures cover a wide range of helix space, a range that covers not only A-helix space but also includes helices with B-helix type conformations (see "Materials and methods"). In addition, A-helices (and B-helices) are well established and defined. Thus, their helix parameter and variances provide a reliable context for testing the validity and accuracy of the chemical shift derived structures. Here we present the main results, while supplementary material contains those of additional tests.

Assessment of precision and accuracy The CHEOPS protocol was run as described in "Materials and methods", starting from 1,000 random structures that cover allowed A-helix space ($\pm 3\sigma_{helixpar}$; Fig. 3, left) with a chemical shift tolerance ranging from 0.05 to 0.30 ppm (Table 1) and employing 15 chemical shifts per basepair. Note that this starting helix space covers a wide range of helices including B-helix type conformations (see above). We consider the smallest chemical shift tolerance first. Out of the 1,000 starting structures, 956 structures converged; they could all be accepted, because none showed extended or compressed hydrogen bonds and thus no backbone regularization was needed. In the narrow bundle of overlaid accepted structures the basepairs can clearly be distinguished (Fig. 3a, right). The $RMSD_{pairwise}$ of this accepted set is quite small (Table 1), indicating that the chemical shifts indeed define the structures with high precision. A more detailed view is obtained from an analysis of the helix parameters of the accepted set (Table S5/S6). First, we note that the RMSDs of the helix parameters of this set range from 0.3 to $1.1\sigma_{helixpar}$ (Table S2/S6) with a mean of $0.55\sigma_{helixpar}$ (Table 1 and Table S2/S6). Expressed in terms of their normal units the translational parameters and rotational parameters are then defined on average within 0.6 Å and 6°, respectively. In other words, the helix parameters are defined to high precision. The variation seen between the RMSDs of the different helix parameters nicely follows the expected trend of how sensitive a helix parameter is to a change in proton chemical shift

(Figure S3/S4; Table S2). In addition, the average converged structure can be built from the mean values of the ensemble averaged helix parameters using 3DNA. Note that in this way an average structure is built with correct bond distances and angles. The $RMSD_{av2target}$ value (Table 1) of this average converged structure versus *target structure* is smaller than the $RMSD_{pairwise}$. Thus, the *target structure* lies well within the bundle of converged structures, which shows that no bias is introduced. This accuracy is also evident from the helix parameter of the average structure being much closer to the target values than the RMSD of the helix parameters of the ensemble of accepted structures (Table S5/S6).

The RMSD of the helix parameters and the $RMSD_{pairwise}$ increase each linearly with the chemical shift tolerance (Table 1, Figure S5/6; Table S5–S12). However, at the highest tolerance (0.30 ppm) the precision remains acceptable (mean helix parameter RMSD, $1.23\sigma_{helixpar}$; RMS $D_{pairwise}$, 1.13 Å, Table 1). The increase in $RMSD_{av2target}$ is quite small (from 0.33 Å to 0.44 Å), so that $RMSD_{av2target} \ll RMSD_{pairwise}$. Consequently, the *target structure* remains well within the bundle of accepted structures for all tolerance settings. In other words, higher tolerance does not introduce bias or inaccuracy. We finally note that the variation in the structures resulting from differences in the two best parameterizations of ring current and anisotropy is less than the uncertainty in the derived structures (Table 1 and Supplementary Material, Table S13/14).

Effect of the number of chemical shift restraints per basepair We find that down to ~4 chemical shifts per basepair, helix structures can be derived from chemical shifts restraints with reasonable precision ($RMSD_{pairwise} < 1.5$ Å, Table S15a, Figure S7) and accuracy, i.e. no bias because $RMSD_{av2target} \ll RMSD_{pairwise}$. The VCSQ-factor of the accepted structures stays below 0.5 indicating that even non-employed chemical shifts remain reproduced rather well (Table S15a, Figure S7). Below 4 chemical shifts per basepair the number of converging and physically feasible resulting structures strongly reduces.

We also investigated three data sets of experimentally 'easily' accessible chemical shifts (base protons plus H1′/2′ plus H1′/4′ or plus H1′, "Materials and methods"). The VCSQ-factors of these accepted structures (black crosses in Figure S7A) reside just below the line of the 'regular' jackknife test results, indicating slightly better correspondence. Similarly, the $RMSD_{pairwise}$ (and $RMSD_{av2target}$) are equal or slightly better (Table S15a) than the trend of removal of random chemical shifts would suggest (Figure S7B). This most likely results from the more even distribution of the chemical shifts over the helix, i.e. they do not cluster in certain residues as may happen in the sets with randomly removed shifts. For all three shift sets the $RMSD_{pairwise}$ is ca. 1 Å, which indicates that per basepair aromatic proton shifts in combination with at least one or two sugar proton shifts (H1′ and/or H2′/4′) suffice to accurately define the helix structure.

Table 1 Characteristics of the accepted structures resulting from structure derivation based on non-exchanging ^{1}H chemical shifts of the RNA target with sequence 5′GGACGA3′/5′UCGUCC3′ and fiber model standard A-helix conformation (Table S3)

Parameter set[a]	GP	GP	GP	GP	DC
Chemical shift tolerance (ppm)[b]	0.05	0.10	0.15	0.30	0.10
N[c]	956	956	955	963	900
$RMSD_{av2target}$ (Å)[d]	0.33	0.38	0.41	0.44	0.57
$RMSD_{pairwise}$ (Å)	0.50	0.68	0.81	1.13	0.84
$RMSD_{chem.shift}$ (ppm)	0.011 (2)	0.019 (5)	0.027 (7)	0.055 (15)	0.023 (3)
$<RMSD_{helixpar2target}>$[e]	0.55	0.74	0.88	1.23	0.97

As 'observed' chemical shifts were used the 93 ^{1}H non-exchangeable proton chemical shifts of the 6 basepair RNA A-helix calculated when in its fiber model standard A-helix conformation (Table S3, see text) and using the Giessner-Prettre (GP) parameter set for ring-current/magnetic anisotropy. During the iterative SVD structure calculations the ^{1}H chemical shifts were calculated from intermediate structures using either the ring-current/magnetic anisotropy parameter set of Giessner-Prettre (GP) or the ring-current parameter set of Case et al. (DC) (see text). The starting ensemble of 1,000 random structures ($\pm 3\sigma_{helixpar}$) has an $RMSD_{pairwise}$ of 5.56 Å and an $RMSD_{chem.shift}$ of 1.35 ppm

[a] GP is the Giessner-Prettre parameter set and DC, Case parameter set

[b] The chemical shift tolerance for convergence in the iterative SVD

[c] N is the number of accepted structures

[d] $RMSD_{av2target}$ is the RMSD of the average final accepted structure (see text) to the *target structure* and represents the accuracy. $RMSD_{pairwise}$ is the pairwise RMSD of the accepted structures and represents the precision. The $RMSD_{chem.shift}$ is the RMSD of the chemical shifts of accepted structures with respect to the 'observed' values, while in parentheses is given the standard deviation in $RMSD_{chem.shift}$.in ppm

[e] The mean RMSD of the helix parameters from *target structure* values normalized to $\sigma_{helixpar}$; the average is taken over all helix parameters in all bases and structures in the final ensemble

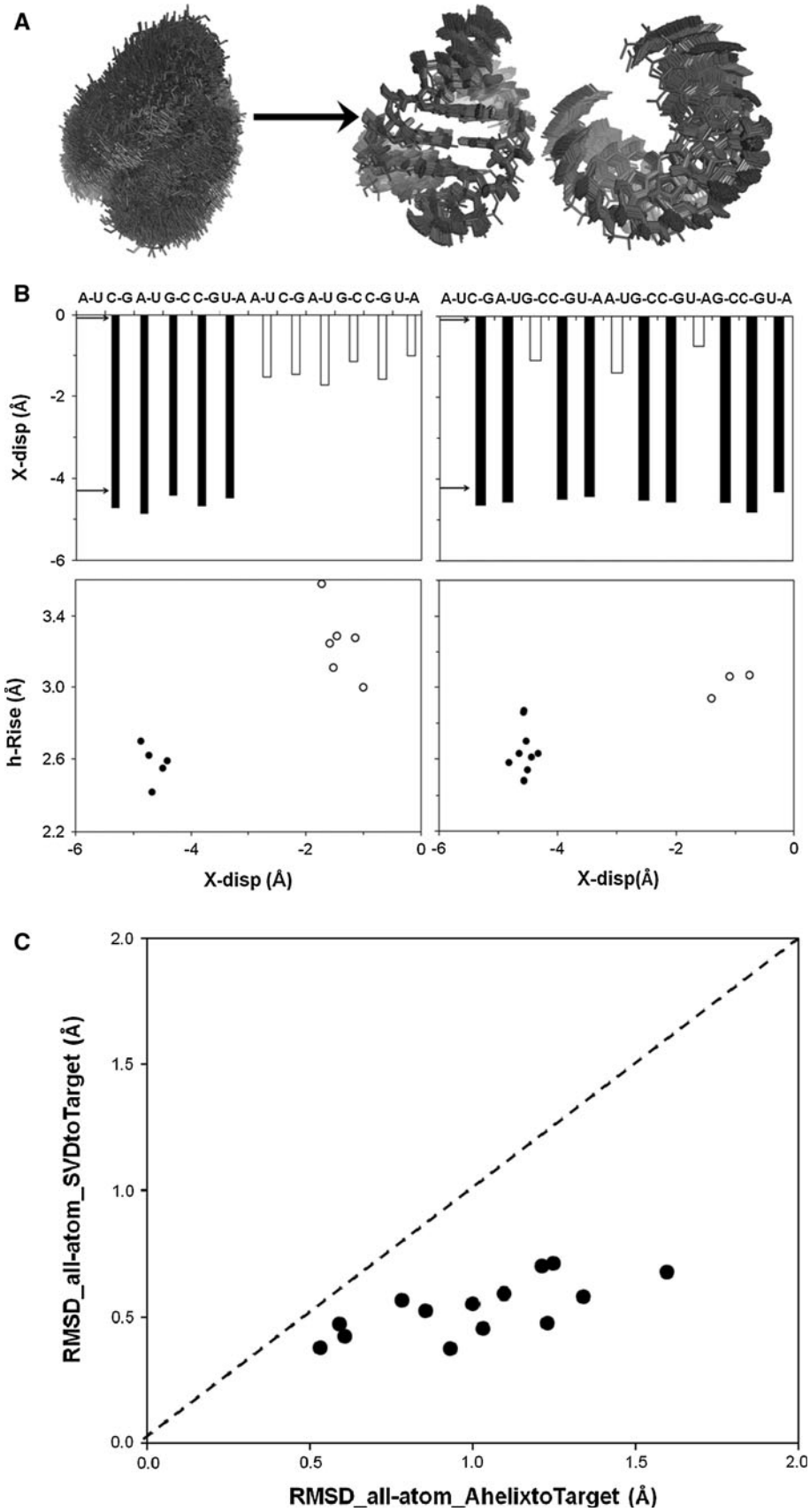
We do not observe, here or in the earlier simulations, final structures that appear locked in a local minimum, i.e. structures that deviate from the target and at the same time show correct or nearly correct chemical shifts. This observation is important because it implies that ill-converged structures, which of course may occur, can be identified. Given a certain preset shift tolerance in combinations with the number of employed chemical shift restraints, converged structures will have a chemical shift deviation that is below a certain limit, a limit that can be read off from Figure S6. It also implies that the standard deviations in the all-atom RMSD and in the helix parameters can be predicted to a large degree in advance given these settings from for instance, Figure S6 and Tables S4/S15a as well as Tables S5–S14.

Non A-helix structures as target

We consider here the performance of CHEOPS employing helices with non-standard A-form conformations as targets and using the chemical shifts of the targets as observed shifts (see "Materials and methods"). For the structure derivation, we started as above for each target from 500 starting structures generated around the standard A-helix conformation. With regard to the choice of these non-A-helix targets, it should be noted that simple random deviations from standard-A-form values in the helix parameters will lead to mostly physically impossible structures/conformations and thus to conformations that are not very relevant. We therefore decided after ample consideration to select two groups of non-standard A-helix targets as follows. Group A: Two $\sim$12 base-pair helices that display A to B transitions in different ways in a number of basepair steps. (a) Helix AtoB is first A-helix like and then halfway proceeds as a B-helix. (b) In helix AtoB2, three basepairs with B-helix conformation reside between A-helix basepairs. Group B: Pdb-deposited RNA helices with non-standard A-helix conformations. First we consider in helix parameter detail the 8 base-pair RNA helix, 1RXA. It displays a range of non-ideal A-helix target conformations in the various C:G basepairs and basepair steps (Egli et al. 1996; Portmann et al. 1995). In this way, a broad range of non-standard A-helix structures, that are physically possible, is tested as targets via the many different non A-helix basepairs and basepair steps. In addition, we selected from crystal structures a set of RNA helices that each shows significant deviations from standard A-helix conformation (see below). The ^{1}H chemical shifts of these targets we used as restraints in the subsequent ^{1}H chemical-shift-based structure determinations (see "Materials and methods"). We have obtained the following results with regard to convergence of the calculated chemical shift based structures.

Group A: A- to B-helix transitions When going from an A-helix to a B-helix conformation, the most significant change lies in the X-disp, -4.3 to 0 Å from A- to B-helix, followed by the change in h-Rise, from 2.7 to 3.3 Å from A- to B-helix. In terms of z-scores these changes are equal to 3.7 (z-score$_{_X\text{-disp}}$) and 1.4 (z-score$_{_h\text{-Rise}}$); here, z-score$_{_X\text{-disp}}$ = $(H_{X\text{-disp,B-helix}} - H_{X\text{-dis,A-helix}})/\sigma_{helixpar}$ and similarly for h-rise. As evident from Fig. 3b, these two most significant parameters are indeed recovered quite well in the calculated structures for both helices with A to B transitions.

Group B: pdb-deposited non-A-helix RNA structures We first consider 1RXA. This 8-basepair RNA helix shows many helix parameters that significantly deviate from standard A-helix values as evident from the target z-scores that reach as far as $\sim\pm3$ (Figure S8, Table S15b) and is thus well suited as a target with physically possible non-A-form conformations. In the calculated structures, derived using ^{1}H chemical shift restraints (see "Materials and methods"), the many non-A-helix target values are essentially all found back (Figure S8, Table S15b). We observe a good correlation with a narrow bandwidth between the z-scores of the helix parameters of the calculated and target structures (Figure S8). Hence, all helix parameters converge well to the non-A-helix targets in the present context. Note that the bandwidth of the |zscore_calc-zscore_target| spreads up to $\sim$1, whereas the |zscore_target| runs up to $\sim$3. In other words, CHEOPS shows a significant improvement over the naïve assumption of just an ideal A-helix as calculated structure. In addition, we selected a further set of $\sim$14 RNA helix fragments from crystal structures using the RNA FRABASE 2.0 (Popenda et al. 2010) and subsequently using the NDB to obtain their helix parameters (Berman et al. 1992). These RNA helix fragments were chosen, because they display significant deviations from standard A-helix conformation. Tables S15d provides an overview of various characteristics of this set of non-standard A-helix RNA fragments. The RNA fragments are extracted from crystal structures with the pdb- and/or NDB codes, 3I56_2/NA0124 (Gurel et al. 2009), 1LC4_1/DR0007 (Vicens and Westhof 2002), 2AOP_1/AR0063 (Haeberli et al. 2005), XXXX_2/NA0253 (Jenner et al. 2010), 3O58_1/NA0743 (Ben-Shem et al. 2010), 3O58_2/NA0743 (Ben-Shem et al. 2010), 2BTE_1/PR0252 (Tukalo et al. 2005), 1RXA, 3R8S_2/NA1049 (Dunkle et al. 2011), 3R8S_1/NA1049 (Dunkle et al. 2011), 3PDR_1/NA0682 (Ramesh et al. 2011), 3OFD_1/NA0682 (Dunkle et al. 2010), 3OFR_1/NA0674 (Dunkle et al. 2010), XXXX_1/NA0535 (Jin et al. 2010). These helix fragments stem from different contexts, i.e. from different ribosomes and/or ribosome fragments, from tRNAs, from a riboswitch, and even from an RNA in which one phosphate-sugar link is substituted for a thio-sugar link (Table S15d).


A
B
A-U C-G A-U G-C C-G U-A A-U C-G A-U G-C C-G U-A
A-UC-GA-UG-CC-GU-AA-UG-CC-GU-AG-CC-GU-A
X-disp (Å)
h-Rise (Å)
X-disp (Å)
X-disp(Å)
C
RMSD_all-atom_SVDtoTarget (Å)
RMSD_all-atom_AhelixtoTarget (Å)

◀**Fig. 3** RNA structure derivation based on ^{1}H chemical shifts, performance tests on simulated chemical shifts. **a** Illustration of the RNA structure-derivation based on chemical shifts via a 6-basepair RNA A-helix with sequence 5′GGACGA3′/5′UCGUCC3′ as target in a fiber model standard A-helix conformation (Table S3). On the *left*, overlays are shown of the 1,000 random starting structures, which cover $\pm 3\sigma_{helixpar}$ of allowed A-helix space. On the right, overlays of the final 956 accepted structures are displayed in two views. A set of restraints consisting of 15 chemical shifts per basepair was used, the threshold for convergence was set to 0.05 ppm, and the GP-parameter set was used in the chemical shift calculations. **b** Performance tests of the chemical shift based structure calculation using as targets two ~12 basepair RNA helices that are each a hybrid of A- and B-conformations. The *left panels* show the data for AtoB target, in which the A-helix to B-helix transition occurs approximately halfway the sequence. Here, 186 chemical shift restraints were used and 99 accepted structures were obtained out of 500 starting structures (convergence threshold 0.10 ppm). The *right panels* show the data for the AtoB2 target, in which the base pairs with B-helix conformation reside in between basepairs with A-helix conformation. Here, 201 chemical shift restraints were used and 144 accepted structures were obtained out of 500 starting structures (convergence threshold 0.10 ppm). The average helix parameters X-disp and h-Rise of the set of accepted chemical-shift-derived structures (converged SVD structures) are displayed, because these helix parameters differ most going from the A-conformation to the B-conformation (see main text). The *top panels* show the average values of the X-disp along the sequence (arrows: X-disp ~−4.3 Å for standard A-conformation and X-disp ~0 Å for standard B-conformation; *filled bars* indicate residues that are A-form in the target, while open bars indicate residues in the B-form in the target). These X-disp values demonstrate that in the chemical shift based structures the B-type conformation is essentially recovered even for basepairs next to or in between basepairs with A-type conformation. The correlation plots of X-disp versus h-Rise of each base pair in the *lower panels*, show that in the chemical shift based structures also the B-type h-Rise values are recovered for the basepairs with B-type conformation (h-Rise ~3.3 Å in standard B-conformation and ~2.7 Å in standard A-conformation). The *filled symbols* indicate residues A-form in the target, while *open symbols* refer to residues in the B-form in the target. **c** Performance test of CHEOPS on a set of ~14 non-standard A-helix RNA targets. The *scatter plot* shows for each target the all-atom RMSDs of target to standard A-helix versus the all-atom RMSDs of the corresponding sets of converged SVD structures to target. The details of targets and structure calculations are given Table S15d. Note that as above the target ^{1}H chemical shifts are used as restraints with ~15 chemical shifts per basepair and the convergence threshold was set to 0.1 ppm

For these non-standard A-helix RNA targets, we calculated the atomic RMSDs of the A-helix conformation versus target, RMSD_all-atom_Ahelixvstarget (Table 15d). As can be seen, in the set the atomic RMSD from standard A-helix of these targets ranges up to ~1.5 Å, e.g. 1RXA shows an RMSD of ~0.78 Å. As for the other non-standard A-helix RNA targtes, we calculated their chemical-shift-based structures (SVD converged) using the target ^{1}H chemical shifts as restraint and with similar parameter settings. To assess the quality of the structure derivation, we calculated for each set of converged SVD structures the atomic RMSD versus target, RMSD_all-atom_SVDvstarget (Table S15d). The scatter plot of RMSD_all-atom_Ahelixvstarget versus RMSD_all-atom_SVDvstarget (Fig. 3c) shows that RMSD_all-atom_SVDvstarget is considerably smaller than the RMSD_all-atom_Ahelixvstarget. In other words, these data show that CHEOPS leads to a significant improvement over simply assuming a standard A-helix.

This improvement can also be understood when considering the ^{1}H chemical shifts. The ^{1}H chemical shifts are quite sensitive to such relatively small conformational changes as in the $\pm 3\sigma_{helixpar}$ range around standard A-helix conformation. This is evident for instance from Fig. 1, but also from RMSD_CS_targetvsAhelix of the RNA target structures considered in Table S15c. For these targets, the RMSD_CS_targetvsAhelix values range from 0 to 0.3 ppm, while their helix parameters spread over $\pm 3\sigma_{helixpar}$ around the standard A-helix conformation. In contrast, for the chemical-shift-derived structures (e.g. SVD converged), the chemical shift RMSDs from target, CS_RMSD_SVDvstarget, are much smaller than this 0.3 ppm. For instance, CS_RMSD_SVDvstarget is much smaller than and uncorrelated with CS_RMSD_targetvsAhelix (Table S15c and Figure S9). Thus, this demonstrates the potential to derive $\pm 3\sigma_{helixpar}$ deviations from A-helix from chemical shift based structures. That the ^{1}H chemical-shift-derived structures in fact do probe such small deviations, is evident from (1) the observations RMSD_all-atom_SVDvstarget is considerably smaller than RMSD_all-atom_Ahelixvstarget for the non-standard A-helix targets considered and (2) the similar observations in terms of z-scores of helix parameter (see e.g. 1RXA and AtoB and AtoB2). In conclusion, we demonstrated, for quite a number of A-helix and non-A-helix target structures, the convergence of the chemical shift based structures.

Evaluating CHEOPS on nucleic acids of known structure using experimental chemical shifts

Strategy and selection/analysis of pdb structures

To evaluate the performance of CHEOPS using experimental chemical shifts, the helical stem structures of three published RNA stem-loops were derived *de novo* from the available proton chemical shifts, namely the HIV-1 frame shift inducing RNA stem-loop (Staple and Butcher 2003) (HIV_RNA, 1PJY), the SIV frame shift inducing RNA element (Marcheschi et al. 2007) (SIV_RNA, 2JTP; see also "Materials and methods"), and an RNA stem containing a tandem GU wobble base pair (McDowell et al. 1997) (UG_RNA, 1QES, see also "Materials and methods"). The three deposited stem-loop structures are based on standard NMR restraints. No X-ray structure is available for these RNAs. However, their helix regions contain only

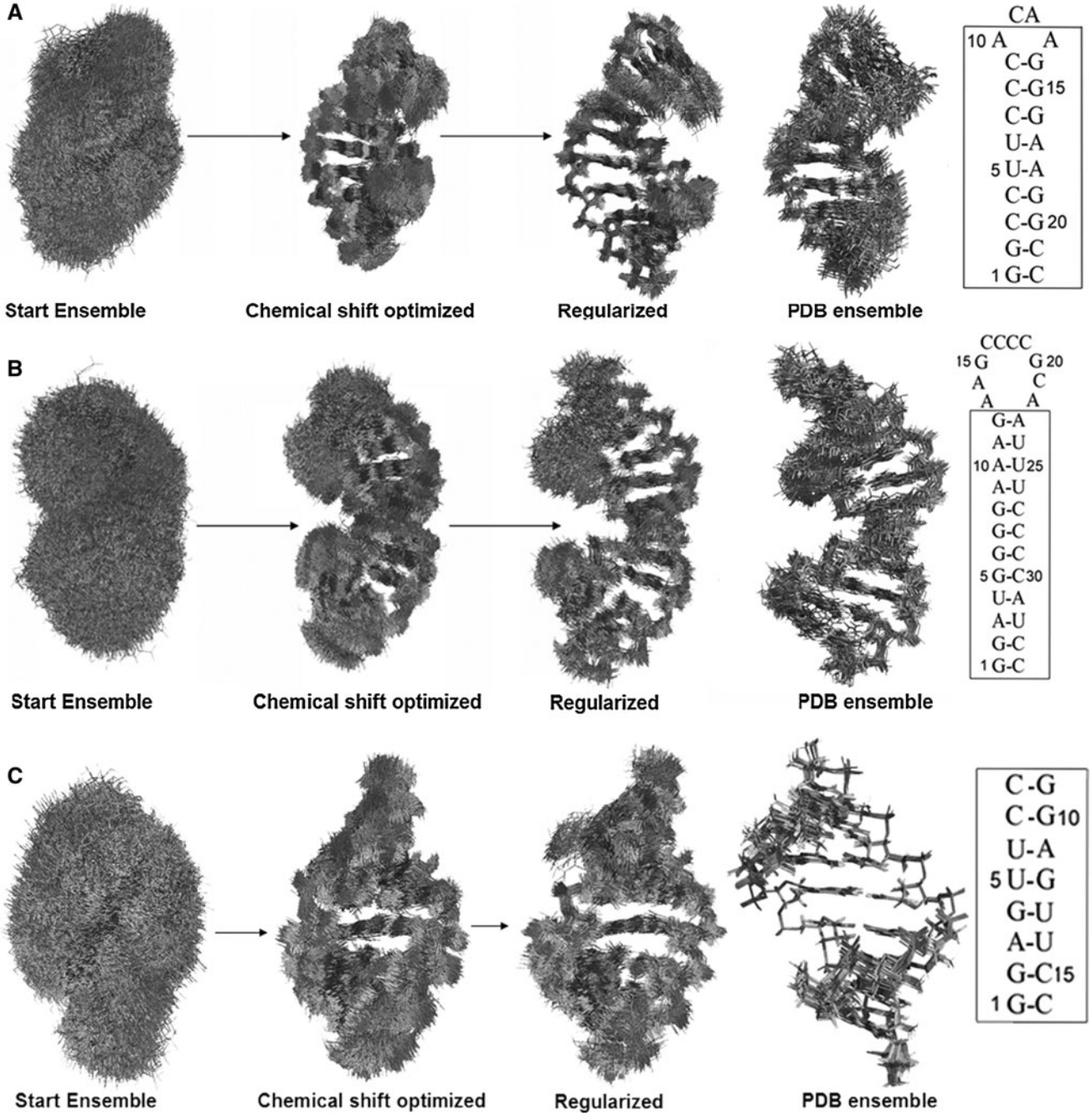


Fig. 4 Structure derivation using experimental proton chemical shifts demonstrated on the HIV-1 (**a**) and SIV (**b**) frame-shift inducing RNA elements and on the duplex RNA containing a tandem GU wobble basepair (**c**). Sequence and secondary structures are shown to the far *right*. The *boxed* regions contain the helical segments whose structures were derived using observed chemical shifts as experimental restraints. For the HIV-1 RNA, the conformation of the non-canonical A:A basepair in the tetraloop and of the preceding C:G basepair were kept fixed by keeping their helix parameters at the mean values of the deposited NMR ensemble 1PJY. No basepairs were kept fixed in the SIV1 RNA. The ensemble of starting, converged-SVD, final-accepted structures, PDB-deposited structures are shown from *left* to *right*. The chemical shift tolerance was set to 0.15, the parameters λ and κ were set to 4.25 and 0.5, respectively, and the maximum number of iterative SVD cycles was set to 150. For the HIV-1 RNA, SIV RNA and UG RNA 87, 108, and 97 chemical shifts were used, respectively. As described in the text, the backbone was regularized by means of a brief restrained molecular dynamics run, using standard helix NOE restraints (for HIV-1 176; for SIV 209, for UG RNA 134) and during which the bases were kept fixed. For SIV, the backbone regularization was for comparison also carried out using 403 published experimental NOE restraints, with or without the published experimental 24 RDCs (Dax $= -53.80$ Hz, R $= 0.10$)

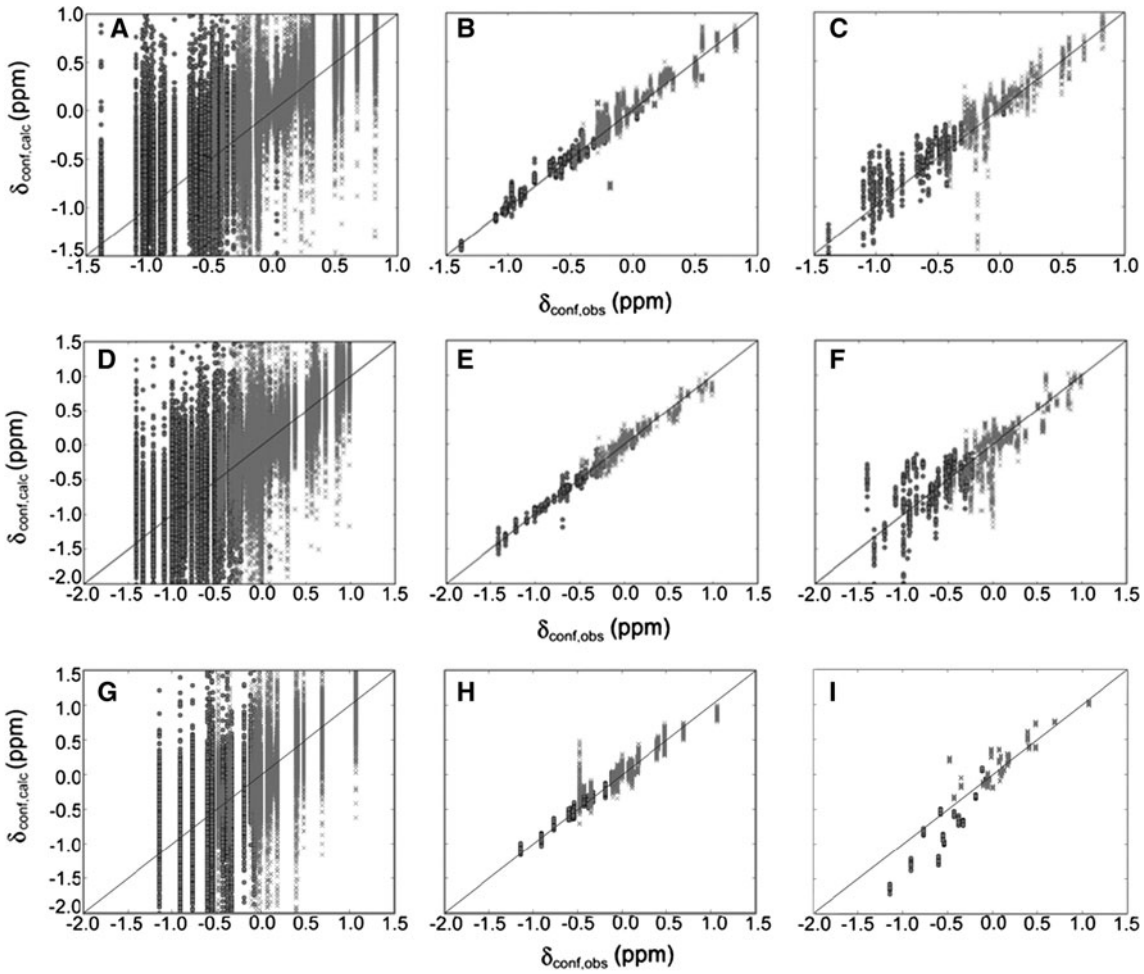


Fig. 5 Correlation of experimental and back-calculated conformational proton chemical shifts. Shown are the correlations between observed (*horizontal axes*) versus back-calculated (*vertical*) conformational proton chemical shifts for HIV-1 RNA (**a**–**c**), SIV RNA (**d**–**f**), and UG_RNA (**g**–**i**). The correlations are displayed for starting ensemble (A: $R_{corr} = 0.01$, $RMSD_{chemshift} = 4.0 \pm 5.6$ ppm; D: $R_{corr} = 0.12$, $RMSD_{chemshift} = 1.6 \pm 3.6$ ppm; G: $R_{corr} = 0.07$, $RMSD_{chemshift} = 5.8$ ppm), final accepted ensemble (B: $R_{corr} = 0.97$, $RMSD_{chemshift} = 0.118 \pm 0.003$ ppm; E: $R_{corr} = 0.99$, $RMSD_{chemshift} = 0.079 \pm 0.009$ ppm; H: $R_{corr} = 0.97$, $RMSD_{chemshift} = 0.105$ ppm), and the deposited PDB ensemble (C: $R_{corr} = 0.92$, $RMSD_{chemshift} = 0.19 \pm 0.02$ ppm; F: $R_{corr} = 0.85$, $RMSD_{chemshift} = 0.27 \pm 0.02$ ppm; I: $R_{corr} = 0.91$, $RMSD_{chemshift} = 0.24$ ppm). Note that for the final accepted set of the SIV RNA the chemical shift correlations are displayed of the final accepted structures obtained by backbone regularization with published distance restraints, but without application of RDC restraints. In the correlation *plots*, the aromatic conformational chemical shifts are mainly at lower values (*filled circles*) and the sugar conformational chemical shifts reside mainly at higher values (*crosses*)

Watson–Crick basepairs, except for UG RNA (Fig. 4). Thus, the helix parameters are expected to fall within allowed ranges for A-helices. The UG RNA with a tandem GU wobble base pair allows investigation of the performance of CHEOPS in establishing the conformation of an A-helix containing non-canonical base pairs.

The helix parameters of the stem sections of the pdb structures of HIV1, SIV RNA and UG_RNA indeed all fall within $\pm 3\sigma_{helixpar}$ of the standard A-helix values (Table S16a–c). Deviations between 2 and $3\sigma_{helixpar}$ are found mostly in terminal residues and thus appear to be due to end-effects. Nevertheless, a number of deviations of 2.0–$2.3\sigma_{helixpar}$ appear in the middle of the helix region. For HIV1 (1PJY), noteworthy are the positive propeller twist values throughout the stem, while the standard A-helix value is $-2° \pm 7°$, as well as two helical rise values that deviate just $2\sigma_{helixpar}$ from the A-helix value of 2.7 Å. For SIV (2JTP), noteworthy are the increased helical rise for three stem residues, a reduced X-displacement and positive propeller twist for two residues. For UG_RNA (1QES) it can be seen that propeller twist, X-displacement and Y-displacement deviate from standard helix values in the UG base pairs and the base pairs that sandwich them. We hypothesize that in the deposited RNA structures, derived with standard NMR methods, such deviations at least in part occur due to limitations in the number and accuracy of NOE-derived distance restraints. We employed 87, 108, and 97 shifts for the stem structures of HIV, SIV,

and UG_RNA, respectively (ca. 9.7 shifts per basepair). The tests with simulated data indicate that given ca. 9.7 shifts/basepair, the $RMSD_{pairwise}$ of accepted chemical shift structures should be ca. 1 Å (Figure S7) and the RMSD in their helix parameters ca 0.5–1.5$\sigma_{helixpar}$ (Figure S5 and S6). It is thus expected that the larger deviations in the pdb structures should disappear in the chemical shift based structures.

Because we want to focus on the stem regions, the proton chemical shifts of loop residues, which includes the non-canonical closing basepair, were not employed as restraints. For HIV RNA (1PJY), we removed in addition five H3′/4′ shifts, because they could be affected by conformational exchange due to repuckering and/or end-effects (G14 H3′, C22 H3′/4′) or by puckering conformational exchange alone (G15 H3′, G19 H3′, C21 H3′). These sugar proton shifts were identified as 'outliers' as described in "Materials and methods". For 2JTP (SIV RNA), we removed the H5′ sugar proton shifts of G1, G2, A3, U4, G5, and U25 (identified as 'outliers, see "Materials and methods"), because they are difficult to assign stereo-specifically and/or affected by end-fraying effects (G1); the removed 'outlier' H1′/2′ of C34 could be affected by such end-effects. For 1QES (UG_RNA) we removed some sugar proton shifts (G1H1′, G4H2′, C8H2′, G12H2′ and C16H2′; H5′ and H5″ of G1, G2, G4, C8, G9, G10, G12, and C16; H4′ of G1, C8, G9, G10, and C16; H3′ of G1, G4, G9, G12, U14, and C16). In total, we removed 6/8/34 out of the 93/116/109 published proton shifts for the HIV/SIV_RNA/UG_RNA stems.

The HIV-1 and SIV frame-shift inducing RNA stem loops (HIV1_RNA and SIV_RNA) and the helix containing a tandem GU wobble (UG_RNA)

For HIV1_RNA (Fig. 4), the $RMSD_{chem.shift}$ drops from ~4 ppm in the 500 starting structures to 0.06 ppm in the ensemble of converged SVD structures (Table S17a), a considerable improvement over the $RMSD_{chem.shift}$ of 0.16 ppm in the deposited NMR structure (Table S17b). The width of the structure bundle ($RMSD_{pairwise}$) decreases from 3.3 Å in the starting ensemble to 0.9 Å. SIV_RNA (SIV_RNA in Fig. 4) and UG RNA gave similar numbers (Table S17b and S17c). These converged SVD structures were subsequently backbone regularized during which the base pairs are kept fixed and the ribose conformation restrained and ultimately checked for violations (see "Materials and methods") to obtain the final accepted structures.

For HIV1-RNA, a relatively small decrease in the $RMSD_{pairwise}$ is seen upon backbone regularization in the final accepted structures (from 0.9 to 0.7 Å; Tables S17a/18a). The $RMSD_{chem.shift}$ usually increases somewhat during backbone regularization, here from 0.06 to 0.12 ppm (Table S17a/S18a). This increase is mostly caused by changes in H1′ and H2′ proton shifts, resulting from slight reorientations in the corresponding C–H bonds (<2°) during backbone regularization. Furthermore, the common ribose unit for N- and S-type puckering used in Xplor does not exactly match the individual N- and S-type ribose units used in 3DNA. This aspect contributes additionally to C–H bond reorientation in the sugar and thus to the slight changes in proton chemical shifts. The impact of backbone regularization on the $RMSD_{chem.shift}$ illustrates again the high sensitivity of the chemical shift to changes in structure. We finally note that the final accepted structures show no violations (>1.0 Å) with respect to the standard A-helix inter-residue sugar–sugar and sugar-aromatic distances (Wijmenga and van Buuren 1998), attesting to the quality of these structures.

For SIV_RNA, the backbone regularization using the standard A-helix NOE restraints, led to an $RMSD_{chem.shift}$ of 0.08 ppm in the final accepted structures, essentially the same as in the converged SVD structures. We also performed backbone regularization with various other combinations of experimental and standard A-helix NOE and RDC restraints. The results are comparable in terms of the various RMSDs (Table S18b). The published NOE and RDC restraints permitted an independent validation of the final accepted structures. In the final accepted structures, the average $RMSD_{RDC}$ ranges from 19 (12 %) to 23 Hz (14 %; Table S18b) in the absence of RDC restraints during backbone regularization. It drops to 13 Hz (8 %) in the presence of RDCs (Table S19b). The percentages are calculated as the ratio of the RMSD over three times the absolute value of the axial component of –54 Hz of the alignment tensor. These values are just above the $RMSD_{RDC}$ of 10 Hz (6 %) of the deposited NMR structure set. The reduction in $RMSD_{RDC}$ upon inclusion of RDC restraints in the backbone regularization is due to sugar C–H RDC restraints that optimize sugar conformation, because the bases are kept fixed during backbone regularization. The final accepted structures do not show any distance violations >1.5 Å with respect to the BMRB experimental NOE distance restraints (Marcheschi et al. 2007).

For UG RNA the final accepted structures do not show any distance violations > 1.5 Å with respect to the BMRB experimental NOE distance restraints (McDowell et al. 1997). The $RMSD_{chem.shift}$ increased somewhat during backbone regularization (from 0.06 to 0.10 ppm, Table S17c/S18c). This increase is caused by slight reorientations in the corresponding C–H bonds during backbone regularization, which is mainly caused by the small mismatch between the generic ribose unit used in Xplor and the one used in 3DNA, which equals the exact N- or S-puckered

X-ray conformation (see above explanation in paragraph on HIV RNA or Supplementary Material).

Figure 5 shows the correlations of $\delta_{conf,obs}$ with $\delta_{conf,calc}$ for the starting, final-accepted, and pdb-deposited ensembles of HIV RNA, SIV RNA, and UG_RNA. The final accepted structures (Fig. 5b, e, h) show a significantly improved correlation over that of the pdb-deposited structures (Fig. 5c, f, i). It is interesting to discuss a few specific examples. (1) In the correlation plot of the final accepted chemical shift structures of HIV RNA (Fig. 5b) one chemical shift (G14 H1′) still lies somewhat off the diagonal. In the pdb ensemble, $\Delta\delta_{conf}$ (= $\delta_{conf,obs} - \delta_{conf,calc}$) of G14 H1′ varies between −1.51 and 0.10 ppm. In the final accepted structures, $\Delta\delta_{conf}$ is smaller and stable around a value of −0.55 ppm, which is still relatively large. Analysis shows that the variation $\Delta\delta_{conf}$ in the pdb ensemble is largely caused by variation in buckle (variance 6°), propeller twist (variance 10°), and opening (variance 12°) of the A10–A13 non-canonical basepair. This basepair was kept fixed during SVD structure calculation in a conformation defined by the average the helix parameter values of the pdb ensemble. (2) For the SIV_RNA ensemble of accepted chemical structures (Fig. 5e), the values are unbiased for all chemical shifts and show little variation. The A11 H2 also shows essentially no bias but still varies over a relatively broad range; $\Delta\delta_{conf}$ runs from 0.41 to −0.22 ppm at $\delta_{conf,obs}$ is −0.69 ppm. This relatively wide range $\Delta\delta_{conf}$ is caused by variation in the conformation of the non-canonical G12–A23. This basepair closes the loop, lies directly above A11 and was left free to adjust within restraint bounds during backbone regularization. (3) The large $\Delta\delta_{conf}$ for A11 H8 of SIV in the ensemble of pdb structures (at $\delta_{conf,obs}$ −1.41 ppm, Fig. 5f), has completely disappeared in the accepted chemical shift structures (Fig. 5e). (4) In Fig. 5h (UG RNA), the chemical shifts of the symmetry-related H2′ sugar proton spins of U13 and U5 (at $\delta_{conf,obs} \sim -0.5$ ppm; crosses sugar protons) are seen to spread somewhat on a vertical line from the diagonal. Note further that in the pdb-deposited NMR-ensemble (Fig. 5i) the same H2′ shifts also deviate from diagonal. Due to the proximity of the H2′ to its own and/or sequential base, some variation in sugar puckering and/or sugar orientation with respect to the base plain may cause a significant change in chemical shift and thus explain the somewhat enhanced spread in the H2′ calculated shift. 5) Finally, in the pdb-deposited UG_RNA NMR structures (Fig. 5i) the aromatic $\delta_{conf,calc}$ are too small for quite a number of them (points on left below diagonal). This deviation disappears in Fig. 5h, attesting to the improved accuracy of the chemical shift based structures.

For HIV1_RNA and SIV RNA, the RMSD of the helix parameters of the ensemble of final accepted structures ranges from 0.6–1.2$\sigma_{helixpar}$, indicating that the structures are determined with high precision. The final accepted structures closer resemble a standard A-helix than the pdb structures (Tables S16a/b vs. S19a/b). For HIV RNA, most striking is the absence of the unusual positive signs of propeller-twist of all basepairs and too large helical rise of basepairs G2:C21, C3:G20, and C4:G19 (Table 16a vs. 19a). Leaving end-effects aside (basepairs G1:C22/G2:C21, C9:G12/A10:A11), deviations from standard A-helix values larger than 2.1$\sigma_{helixpar}$ have disappeared in the final accepted structures, except for C8:G15 (deviation 2.2–2.6$\sigma_{helixpar}$). The latter is most likely transference of end-effects from the closing A:A/C:G basepairs of the tetra-loop. SIV_RNA shows no deviations larger than 2.2$\sigma_{helixpar}$, leaving end-effects aside (Table 16b/19b). In conclusion, the average helix parameters of the final accepted structure of the two RNAs with regular A-helix stems indeed fall within the allowed ranges for A-helices with Watson–Crick basepairs, and the structures satisfy the experimental NOE and RDC restraints.

In UG_RNA (Table S16c) a tandem of non-canonical UG basepairs is situated in the middle of the A-helix stem and certain helix parameters are expected to deviate from standard A-helix values. The UG basepairs indeed display in the deposited structure ensemble shear and to lesser degree opening helix parameter values that significantly deviate from regular A-helix values (Table S16c). The tandem of UG basepairs also affects the base-pair helix parameters of neighboring Watson–Crick residues, i.e. Prop-Tw and to a lesser degree Buckle. In addition, some base-pair step parameters between and to or from the UG tandem basepair deviate significantly from regular A-helix values (X-disp, Y-disp, h-Twist). The Watson–Crick basepairs further away show helix parameters within the range of 0.6–1.2$\sigma_{helixpar}$, excluding some end-effects. In summary, the deposited set UG_RNA adheres to A-helix parameters except for the UG and their flanking basepairs. In the final chemical shift based structures (Table 19c), the helix parameters essentially show the same deviations from standard A-helix parameters. Note for example the basepair step parameters, Y-disp and h-Twist, which show the same pattern of deviation from A-helix values as in the deposited structures. First, the pattern of a negative versus a positive Y-disp value for A:U/G:U versus U:G/U:A, respectively, is seen in both deposited and chemical-shift derived structures. Also, the significantly increased h-Twist value (>40°) over the mean A-helix value for the G:U/U:G basepair step comes back in both deposited and chemical-shift structures. Note further the symmetry along the sequence in the h-Twist deviations in the chemical-shift based structures in the basepair steps to and from the tandem U:G basepair, one observes a slightly smaller than standard A-helix h-Twist value versus a slightly larger value for A:U/G:U versus U:G/U:A, respectively (Table 19c).

In conclusion, the average helix parameters of the final accepted structure of the three RNAs indeed fall within the allowed ranges for A-helices with Watson–Crick basepairs, and the structures satisfy the experimental NOE and RDC restraints. In addition, CHEOPS can be used to derive the helix structures, which include non-canonical basepairs such as UG-basepairs.

Concluding remarks

We presented an efficient method for determining the nucleic acid helix structure with high precision and accuracy that uses exclusively proton chemical shifts as experimental restraints. We have used simulated data to establish in a defined manner the effect of errors and data limitations of the method. We further demonstrated via three deposited RNA structures that also experimental proton chemical shifts suffice to derive RNA helix structures with high precision and accuracy. A few concluding remarks on the possibilities and limitations of CHEOPS appear in order.

1. Structures derived with CHEOPS are fully based on chemical shifts as no use is made of sequence homology like in chemical shift based structure programs like CS ROSETTA (Shen et al. 2008), CS23D (Wishart et al. 2008), and CHESHIRE (Cavalli et al. 2007).
2. The CHEOPS method works and has been demonstrated to work for RNA helices with canonical and non-canonical basepairs. We explicitly demonstrate that CHEOPS can derive non-canonical basepairs by way of the derivation of the structure of a tandem UG basepairs embedded in RNA Watson–Crick Stem. Other non-canonical basepairs can be handled in CHEOPS as well. We note in this respect that the sensitivity of ^{1}H chemical shifts to helix parameters that affect the conformation of the basepair or the associated basepair steps could simply be checked by way of test calculations in CHEOPS (see also point 5 below, grid search). CHEOPS also works for DNA helices, which will be demonstrated in a forthcoming paper.
3. We implemented the SVD procedure and derived the structures in helix parameter space. The helix parameters lead to smooth trajectories and their usage prevents conformations from becoming locked into local minima.
4. Apart from the SVD minimization, the CHEOPS method also allows for performing a grid search of the helix parameter space to find for instance an optimal conformation, e.g. to establish the conformation of a non-canonical basepair or separate optimization of a helix parameter.
5. Moreover, it is also possible to keep certain helix parameters fixed during SVD or grid search.
6. The calls to the 3DNA build function also allow for atomic models of single-stranded RNA/DNA molecules to be constructed from the helix parameters. To employ the helix parameter description within the context of a single strand is therefore an obvious and rather straightforward extension of the method, but has of yet not been implemented.
7. The CHEOPS method includes a restrained Molecular Dynamics step, which we use for regularization of the backbone as explicitly shown. During the backbone regularization, the base positions are kept fixed and the ribose sugar conformation is restrained so that only the ribose-phosphate backbone is optimized and most importantly, the helix conformation is not affected. The inclusion of the rMD has however the advantage of flexibility and allows in the present state of CHEOPS for a hybrid approach in which the helix is SVD optimized based on chemical shifts and the full structure including the loops is ultimately calculated using rMD.
8. Structure deviations due to interactions with proteins can be assessed with the present method. An interesting situation that may also occur is interaction with metal ions, like Mg^{2+}. Charge effects on ^{1}H chemical shift are accounted for in the chemical shift calculations in NUCHEMICS. However, the effect of full or partial charges on the ^{1}H chemical shifts have in nucleic acids been found relatively small (Cromsigt et al. 2001; Wijmenga et al. 1997).
9. A further and possibly even more important point is proper accounting for dynamics. Loop structures are commonly dynamic to some degree (e.g. (Flodell et al. 2006; Mooren et al. 1994; Petzold et al. 2007)) and accounting for such dynamics is then relevant. CHEOPS minimizes for each structure in the ensemble the error function. In that respect, CHEOPS is similar to the usual rMD NMR structure calculation. Hence, the single structure obtained best accounts for the measured restraints, which are in fact averages over many molecules and in case of conformational dynamics also averages over many different conformations. In the absence of dynamics or in the presence of small-scale dynamics, as in libration motions, the structure derived with this usual rMD or CHEOPS approach leads to a single structure that equals the actual physical conformation or to one that equals the average conformation, which is still physically correct. The effect of such dynamics on the restraints is to increase the minimum in the error function and thus it increases the width of

the bundle of minimized structures, in which case the width is reflective of such dynamics. For more complex dynamics, the average conformation may not be physically correct, an example of such complex dynamics is formed by conformational exchange between N- and S-puckered states; consequently, a physically incorrect structure would then also be derived. Alternatively, when the pucker is kept fixed during the structure derivation, a physically correct structure is enforced, but then the minimum error deviates more strongly from its real minimum. The width of the bundle of structures is likely to increase, but the structures do not really reflect the physical events. To account for such complex dynamics requires additional method development, e.g. inclusion of ensemble averaging. Nevertheless, the chemical shift is highly sensitive to conformational differences and to highly susceptible to dynamics, which opens new opportunities for assessing dynamics.

Acknowledgments The 6th framework program of the EU, project FSG-V-RNA, is acknowledged for funding.

References

Allain FHT, Varani G (1997) How accurately and precisely can RNA structure be determined by NMR? J Mol Biol 267:338–351

Altona C, Faber DH, Hoekzema AJAW (2000) Double-helical DNA ^{1}H chemical shifts: an accurate and balanced predictive empirical scheme. Magn Reson Chem 38:95–107

Arnott S, Hukins DWL (1973) Refinement of the structure of B-DNA and implications for the analysis of X-ray diffraction data from fibers of biopolymers. J Mol Biol 81:93–105

Bailor MH, Musselman C, Hansen AL, Gulati K, Patel DJ, Al-Hashimi HM (2007) Characterizing the relative orientation and dynamics of RNA A-form helices using NMR residual dipolar couplings. Nat Protoc 2:1536–1546

Ben-Shem A, Jenner L, Yusupova G, Yusupov M (2010) Crystal structure of the eukaryotic ribosome. Science 330:1203–1209

Berman HM, Olson WK, Beveridge BL, Westbrook J, Gelbin A, Demeny T, Hsieh SH, Srinivasan AR, Schneider B (1992) The nucleic acid database: a comprehensive relational database of three-dimensional structures of nucleic acids. Biophys J 63: 751–759

Borgias BA, James TL (1990) Mardigras—a procedure for matrix analysis of relaxation for discerning geometry of an aqueous structure. J Magn Reson 87:475–487

Brunger AT (1992) X-PLOR version 3.1 a system for X-ray crystallography and NMR. Yale University Press, New Haven

Case DA (1995) Calibration of ring-current effects in proteins and nucleic acids. J Biomol NMR 6:341–346

Cavalli A, Salvatella X, Dobson CM, Vendruscolo M (2007) Protein structure determination from NMR chemical shifts. Proc Natl Acad Sci USA 104:9615–9620

Cornilescu G, Hu JS, Bax A (1999) Identification of the hydrogen bonding network in a protein by scalar couplings. J Am Chem Soc 121:2949–2950

Cromsigt JAMTC, Hilbers CW, Wijmenga SS (2001) Prediction of proton chemical shifts in RNA—their use in structure refinement and validation. J Biomol NMR 21:11–29

Dejaegere A, Bryce RA, Case DA (1999) An empirical analysis of proton chemical shifts in nucleic acids. In: Facelli JC, de Dios AC (eds) Modeling NMR chemical shifts. Gaining insight into structure and environment. American Chemical Society, Washington, DC

Dickerson RE (1989) Definitions and nomenclature of nucleic acid structure parameters. J Biomol Struct Dyn 6:627–634

Dunkle JA, Xiong L, Mankin AS, Cate JH (2010) Structures of the *Escherichia coli* ribosome with antibiotics bound near the peptidyl transferase center explain spectra of drug action. Proc Natl Acad Sci USA 107:17152–17157

Dunkle JA, Wang L, Feldman MB, Pulk A, Chen VB, Kapral GJ, Noeske J, Richardson JS, Blanchard SC, Cate JH (2011) Structures of the bacterial ribosome in classical and hybrid states of tRNA binding. Science 332:981–984

Egli M, Portmann S, Usman N (1996) RNA hydration: a detailed look. Biochemistry 35:8489–8494

Flodell S, Petersen M, Girard F, Zdunek J, Kidd-Ljunggren K, Schleucher J, Wijmenga S (2006) Solution structure of the apical stem-loop of the human hepatitis B virus encapsidation signal. Nucleic Acids Res 34:4449–4457

Gelbin A, Schneider B, Clowney L, Hsieh SH, Olson WK, Berman HM (1996) Geometric parameters in nucleic acids: sugar and phosphate constituents. J Am Chem Soc 118:519–529

Giessner-Prettre C, Pullman B (1987) Quantum mechanical calculations of NMR chemical shifts in nucleic acids. Q Rev Biophys 20:113–172

Gronenborn AM, Clore GM (1989) Analysis of the relative contributions of the nuclear overhauser interproton distance restraints and the empirical energy function in the calculation of oligonucleotide structures using restrained molecular dynamics. Biochemistry 28:5978–5984

Grzesiek S, Sass HJ (2009) From biomolecular structure to functional understanding: new NMR developments narrow the gap. Curr Opin Struct Biol 19:585–595

Gurel G, Blaha G, Steitz TA, Moore PB (2009) Structures of triacetyloleandomycin and mycalamide A bind to the large ribosomal subunit of Haloarcula marismortui. Antimicrob Agents Chemother 53:5010–5014

Haeberli P, Berger I, Pallan PS, Egli M (2005) Syntheses of 4'-thioribonucleosides and thermodynamic stability and crystal structure of RNA oligomers with incorporated 4'-thiocytosine. Nucleic Acids Res 33:3965–3975

Jenner LB, Demeshkina N, Yusupova G, Yusupov M (2010) Structural aspects of messenger RNA reading frame maintenance by the ribosome. Nat Struct Mol Biol 17:555–560

Jin H, Kelley AC, Loakes D, Ramakrishnan V (2010) Structure of the 70S ribosome bound to release factor 2 and a substrate analog provides insights into catalysis of peptide release. Proc Natl Acad Sci USA 107:8593–8598

Keepers JW, James TL (1984) A theoretical studie of distance determinations from NMR. Two-dimensional nuclear overhauser effect spectra. J Magn Reson 57:404–426

Kuszewski J, Schwieters C, Clore GM (2001) Improving the accuracy of NMR structures of DNA by means of a database potential of mean force describing base–base positional interactions. J Am Chem Soc 123:3903–3918

Lam SL, Chi LM (2010) Use of chemical shifts for structural studies of nucleic acids. Prog Nucl Magn Reson Spectrosc 56:289–310

Lu XJ, Olson WK (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. Nucleic Acids Res 31:5108–5121

Marcheschi RJ, Staple DW, Butcher SE (2007) Programmed ribosomal frameshifting in SIV is induced by a highly STRUCTURED RNA stem–loop. J Mol Biol 373:652–663

McDowell JA, He LY, Chen XY, Turner DH (1997) Investigation of the structural basis for thermodynamic stabilities of tandem GU wobble pairs: NMR structures of (rGGAGUUCC)(2) and (rGGAUGUCC)(2). Biochemistry 36:8030–8038

Mielke SP, Krishnan VV (2009) Characterization of protein secondary structure from NMR chemical shifts. Prog Nucl Magn Reson Spectrosc 54:141–165

Mollova E, Pardi A (2000) NMR Solution structure determination of RNAs. Curr Opin Struct Biol 10:298–302

Mooren MMW, Pulleyblank DE, Wijmenga SS, Vandeven FJM, Hilbers CW (1994) The solution structure of the hairpin formed by d(TCTCTC-TTT-GAGAGA). Biochemistry 33:7315–7325

Musselman C, Pitt SW, Gulati K, Foster LL, Andricioaei I, Al-Hashimi HM (2006) Impact of static and dynamic A-form heterogeneity on the determination of RNA global structure dynamics using NMR residual dipolar couplings. J Biomol NMR 36:235–249

Olson WK, Bansal M, Burley SK, Dickerson RE, Gerstein M, Harvey SC, Heinemann U, Lu XJ, Neidle S, Shakked Z, Sklenar H, Suzuki M, Tung CS, Westhof E, Wolberger C, Berman HM (2001) A standard reference frame for the description of nucleic acid base-pair geometry. J Mol Biol 313:229–237

Olson WK, Esguerra M, Xin Y, Lu X (2009) New information content in RNA base pairing deduced from quantitative analysis of high-resolution structures. Methods 47:177–186

Pardi A, Hare DR, Wang C (1988) Determination of DNA structures by NMR and distance geometry techniques: a computer simulation. Proc Natl Acad Sci USA 85:8785–8789

Petzold K, Duchardt E, Flodell S, Larsson G, Kidd-Ljunggren K, Wijmenga S, Schleucher J (2007) Conserved nucleotides in an RNA essential for hepatitis B virus replication show distinct mobility patterns. Nucleic Acids Res 35:6854–6861

Popenda M, Szachniuk M, Blazewicz M, Wasik S, Burke EK, Blazewicz J, Adamiak RW (2010) RNA FRABASE 2.0: an advanced web-accessible database with the capacity to search the three-dimensional fragments within RNA structures. BMC Bioinformatics 11:231–237

Portmann S, Usman N, Egli M (1995) The crystal-structure of R(Ccccgggg) in 2 distinct lattices. Biochemistry 34:7569–7575

Ramesh A, Wakeman CA, Winkler WC (2011) Insights into metalloregulation by M-box riboswitch RNAs via structural analysis of manganese-bound complexes. J Mol Biol 407: 556–570

Ribas-Prado R, Giessner-Prettre C (1981) Parameters for the calculation of the ring current and atomic magnetic anisotropy contributions to magnetic shielding constants: nucleic acid bases and intercalating agents. J Mol Struct THEOCHEM 76:81–92

Shajani Z, Varani G (2007) NMR studies of dynamics in RNA and DNA by C-13 relaxation. Biopolymers 86:348–359

Shen Y, Lange O, Delagio F, Rossi P, Aramini JM, Liu G, Eletsky A, Wu Y, Singarapu KK, Lemak A, Ignatchenko A, Arrowsmith CH, Szyperski T, Montelione GT, Baker D, Bax A (2008) Consistent blind protein structure generation from NMR chemical shift data. Proc Natl Acad Sci USA 105:4685–4690

Shen Y, Vernon R, Baker D, Bax A (2009) De novo protein structure generation from incomplete chemical shift assignments. J Biomol NMR 43:63–78

Sibille N, Pardi A, Simorre JP, Blackledge M (2001) Refinement of local and long-range structural order in theophylline-binding RNA using C-13-H-1 residual dipolar couplings and restrained molecular dynamics. J Am Chem Soc 123:12135–12146

Staple DW, Butcher SE (2003) Solution structure of the HIV-1 frameshift inducing stem-loop RNA. Nucleic Acids Res 31:4326–4331

Tolbert BS, Miyazaki Y, Barton S, Kinde B, Starck P, Singh R, Bax A, Case DA, Summers MF (2010) Major groove width variations in RNA structures determined by NMR and impact of 13C residual chemical shift anisotropy and 1H–13C residual dipolar coupling on refinement. J Biomol NMR 47:205–219

Tukalo M, Yaremchuk A, Fukunaga R, Yokoyama S, Cusack S (2005) The crystal structure of leucyl-tRNA synthetase complexed with tRNA(Leu) in the post-transfer- editing conformation. Nat Struct Mol Biol 12:923–930

Varani G, Aboul-ela F, Allain FHT (1996) NMR investigation of RNA structure. Prog Nucl Magn Reson Spectrosc 29:51–127

Vicens Q, Westhof E (2002) Crystal structure of a complex between the aminoglycoside tobramycin and an oligonucleotide containing the ribosomal decoding A site. Chem Biol 9:747–755

Wijmenga SS, van Buuren BNM (1998) The use of NMR methods for conformational studies of nucleic acids. Prog Nucl Magn Reson Spectrosc 32:287–387

Wijmenga SS, Heus HA, Werten B, Vandermarel GA, Vanboom JH, Hilbers CW (1994) Assignment strategies and analysis of cross-peak patterns and intensities in the 3-dimensional homonuclear TOCSY-NOESY of RNA. J Magn Reson Ser B 103:134–141

Wijmenga SS, Heus HA, Leeuw HAE, Hoppe H, van der Graaf M, Hilbers CW (1995) Sequential backbone assignment of uniformly C13-labeled RNAs by a 2-dimensional P(CC)H-TOCSY triple-resonance NMR experiment. J Biomol NMR 5:82–86

Wijmenga SS, Kruithof M, Hilbers CW (1997) Analysis of H-1 chemical shifts in DNA: assessment of the reliability of H-1 chemical shift calculations for use in structure refinement. J Biomol NMR 10:337–350

Wishart DS, Arndt D, Berjanskii M, Tang P, Zhou J, Lin G (2008) CS23D: a web server for rapid protein structure generation using NMR chemical shifts and sequence data. Nucleic Acids Res 36:W496–W502