Journal of Biomolecular NMR, **18:** 119–128, 2000.
KLUWER/ESCOM

Refined solution structure of the dimeric N-terminal HHCC domain of HIV-2 integrase

Astrid P.A.M. Eijkelenboom[a,*], Fusinita M.I. van den Ent[b,**], Rainer Wechselberger[a], Ronald H.A. Plasterk[b], Robert Kaptein[a] & Rolf Boelens[a,***]
[a]*Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands;*
[b]*Division of Molecular Biology, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands*

Received 4 May 2000; Accepted 19 July 2000

Key words: helix-turn-helix motif, HIV, integrase, protein dimer, protein structure, zinc-binding domain

Abstract

The solution structure of the dimeric N-terminal domain of HIV-2 integrase (residues 1–55, named IN_{1-55}) has been determined using NMR spectroscopy. The structure of the monomer, which was already reported previously [Eijkelenboom et al. (1997) *Curr. Biol.*, **7**, 739–746], consists of four α-helices and is well defined. Helices α1, α2 and α3 form a three-helix bundle that is stabilized by zinc binding to His12, His16, Cys40 and Cys43. The dimer interface is formed by the N-terminal tail and the first half of helix α3. The orientation of the two monomeric units with respect to each other shows considerable variation. ^{15}N relaxation studies have been used to characterize the nature of the intermonomeric disorder. Comparison of the dimer interface with that of the well-defined dimer interface of HIV-1 IN_{1-55} shows that the latter is stabilized by additional hydrophobic interactions and a potential salt bridge. Similar interactions cannot be formed in HIV-2 IN_{1-55} [Cai et al. (1997) *Nat. Struct. Biol.*, **4**, 567–577], where the corresponding residues are positively charged and neutral ones.

Introduction

The insertion of a DNA-copy (cDNA) of the viral RNA into the genome of the infected cell is essential for replication of the human immunodeficiency virus (HIV). This integration process comprises two reactions that are catalyzed by the viral enzyme integrase (IN): site-specific cleavage of two (occasionally three) nucleotides from each 3′ end of the viral cDNA next to a conserved CA dinucleotide (cleavage or 3′ processing reaction) and insertion of the recessed viral cDNA end into the host DNA (integration or strand transfer reaction; for recent reviews on retroviral integration, see Puras Lutkze and Plasterk, 1998a; Esposito and Craigi, 1999).

HIV IN contains three functional domains (Bushman et al., 1993; Vink et al., 1993): an N-terminal zinc-binding HHCC domain, a central catalytic core and a C-terminal DNA-binding domain. These domains are all required for cleavage and integration (Schauer and Billich, 1992; Vink et al., 1993), but the catalytic core domain alone can catalyze the apparent reversal of the integration reaction, termed disintegration, in vitro (Chow et al., 1992; Bushman et al., 1993; Vink et al., 1993). X-ray structures of the dimeric catalytic core domain of HIV IN (Dyda et al., 1994; Goldgur et al., 1998; Maignan et al., 1998) and Avian Sarcoma Virus IN (Bujacz et al., 1995; Bujacz et al., 1996) have been determined in the presence and absence of metal cofactors.

The C-terminal domain binds DNA in a non-specific manner (Kahn et al., 1991; Vink et al.,

*Present address: Unilever Research Vlaardingen, Olivier van Noortlaan 120, 3133 AT Vlaardingen, The Netherlands.
**Present address: MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB 2 QH, U.K.
***To whom correspondence should be addressed. E-mail: boelens@nmr.chem.uu.nl

1993; Woerner and Marcus Secura, 1993; Engelman et al., 1994; Puras Lutzke et al., 1994). Recent photocrosslinking studies indicate that the C-terminal domain is mainly involved in stabilizing the specific interaction of integrase with the viral DNA ends (Heuer and Brown, 1997; Esposito and Craigie, 1998). Furthermore, this domain also contributes to the functional multimerization of IN (Jenkins et al., 1996; Puras Lutzke and Plasterk, 1998b). The structure of the C-terminal DNA-binding domain has been determined by NMR spectroscopy (Eijkelenboom et al., 1995, 1999; Lodi et al., 1995).

The N-terminal domain contains an HHCC motif that is conserved in all retroviral integrases and retrotransposon integrase proteins (Johnson et al., 1986; Doolittle et al., 1989; Khan et al., 1991). Binding of a single zinc ion to the HHCC motif induces folding in the isolated N-terminal domain of HIV IN (IN_{1-55}; Burke et al., 1992) as well as in full-length HIV IN (Zheng et al., 1996). In full-length IN, the presence of zinc induces tetramerization and octamerization and enhances Mg^{2+} dependent catalytic activity (Lee and Han, 1996; Zheng et al., 1996; Lee et al., 1997). These observations suggest that the N-terminal domain is involved in protein–protein interactions in order to form the active multimer. Recently, it has been shown that a structured N-terminal domain restores IN activity of an N-terminal deletion mutant, which indicates that an interaction between the N-terminal domain and another domain of IN is essential for the formation of the active multimer (van den Ent et al., 1999; Yang et al., 1999). There are indications that IN oligomerizes through interactions between the N-terminal and the core domains (Ellison et al., 1995). Furthermore, experiments with monoclonal antibodies suggest that the N-terminal and C-terminal domains are close together in space (Bizub-Bender et al., 1994).

The N-terminal domain may also play a role in the recognition of the viral DNA ends. Disintegration activities of HHCC mutant proteins on various substrates suggest that the N-terminal domain contributes to the interaction of HIV IN with the viral DNA ends internal to the conserved CA dinucleotide (Vincent et al., 1993). In addition, cross-complementation experiments using the isolated N-terminal domain of HIV or FIV IN and an N-terminal deletion mutant of HIV IN on different substrates indicate that the N-terminal domain supports the recognition of the viral DNA ends by IN (van den Ent et al., 1999). No direct DNA binding activity by the N-terminal domain has been observed thus far (Khan et al., 1991; Mumm and Grandgenett, 1991; Schauer and Billich, 1992; Woerner and Marcus-Secura, 1993), but this domain may facilitate viral DNA recognition indirectly via protein–protein interactions.

The isolated N-terminal domain is a dimer in solution in the presence of zinc (Cai et al., 1997; Eijkelenboom et al., 1997). The structure of IN_{1-55} has been solved by NMR spectroscopy for HIV-1 IN (Cai et al., 1997) and a mutant thereof (Cai et al., 1998), as well as for the monomeric unit of HIV-2 IN (Eijkelenboom et al., 1997). In contrast to HIV-2 IN_{1-55}, the monomer of HIV-1 IN_{1-55} exists in two interconverting folded forms. One of them closely resembles the fold of HIV-2 IN_{1-55}. The structure consists of a three-helix bundle, which is stabilized by zinc binding to the HHCC motif. The arrangement of the helices is similar to that found in several DNA-binding proteins, in which the second and the third helix form a helix–turn–helix motif. The helix that is used by these proteins to bind to DNA is part of the dimer interface of HIV-1 integrase (Cai et al., 1997).

In this study, we characterize the HIV-2 IN_{1-55} dimer in more detail by NMR spectroscopy. To detect intermonomer NOEs, we recorded 2D and 3D isotope filtered NOE experiments on a 1:1 mixture of unlabelled and $^{15}N/^{13}C$-labelled IN_{1-55} in 99.9% D_2O, which allowed for more sensitive measurements as compared to those previously performed in 95% H_2O/5% D_2O (Eijkelenboom et al., 1997). From these experiments, several intersubunit NOEs could be identified, which were used as a starting point for structure calculations of dimeric IN_{1-55}. In the final structures that we present here, the monomers are very well defined. Interestingly, the relative orientation of the two subunits with respect to each other shows considerable variation. ^{15}N relaxation studies indicate that this disorder may be caused by flexibility and conformational averaging, in line with our previous hypothesis (Eijkelenboom et al., 1997).

Materials and methods

IN_{1-55} preparation

The cloning, expression and purification of unlabelled, uniformly ^{15}N-labelled and $^{15}N/^{13}C$-labelled IN_{1-55} protein has been described previously (Eijkelenboom et al., 1997). For the NMR experiments, the resulting lyophilized protein was dissolved in 50 mM deuterated Tris, 150 mM NaCl and 2 mM β-mercaptoethanol with 1.1 equiv $ZnCl_2$, and the pH was adjusted to 6.5. In

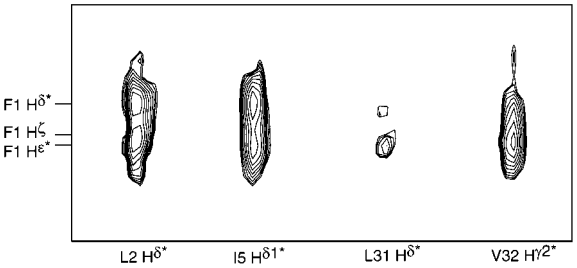

Figure 1. Strips of a NOESY-(^{1}H,^{13}C)-HSQC showing NOEs between ring protons of F1 and methyl groups of the interfacial residues L2, I5, L31 and V32. From these residues, only the methyl groups that have the strongest NOEs to F1 are shown. The NOEs from F1 to L2 are intramonomer, from F1 to I5 and from F1 to V32 ambiguous, and from F1 to L31 intermonomer.

addition to the previously used NMR samples (^{15}N-labelled IN_{1-55} in 95% H_2O/5% D_2O, ^{15}N-labelled IN_{1-55} in 99.99% D_2O, and a 1:1 mixture of unlabelled and ^{15}N/^{13}C-labelled IN_{1-55} in 95% H_2O/5% D_2O), a 1:1 mixture of unlabelled and ^{15}N/^{13}C-labelled IN_{1-55} in 99.99% D_2O was prepared.

NMR spectroscopy and spectral assignments

NMR experiments were performed essentially as described in Cavanagh et al. (1996) on a Bruker AMX-600, a Varian Unity+ 500 and a Varian Unity+ 750 MHz spectrometer, equipped with triple resonance gradient probes, at 300 K. The sequential assignment of ^{1}H and ^{15}N resonances has been presented previously (Eijkelenboom et al., 1997). In addition, ^{13}C resonances were assigned using HNCA, HN(CO)CA, H(C)CH DIPSI and HCC(H) DIPSI spectra. Stereospecific assignments of β-methylene protons (for 11 residues) and χ_1 dihedral angle restraints (for 10 residues) were obtained using 2D NOE and 2D clean-TOCSY spectra recorded with short mixing times (25 ms and 30 ms, respectively) and a 3D HNHB spectrum (Düx et al., 1997). Furthermore, stereospecific assignments of the methyl groups of three valines and their χ_1 dihedral angle restraints were obtained from the 2D NOE spectrum with short mixing time. $^3J_{HNH\alpha}$ coupling constants were derived from a 3D HNHA experiment and converted into ϕ dihedral angle restraints (for 32 residues). NOE distance restraints were obtained from 2D NOE, 3D NOESY-(^{1}H,^{15}N)-HSQC, 3D NOESY-(^{1}H,^{13}C)-HSQC and 3D (^{1}H,^{13}C)-HMQC-NOESY-(^{1}H,^{13}C)-HSQC spectra with mixing times ranging from 50 to 150 ms. In addition, three restraints involving the backbone amide proton of residue N30, which is not visible at pH 6.5, were obtained from a 2D NOE spectrum of IN_{1-55} at pH 5.4. At this pH, the conformation of folded IN_{1-55} is the same as at pH 6.5, but since IN_{1-55} loses zinc below pH 5 (Eijkelenboom et al., 1997), a higher pH was preferred for the structural studies. To distinguish between intra- and intermonomer NOEs, ^{13}C-filtered 2D and 3D NOE experiments were performed on the 1:1 mixture of unlabelled and ^{15}N/^{13}C-labelled IN_{1-55} in 99.99% D_2O as described by Burgering et al. (1993), Folmer et al. (1995) and Zwahlen et al. (1997). ^{15}N T_1, ^{15}N $T_{1\rho}$ and ^{15}N{^{1}H} NOE relaxation experiments were recorded to study the dynamic behavior of IN_{1-55}.

Structure calculations and analysis

Approximate NOE distance restraints were obtained from the above mentioned 2D and 3D NOE spectra. Where possible, the distance restraints were derived from the 2D NOE spectra. In addition, the 3D NOE spectra were used to resolve overlap. The restraints were classified as strong, medium, weak and very weak, and the corresponding upper bounds were set to 2.9 Å, 3.6 Å, 5.0 Å and 6.0 Å, respectively. The lower distance bounds were all effectively 1.8 Å. In addition, ϕ and χ_1 dihedral angle restraints were used as experimental input in the structure calculations. The structures were calculated with X-PLOR version 3.851 (Brünger, 1992) using a dynamical simulated annealing protocol starting from randomized coordinates (Nilges, 1988), in which two extra terms were introduced to deal with the symmetric dimer (Nilges, 1993). Non-crystallographic symmetry (NCS) restraints were added to keep the monomers superimposable, and a global symmetry potential was included to keep the dimer symmetric (Nilges, 1993). Sum averaging (Nilges, 1993) was used to correct for multiple atom selections and to include distance restraints resulting from ambiguous NOE assignments. Distance restraints containing non-stereospecifically assigned diastereotopic groups were corrected as described by Fletcher et al. (1996). Furthermore, a set of restraints to provide the tetrahedral coordination geometry of the zinc ion was included in the structure calculations (Eijkelenboom et al., 1997). During the final refinement step the following force constants were used in the X-PLOR target function for minimization: 1000 kcal mol^{-1}Å^{-2} for bond lengths, 500 kcal mol^{-1}rad^{-2} for bond angles and impropers, 100 kcal mol^{-1}Å^{-2} for NCS restraints, 200 kcal mol^{-1}rad^{-2} for dihedral angle restraints, 50 kcal mol^{-1}Å^{-2} for NOE distance restraints, 1 kcal mol^{-1}Å^{-2} for sym-

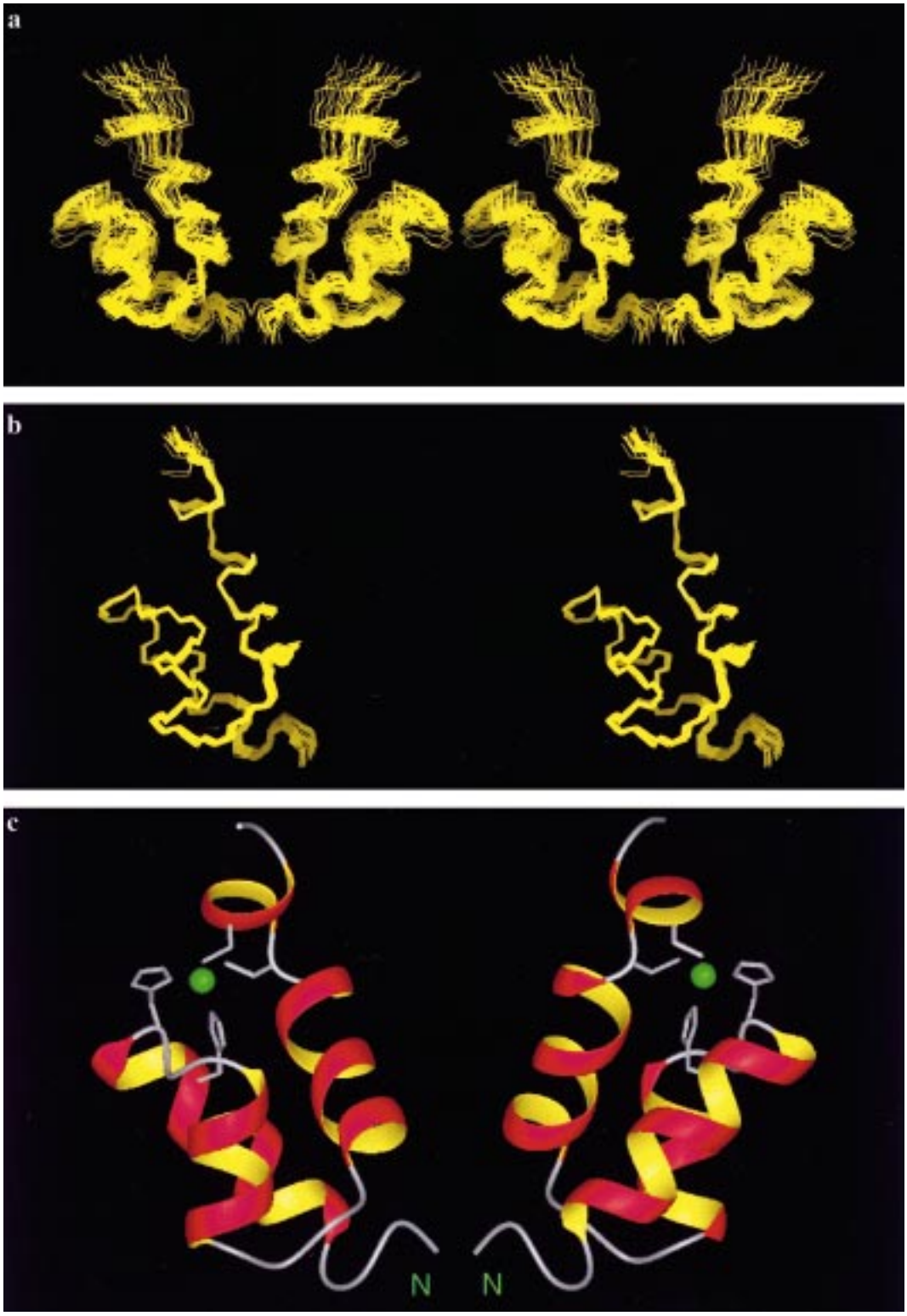

Figure 2. Structure of the N-terminal domain of HIV-2 integrase (residues 1–46). (a) Superposition of backbone traces of the dimer. (b) Superposition of backbone traces of the monomer only. (c) Ribbon diagram of the dimer structure that is closest to the average. The zinc-coordinating residues and the zinc ion are indicated in black.

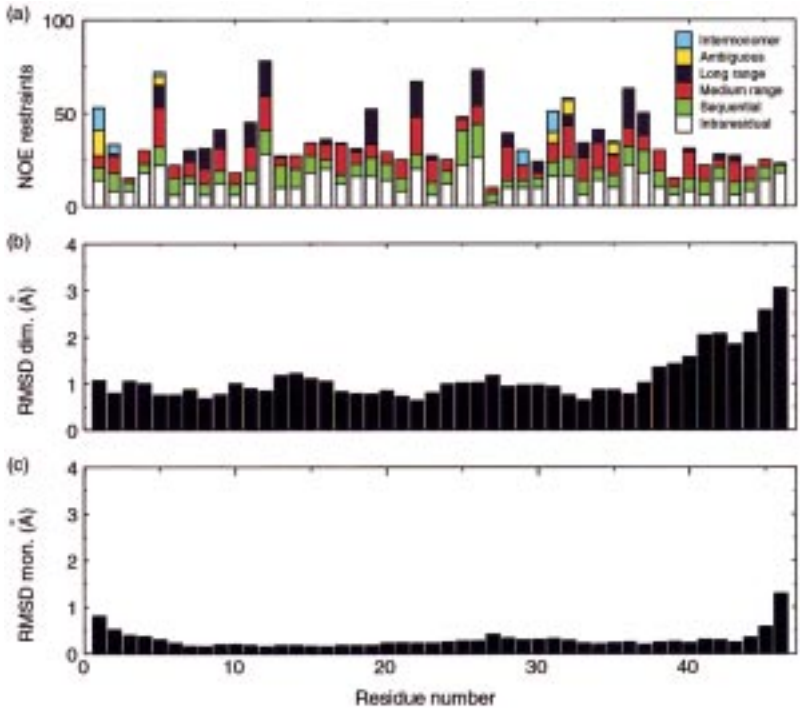

Figure 3. Overview of structural parameters. (a) Number of distance restraints per residue. The restraints are classified as intraresidual, sequential, medium range, long range, ambiguous and intermonomer. (b, c) Root Mean Square Deviation (RMSD) for the backbone N, C^{α}, C' atoms. The structures are superimposed on the backbone atoms of residues 1–45 of the dimer, and of the monomer, respectively.

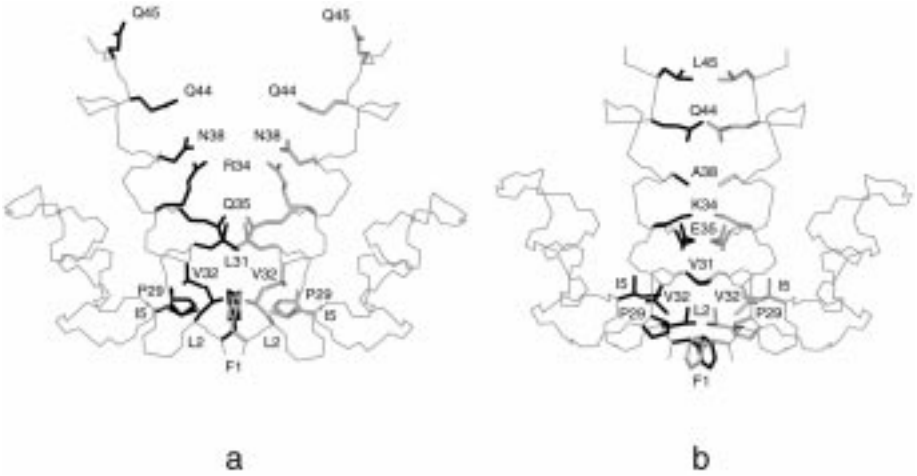

Figure 4. Comparison of HIV-2 and HIV-1 IN_{1-55} (residues 1–46). (a) HIV-2 IN_{1-55} dimer. Shown are the side chains that are part of the dimer interface (F1, L2, I5, P29, L31, V32 and Q35), as well as residues that correspond to those that are part of the dimer interface of HIV-1 IN_{1-55}. (b) HIV-1 IN_{1-55} dimer. Shown are the side chains that are part of the dimer interface, except for P30 (see text).

metry table distance restraints and 4 kcal $mol^{-1}Å^{-4}$ for the van der Waals repulsion term.

The calculated structures were analyzed using PROCHECK-NMR (Laskowski et al., 1996), and secondary structure elements were identified according to the DSSP algorithm (Kabsch and Sander, 1983). The program MOLMOL (Koradi et al., 1996) was used to produce Figures 2 and 4. The coordinates of the final 32 structures, the restraint data and the chemical shifts have been submitted to the RCSB Protein Databank, pdb accession code 1e0e.

Results and discussion

Structure determination

Our previously determined structure of the monomeric unit of IN_{1-55} was solved using 2D NOE and 3D ^{15}N-separated NOE experiments. For the refinement of the structure we have now also used 3D ^{13}C-separated NOE experiments. Furthermore, we recorded 2D and 3D ^{13}C-filtered NOE experiments on an equimolar mixture of unlabelled and ^{15}N/^{13}C-labelled IN_{1-55} in 99.99% D_2O. From these spectra, several intersubunit NOEs could clearly be identified, involving the residues F1, L31, V32 and Q35. These NOEs were used as a starting point for the structure calculations of the dimer. During the refinement procedure, more intermonomer and co-monomer (which have both an inter- and intramonomer contribution) NOEs could be identified. The use of 3D ^{13}C-separated NOE experiments turned out to be essential for the structure determination of the dimer, since methyl proton resonances of the interfacial residues L2, I5, L31 and V32 overlap. Distance restraints between methyl groups of these residues and for instance the aromatic ring protons of F1 could therefore only be assigned using the (^{13}C-filtered) 3D NOESY-(^{1}H,^{13}C)-HSQC spectra. Several cross peaks between the ring protons of F1 and methyl groups of interfacial residues are shown in Figure 1. In addition, distance restraints between the overlapping methyl groups could only be assigned from a 3D (^{1}H,^{13}C)-HMQC-NOESY-HSQC spectrum, e.g. from L2 to I5 (intramonomer) or from I5 to L31 (intermonomer).

The final structure calculations were performed using 1658 distance restraints (40 of which were intermonomer, and 38 ambiguous with respect to intra- versus intermonomer), 64 ϕ and 26 χ_1 dihedral angle restraints. As before, the largely disordered residues 47–55, for which no medium or long range NOEs were observed, were not included in the calculations. Fifty structures were calculated, of which 32 were selected on the basis of low overall energy. A superposition of the 32 final structures is shown in Figure 2a, a summary of the structural statistics is given in Table 1, and an overview of structural parameters is shown in Figure 3. As can be seen from Figure 2a, the structure of the dimer shows considerable disorder. This is caused by a loose orientation of the two subunits with respect to each other, since a superposition of the 32 monomers of one subunit shows that the monomers themselves are well defined (see Figure 2b). The backbone RMSD versus the average for the dimer is 1.2 Å for the well-defined residues 1–45 and 0.30 Å when comparing the monomers only. The precision of the monomers is comparable with our previously determined structure of IN_{1-55}, but the quality has improved significantly. The structures show smaller deviations from idealized covalent geometry and the percentage of residues in the most favorable region of the Ramachandran plot increased from 78% to 90%.

Description of the structure

Figure 2c shows the folding topology of the structure that is closest to the average of the IN_{1-55} dimer. The structure of the monomer consists of four α-helices, which comprise residues 5–15 (α1), 19–25 (α2), 30–39 (α3) and 41–44 (α4) in 97% of the structures. In 47% of the structures helix α1 ranges from residue 2–15. Helices α1, α2 and α3 form a three-helix bundle that is stabilized by the zinc-binding unit, which is formed by His12, His16, Cys40 and Cys43. His12 coordinates the zinc ion with its $N^{\varepsilon 2}$ atom, His16 with $N^{\delta 1}$, and Cys40 and Cys43 coordinate zinc via their sulfur atoms (Eijkelenboom et al., 1997). The dimer interface is formed by the N-terminal tail and part of helix α3, which are positioned in a parallel manner with respect to each other. The interface comprises the hydrophobic residues F1, L2, I5, P29, L31 and V32, and the hydrophilic residue Q35 (see Figure 4a). The interface is stabilized in particular by hydrophobic interactions between the side chains of residues F1 of one subunit and F1′, L31′ and V32′ of the other subunit (interproton distances up to 3 Å in at least 50% of the structures). Additional hydrophobic contacts (interproton distances up to 5 Å) are mainly observed between F1 and I5′ and P29′, L2 and P29′ and L31′, I5 and L31′, L31 and Q35′, and Q35 and Q35′. In 53% of the structures, one or more atoms of the side chain carboxyamide groups of Q35 of the two subunits are close together in space (within 3 Å),

Table 1. Structural statistics for HIV-2 IN^{a}_{1-55}

Type and number of restraints per dimer	
Distance restraints	1658
Intramonomer	1580
Intraresidual	586
Sequential	362
Medium range	418
Long range	214
Intermonomer	40
Ambiguous	38
Dihedral restraints	90
ϕ	64
χ_1	26
Maximum experimental violations	
Distance restraints (Å)	0.20
Dihedral restraints (deg)	1.9
RMSD from average structure (Å)[b]	
Backbone (N,C^{α},C′)	1.19 ± 0.56 (0.30 ± 0.07)
All heavy atoms	1.49 ± 0.50 (0.80 ± 0.06)
Deviation from experimental restraints	
Distance restraints (Å)	0.007 ± 0.001
Dihedral restraints (deg)	0.18 ± 0.07
Deviation from idealized covalent geometry	
Bonds (Å)	0.0017 ± 0.0001
Angles (deg)	0.39 ± 0.01
Impropers (deg)	0.27 ± 0.02
Percentage of residues with ϕ/ψ[c] *in*	
Most favoured regions	90.3
Additionally allowed regions	9.6
Generously allowed regions	0.1
Disallowed regions	0.0
Average number of bad contacts per 100 residues	0.61

[a] Given for final 32 structures for residues 1–46.
[b] For residues 1–45. RMSD for the monomer is given in parentheses.
[c] Excluding glycine and proline residues.

suggesting that they could potentially form a hydrogen bond, as has been observed for the carboxyamide groups of an asparagine residue in the dimer interface of leucine zippers (see Junius et al., 1995). In these structures, the conformation of the Asn side chains is asymmetric: the NH_2 group of one carboxyamide is hydrogen bonded to the CO group of the other carboxyamide. A 180° flip around the C^{β}-C^{γ} bond of both Asn residues reverses the conformation; the CO group of the first carboxyamide is then hydrogen bonded to the NH_2 group of the second. On the basis of NMR studies, Junius et al. (1995) proposed a model in which rapid 180° flipping around the C^{β}-C^{γ} bond of the Asn side chain continuously exchanges the hydrogen bonding conformations in the dimer interface of Jun-c. The asymmetric carboxyamide conformation cannot be obtained from our structure calculations, since we used NCS restraints to keep the monomers superimposable.

Comparison with the structure of HIV-1 IN_{1-55}

HIV-2 and HIV-1 IN_{1-55} have 56% sequence identity, and it has already been mentioned previously that the structures of the monomers of HIV-2 IN_{1-55} (Eijkelenboom et al., 1997) and one of the two interconverting forms of HIV-1 IN_{1-55} (Cai et al., 1997)

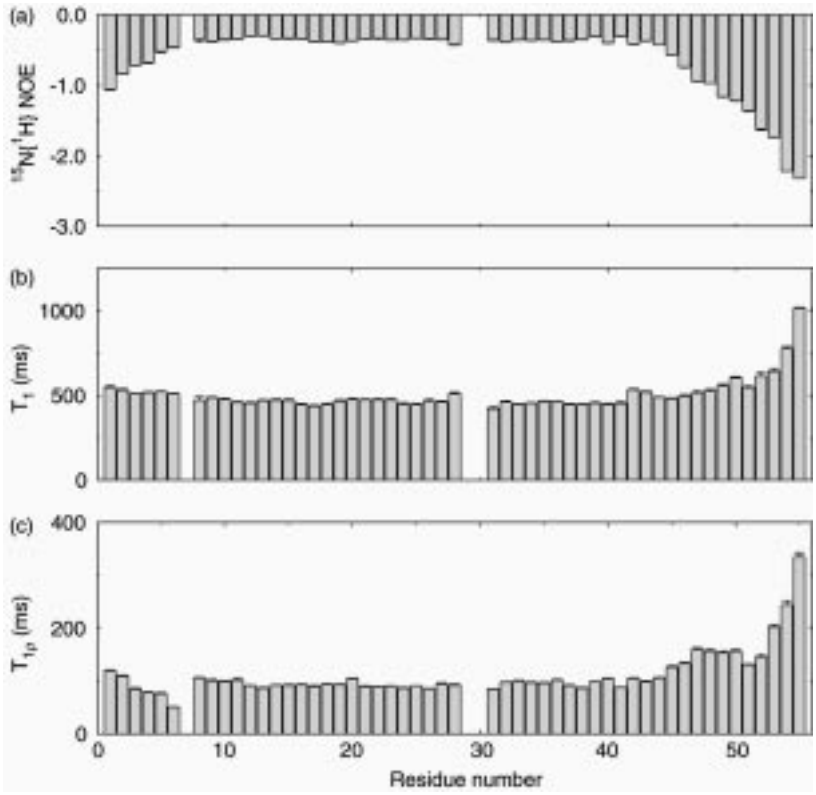

Figure 5. ^{15}N backbone relaxation data of IN_{1-55}. (a) Heteronuclear $^{15}N\{^{1}H\}$ NOEs. (b) Longitudinal relaxation time constants T_1. (c) Rotating frame relaxation time constants $T_{1\rho}$.

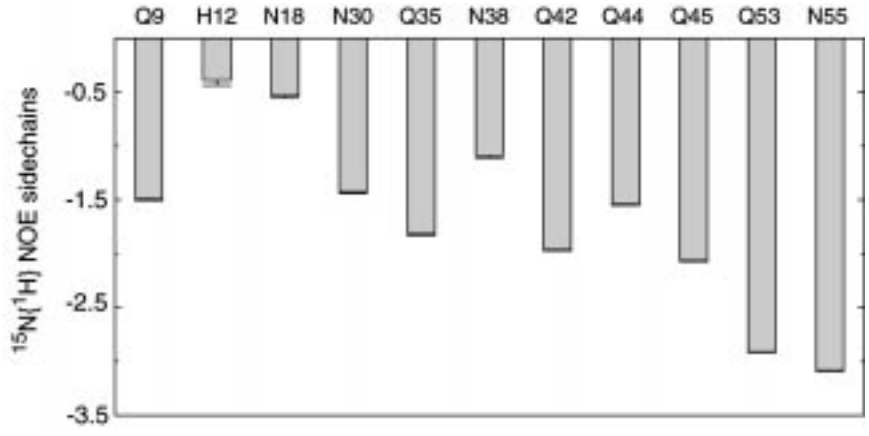

Figure 6. Side chain heteronuclear $^{15}N\{^{1}H\}$ NOEs of $N^{\delta 2}$ of asparagines, $N^{\varepsilon 2}$ of glutamines and $N^{\delta 1}$ of histidine 12.

are very similar. A superposition of the monomer structures that are closest to the average for HIV-2 and HIV-1 IN_{1-55} results in a pairwise backbone RMSD of 0.80 Å for residues 1–45. However, there are significant differences between the structures regarding the orientation of the monomers within the dimer, resulting in a pairwise backbone RMSD of 2.8 Å. Figure 4a shows the HIV-2 IN_{1-55} dimer, in which the side chains that form the dimer interface of HIV-2 IN_{1-55} and the corresponding side chains that are part of the dimer interface of HIV-1 IN_{1-55} are depicted. Figure 4b shows the dimer interface of HIV-1

IN_{1-55}. It can be seen from the figure that the hydrophobic residues that stabilize the interface of HIV-2 IN_{1-55} are also part of the dimer interface of HIV-1 IN_{1-55}: F1, L2, I5, P29, L31 and V32 (of these residues, L31 of HIV-2 IN_{1-55} corresponds to V31 in HIV-1 IN_{1-55}). Note that the orientation of the aromatic ring of F1 is different in both interfaces; it is buried more in the dimer interface of HIV-2 IN_{1-55}. Furthermore, in HIV-1 IN_{1-55} residue P30 is part of the dimer interface, whereas the side chain of residue N30 of HIV-2 IN_{1-55} points away from the interface (these residues are not shown in Figure 4). The major difference between the two dimer interfaces is the fact that the upper halves of helix α3 and helix α4 are also part of the interface in HIV-1 IN_{1-55}. The residues in this part of the interface are K34, A38, Q44 and L45. In HIV-2 IN_{1-55}, the corresponding residues are R34, N38, Q44 and Q45. Note that the hydrophobic residues in this part of the interface of HIV-1 IN_{1-55} are hydrophilic in HIV-2 IN_{1-55}. Furthermore, the potential salt bridge in the HIV-1 IN_{1-55} dimer interface between K34 and E35 can not be formed in HIV-2 IN_{1-55}, where the corresponding residues are an arginine and a glutamine, respectively. In our isotope filtered experiments, both in 95% H_2O/5% D_2O and 99.9% D_2O, we found no evidence that R34, N38, Q44 or Q45 are involved in dimerization. However, it may be possible that these residues are part of a flexible dimer interface where hydrogen bonds between the side chains of these residues are formed transiently.

Dynamic processes near the dimer interface

In order to test whether the observed disorder near the dimer interface correlates with dynamic processes, we performed a set of ^{15}N relaxation experiments. The results for the backbone dynamics are shown in Figure 5. From the heteronuclear NOE data it can be seen that significant flexibility of the backbone on the nanosecond to picosecond time scale is limited to the termini of the protein. The C-terminal tail, which is largely disordered due to the lack of medium and long range NOEs and was excluded from the structure calculations, forms – as expected – the most flexible part of the protein. Furthermore, the heteronuclear NOE experiment shows that the first six residues, which comprise the interfacial residues F1, L2 and I5, are more flexible on the nanosecond to picosecond time scale than the core of the protein. Interestingly, the backbone amide nitrogen of the interfacial residue F1 (which is preceded by three amino acids in our peptide construct) shows the largest flexibility, whereas its side chain is involved in most of the intermonomer interactions (see above). In addition, the $T_{1\rho}$ data indicate that the N-terminal tail is involved in conformational exchange processes on the millisecond to microsecond time scale, since the $T_{1\rho}$ values for residues 3–6 drop below those of the core of the protein. Figure 6 shows the heteronuclear NOE data from the side chain nitrogens. From the figure it can be seen that the side chain of the interfacial residue Q35 shows considerable flexibility. Taken together, these results suggest that the observed disorder is due to dynamic processes.

Since we suggested in the preceding section that N38, Q44 and Q45 could perhaps transiently be involved in hydrogen bonds across the dimer interface, it is interesting to take a look at the side chain dynamics of these residues as well in Figure 6 (note that the side chain nitrogens of residue R34, which was also mentioned as a possible candidate for momentary dimerization, were not identified in our spectra). First, it should be mentioned that, apart from His12 $N^{\delta 1}$, all side chain nitrogens show considerably larger flexibility than the backbone nitrogen atoms of the core of the protein at the nanosecond to picosecond time scale. Comparison among the carboxyamide nitrogens of the asparagines suggests that the side chains of N18 and – to a lesser extent – N38 are stabilized.

Conclusions

The solution structure of the dimeric N-terminal domain of HIV-2 integrase shows considerable variation in the orientation of the two monomers with respect to each other. The structure of the monomeric unit, which is well defined, contains four α-helices comprising residues 5–15 (α1), 19–25 (α2), 31–39 (α3) and 41–44 (α4). Helices α1, α2 and α3 form a three-helix bundle that is stabilized by zinc binding to the conserved HHCC motif. The interface is mainly hydrophobic and is formed by residues of the N-terminal tail and in the first half of helix α3. ^{15}N relaxation studies indicate that the observed intermonomer disorder can be accounted for by dynamic processes on the nanosecond to picosecond time scale, as well as conformational averaging on the microsecond to millisecond time scale. In HIV-1 IN_{1-55}, also the second halves of helix α3 and helix α4 are part of the dimer interface. This additional part of the dimer interface of HIV-1 IN_{1-55} is mainly stabilized by hydrophobic interactions and a potential salt bridge. These interactions cannot be formed in HIV-2 IN_{1-55}, as the

corresponding residues are glutamines, an asparagine and an arginine. Given the observed flexibility near the dimer interface, it may be possible that these residues interact transiently with each other, during which hydrogen bonds across the dimer interface could be formed.

Acknowledgements

We thank M. Czich for technical assistence and K. Hård and A. Bonvin for many helpful discussions. A.E. was supported by the Netherlands Organization for Scientific Research (NWO-CW). The 750 MHz spectra were recorded at the SON NMR large scale facility (Utrecht), which is supported by the Large Scale Facility program of the European Union (Contract ERBFMGECT950032).

References

Bizub-Bender, D., Kulkosky, J. and Skalka, A.M. (1994) *AIDS Res. Hum. Retroviruses*, **10**, 1105–1115.

Brünger, A.T. (1992) *X-PLOR, version 3.1: A system for X-ray crystallography and NMR*, Yale University Press, New Haven, CT.

Bujacz, G., Jaskolski, M., Alexandratos, J., Wlodawer, A., Merkel, G., Katz, R.A. and Skalka, A.M. (1995) *J. Mol. Biol.*, **253**, 333–346.

Bujacz, G., Jaskolski, M., Alexandratos, J., Wlodawer, A., Merkel, G., Katz, R.A. and Skalka, A.M. (1996) *Structure*, **4**, 89–96.

Burgering, M.J.M., Boelens, R., Caffrey, M., Breg, J.N. and Kaptein, R. (1993) *FEBS Lett.*, **330**, 105–109.

Burke, C.J., Sanyal, G., Bruner, M.W., Ryan, J.A., LaFemina, R.L., Robbins, H.L., Zeft, A.S., Middaugh, C.R. and Cordingley, M.G. (1992) *J. Biol. Chem.*, **267**, 9639–9644.

Bushman, F.D., Engelman, A., Palmer, I., Wingfield, P. and Craigie, R. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 3428–3432.

Cai, M., Huang, Y., Caffrey, M., Zheng, R., Craigie, R., Clore, G.M. and Gronenborn, A.M. (1998) *Protein Sci.*, **7**, 2669–2674.

Cai, M., Zheng, R., Caffrey, M., Craigie, R., Clore, G.M. and Gronenborn, A.M. (1997) *Nat. Struct. Biol.*, **4**, 567–577.

Cavanagh, J., Fairbrother, W.J., Palmer, A.G. and Skelton, N.J. (1996) *Protein NMR Spectroscopy. Principles and Practice*, Academic Press, Inc., San Diego, CA.

Chow, S.A., Vincent, K.A., Ellison, V. and Brown, P.O. (1992) *Science*, **255**, 723–726.

Doolittle, R.F., Feng, D.-F., Johnson, M.S. and McClure, M.A. (1989) *Quart. Rev. Biol.*, **64**, 1–30.

Düx, P., Whitehead, B., Boelens, R., Kaptein, R. and Vuister, G.W. (1997) *J. Biomol. NMR*, **10**, 301–306.

Dyda, F., Hickman, A.B., Jenkins, T.M., Engelman, A., Craigie, R. and Davies, D.R. (1994) *Science*, **266**, 1981–1986.

Eijkelenboom, A.P.A.M., Puras Lutzke, R.A., Boelens, R., Plasterk, R.H.A., Kaptein, R. and Hård, K. (1995) *Nat. Struct. Biol.*, **2**, 807–810.

Eijkelenboom, A.P.A.M., Sprangers, R., Hård, K., Puras Lutzke, R.A., Plasterk, R.H.A., Boelens, R. and Kaptein, R. (1999) *Proteins Struct. Funct. Genet.*, **36**, 556–564.

Eijkelenboom, A.P.A.M., van den Ent, F.M.I., Vos, A., Doreleijers, J.F., Hård, K., Tullius, T.D., Plasterk, R.H.A., Kaptein, R. and Boelens, R. (1997) *Curr. Biol.*, **7**, 739–746.

Ellison, V., Gerton, J., Vincent, K.A. and Brown, P.O. (1995) *J. Biol. Chem.*, **270**, 3320–3326.

Engelman, A., Hickman, A.B. and Craigie, R. (1994) *J. Virol.*, **68**, 5911–5917.

Esposito, D. and Craigie, R. (1998) *EMBO J.*, **17**, 5832–5843.

Esposito, D. and Craigie, R. (1999) *Adv. Virus Res.*, **52**, 319–333.

Fletcher, C.M., Jones, D.N., Diamond, R. and Neuhaus, D. (1996) *J. Biomol. NMR*, **8**, 292–310.

Folmer, R.H.A., Hilbers, C.W., Konings, R.N.H. and Hallenga, K. (1995) *J. Biomol. NMR*, **5**, 427–432.

Goldgur, Y., Dyda, F., Hickman, A.B., Jenkins, T.M., Craigie, R. and Davies, D.R. (1998) *Proc. Natl. Acad. Sci. USA*, **95**, 9150–9154.

Jenkins, T.M., Engelman, A., Ghirlando, R. and Craigie, R. (1996) *J. Biol. Chem.*, **271**, 7712–7718.

Johnson, M.S., McClure, M.A., Feng, D.F., Gray, J. and Doolittle, R.F. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 7648–7652.

Junius, F.K., Mackay, J.P., Bubb, W.A., Jensen, S.A., Weiss, A.S. and King, G.F. (1995) *Biochemistry*, **34**, 6164–6174.

Kabsch, W. and Sander, C. (1983) *Biopolymers*, **22**, 2577–2637.

Khan, E., Mack, J.P., Katz, R.A., Kulkosky, J. and Skalka, A.M. (1991) *Nucleic Acids Res.*, **19**, 851–860.

Koradi, R., Billeter, M. and Wüthrich, K. (1996) *J. Mol. Graph.*, **14**, 51–55.

Laskowski, R.A., Rullman, J.A., MacArthur, M.W., Kaptein, R. and Thornton, J.M. (1996) *J. Biomol. NMR*, **8**, 477–486.

Lee, S.P. and Han, M.K. (1996) *Biochemistry*, **35**, 3837–3844.

Lee, S.P., Xiao, J., Knutson, J.R., Lewis, M.S. and Han, M.K. (1997) *Biochemistry*, **36**, 173–180.

Lodi, P.J., Ernst, J.A., Kuszewski, J., Hickman, A.B., Engelman, A., Craigie, R., Clore, G.M. and Gronenborn, A.M. (1995) *Biochemistry*, **34**, 9826–9833.

Maignan, S., Guilloteau, J.-P., Zhou-Liu, Q., Clement-Mella, C. and Mikol, V. (1998) *J. Mol. Biol.*, **282**, 359–368.

Mumm, S.R. and Grandgenett, D.P. (1991) *J. Virol.*, **65**, 1160–1167.

Nilges, M., Clore, G.M. and Gronenborn, A.M. (1988) *FEBS Lett.*, **239**, 129–136.

Nilges, M. (1993) *Proteins*, **17**, 297–309.

Puras Lutzke, R.A. and Plasterk, R.H.A. (1998a) *Genes Function*, **1**, 289–307.

Puras Lutzke, R.A. and Plasterk, R.H.A. (1998b) *J. Virol.*, **72**, 4841–4848.

Puras Lutzke, R.A., Vink, C. and Plasterk, R.H.A. (1994) *Nucleic Acids Res.*, **22**, 4125–4131.

Schauer, M. and Billich, A. (1992) *Biochem. Biophys. Res. Commun.*, **185**, 874–880.

van den Ent, F.M.I., Vos, A. and Plasterk, R.H.A. (1999) *J. Virol.*, **73**, 3176–3183.

Vincent, K.A., Ellison, V., Chow, S.A. and Brown, P.O. (1993) *J. Virol.*, **67**, 425–437.

Vink, C., Oude Groeneger, A.A.M. and Plasterk, R.H.A. (1993) *Nucleic Acids Res.*, **21**, 1419–1425.

Woerner, A.M. and Marcus-Sekura, C.J. (1993) *Nucleic Acids Res.*, **21**, 3507–3511.

Yang, F., Leon, O., Greenfield, N.J. and Roth, M.J. (1999) *J. Virol.*, **73**, 1809–1817.

Zheng, R., Jenkins, T.M. and Craigie, R. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 13659–13664.

Zwahlen, C., Legault, P., Vincent, S.J.F., Greenblatt, J., Konrat, R. and Kay, L.E. (1997) *J. Am. Chem. Soc.*, **119**, 6711–6721.